

## Research



**Cite this article:** Childs ML *et al.* 2021 The impact of long-term non-pharmaceutical interventions on COVID-19 epidemic dynamics and control: the value and limitations of early models. *Proc. R. Soc. B* **288**: 20210811. <https://doi.org/10.1098/rspb.2021.0811>

Received: 6 April 2021

Accepted: 26 July 2021

**Subject Category:**

Ecology

**Subject Areas:**

ecology, health and disease and epidemiology, computational biology

**Keywords:**

SARS-CoV-2, non-pharmaceutical interventions, mathematical models, epidemics, reproduction number

**Authors for correspondence:**

Marissa L. Childs

e-mail: [marissac@stanford.edu](mailto:marissac@stanford.edu)

Morgan P. Kain

e-mail: [morganpkain@gmail.com](mailto:morganpkain@gmail.com)

Mallory J. Harris

e-mail: [mharris9@stanford.edu](mailto:mharris9@stanford.edu)

†Denotes equal authorship.

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5545379>.

# The impact of long-term non-pharmaceutical interventions on COVID-19 epidemic dynamics and control: the value and limitations of early models

Marissa L. Childs<sup>1,†</sup>, Morgan P. Kain<sup>2,3,†</sup>, Mallory J. Harris<sup>2,†</sup>, Devin Kirk<sup>2,6</sup>, Lisa Couper<sup>2</sup>, Nicole Nova<sup>2</sup>, Isabel Delwel<sup>4</sup>, Jacob Ritchie<sup>5</sup>, Alexander D. Becker<sup>2</sup> and Erin A. Mordecai<sup>2</sup>

<sup>1</sup>Emmett Interdisciplinary Program in Environment and Resources, <sup>2</sup>Department of Biology, <sup>3</sup>Natural Capital Project, Woods Institute for the Environment, <sup>4</sup>Department of Microbiology and Immunology, and <sup>5</sup>Department of Computer Science, Stanford University, Stanford, CA 94305, USA

<sup>6</sup>Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4

**ORCID** MLC, 0000-0002-8597-2161; MPK, 0000-0003-0605-7289; MJH, 0000-0001-8196-0430; DK, 0000-0001-9588-1004; NN, 0000-0001-8585-1215; ID, 0000-0001-6671-3669; EAM, 0000-0002-4402-5547

Mathematical models of epidemics are important tools for predicting epidemic dynamics and evaluating interventions. Yet, because early models are built on limited information, it is unclear how long they will accurately capture epidemic dynamics. Using a stochastic SEIR model of COVID-19 fitted to reported deaths, we estimated transmission parameters at different time points during the first wave of the epidemic (March–June, 2020) in Santa Clara County, California. Although our estimated basic reproduction number ( $\mathcal{R}_0$ ) remained stable from early April to late June (with an overall median of 3.76), our estimated effective reproduction number ( $\mathcal{R}_E$ ) varied from 0.18 to 1.02 in April before stabilizing at 0.64 on 27 May. Between 22 April and 27 May, our model accurately predicted dynamics through June; however, the model did not predict rising summer cases after shelter-in-place orders were relaxed in June, which, in early July, was reflected in cases but not yet in deaths. While models are critical for informing intervention policy early in an epidemic, their performance will be limited as epidemic dynamics evolve. This paper is one of the first to evaluate the accuracy of an early epidemiological compartment model over time to understand the value and limitations of models during unfolding epidemics.

## 1. Introduction

COVID-19, caused by the emerging virus SARS-CoV-2, rapidly expanded across the globe, overwhelmed healthcare systems, and has led to just under 4.0 million deaths with the pandemic still underway as of July 2021 [1]. It is just the latest and most widespread in a series of (re)emerging and expanding infectious disease outbreaks, including SARS-CoV in 2003, H1N1 influenza virus in 2009, Ebola virus in 2014 and Zika virus in 2016. Before effective vaccines and specific drug therapies are available at the start of emerging epidemics, non-pharmaceutical interventions such as social distancing, mask-wearing, diagnostic and serological testing, contact tracing, and quarantine are the best available tools to slow epidemics and to mitigate their health toll. Early in the COVID-19 epidemic, when epidemiological information was limited, governments and other decision-makers used models (e.g. [2–5]) to predict the spread of COVID-19 under various non-pharmaceutical interventions and to show the benefits of social distancing for reducing and delaying the

epidemic peak (i.e. flattening the curve) in an effort to prevent medical systems from becoming overwhelmed and to buy time for more effective treatments, testing capacity and potential vaccines to become available.

Early models can and should inform policy decisions in an epidemic as they are our primary tool for synthesizing early knowledge of transmission in order to define a plausible range of epidemic outcomes [6]. Such models also quantify trade-offs among proposed intervention scenarios to identify which responses will most efficiently slow exponential growth. During the 2014 Ebola epidemic, for example, early models promoted the use of contact tracing and sanitary burials to reduce transmission [7]. At the beginning of the COVID-19 pandemic, models identified the critical importance of social distancing to slow viral spread [2,8,9]. Early models also serve to illuminate difficult-to-observe processes and to test the implications of new information; for example, in February 2020, Hellwell *et al.* [3] showed that because of SARS-CoV-2 presymptomatic and asymptomatic transmission, contact tracing would be insufficient to curb the spread of the disease and strong social distancing would also be necessary.

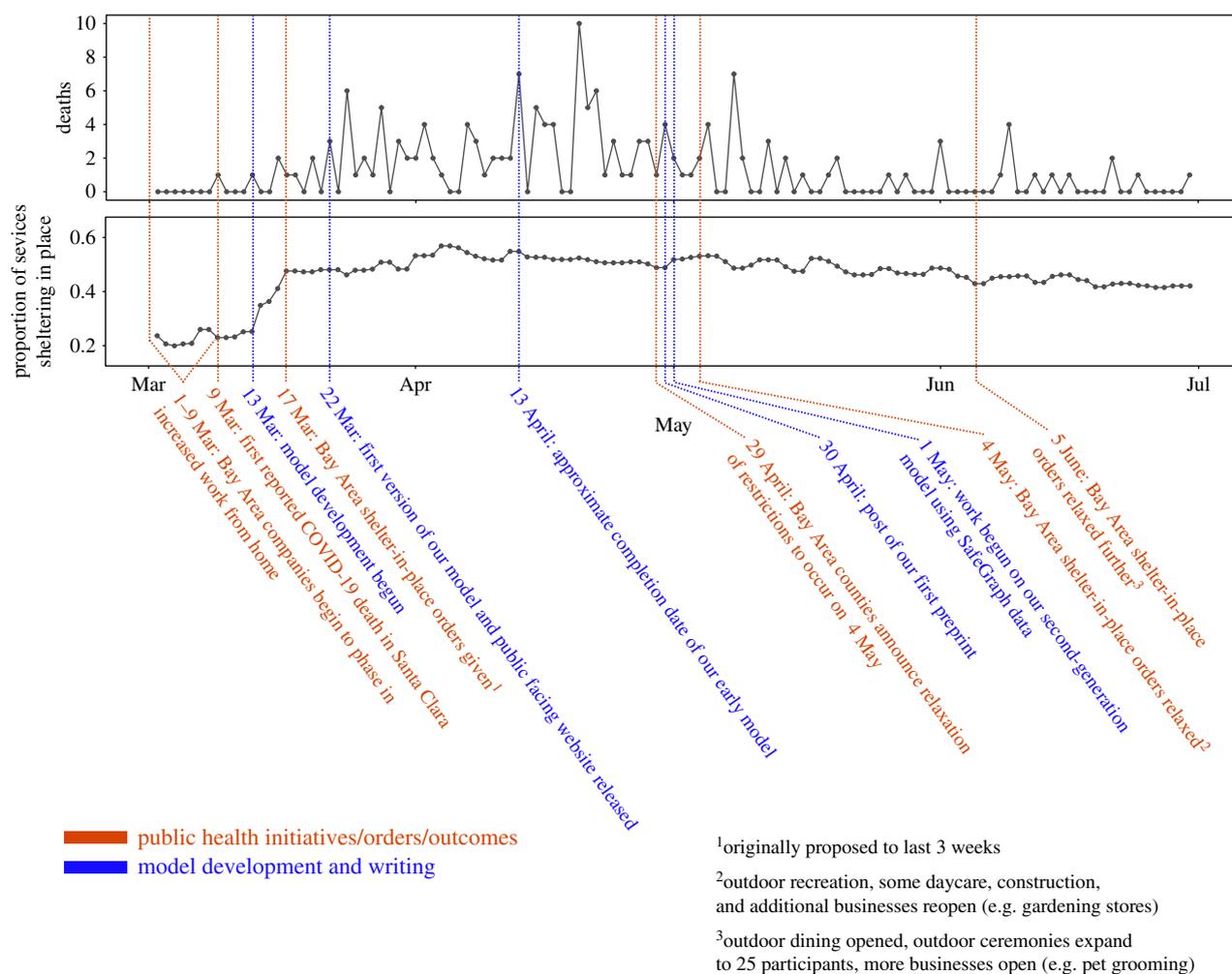
Early models must be built rapidly and calibrated to data that are incomplete and of unknown quality, therefore it can be difficult to appropriately quantify uncertainty and to assess model accuracy in order to compare policy decisions. With ample time, alternative transmission scenarios can be thoughtfully compared and uncertainty well characterized [10]; however, the need for rapid decisions in the face of exponential epidemic growth makes such efforts infeasible for most early models. Though such real-time model assessment is rare, post-epidemic retrospective analyses of SARS [11–13], H1N1 influenza [14–21], Ebola [22–28] and Zika [29–33] have illustrated that much can be learned from emerging epidemics about the fundamental principles of disease transmission and epidemic modelling (e.g. the limits and utility of model complexity [15,23,32], the effects of population and geographic heterogeneity on disease dynamics [20,21,27], and the importance of stochastic models to capture uncertainty [13,25]). However, few of these analyses focus on models developed early in the epidemics. The accuracy and appropriate use of early models is rarely assessed, first because of a lack of emphasis and resources while the epidemic is underway, and later because early models tend to give way to more sophisticated and better-fitting models as more information and data are acquired (e.g. moving from using single transmission rate during the early COVID-19 lockdowns [34] to continuous-time human movement data [35]). Thus, we know less about the value and limitations of models that are built and applied early in an epidemic when epidemiological information is severely limited. Yet, this early model assessment remains vitally important because such models often inform policy and public opinion in real time, even if they are later revised. To understand and anticipate problems for future emerging infectious diseases, and to produce models that will be taken up by policymakers, it is critical to reflect upon the value and limitations of early models and to assess their accuracy over time.

The need for rapid model development with incomplete and uncertain data forces modellers to make a series of decisions and assumptions, many of which must be made with relatively little empirical evidence (e.g. about the proportion of infections that are asymptomatic and the transmission potential of asymptomatic infections). All early models of COVID-19 dynamics, for example, were constrained by

limited: (1) observations of unmitigated epidemic dynamics from which to inform key epidemiological parameters like  $\mathcal{R}_0$ ; (2) information about the impact of preliminary interventions and (3) availability of testing, which made case data an unreliable indicator of epidemic magnitude and dynamics. Further, because of regional differences in socio-economic conditions and demography paired with difficult-to-observe case importations, disentangling local epidemic dynamics from policy interventions (and estimating the effectiveness of those interventions) proved difficult [36]. Data limitations also generate trade-offs between realistic complexity, parameter identifiability, computational feasibility and accuracy (see [37]), which require models to be designed around a targeted purpose rather than comprehensively describing all aspects of disease dynamics. Early in the COVID-19 pandemic, the need for rapid model deployment resulted in some researchers adopting a minimally complex statistical approach with the aim of producing near-term (e.g. one to five month) epidemic forecasts (e.g. [38,39]). However, because statistical models do not capture the underlying mechanistic transmission process, early COVID-19 statistical models were poorly suited to predict the effects of non-pharmaceutical interventions. On the opposite end of the spectrum, mechanistic, agent-based models sought to more precisely estimate the potential effects of different social distancing policies by incorporating population structures and individual movement (e.g. [2]). Such models were, however, computationally intensive, contained a large number of parameters, and did not always have publicly available code from the outset, making them infeasible to rapidly fit and simulate in many locations and relatively inaccessible to outside research groups and decision makers. Similar types of uncertainty and the need for rapid model development also led to the mixed success of early models of SARS [13], H1N1 pandemic influenza [14–17], Ebola [22–28] and Zika [32].

In addition to limited information, unreliable data and model trade-offs, a hurdle in the use of models early in the COVID-19 epidemic was the rapidity and heterogeneity of policy changes (especially in locations where interventions varied at a local level, for example in the USA). These changes, along with temporal variation in human behaviour (which often changed in advance of government interventions [40]) and extrinsic factors (e.g. seasonality), quickly rendered many models obsolete. Though many models were being continually updated (using, for example, more reliable case data due to expanded testing and cell phone-based mobility data), delays between changes in epidemic drivers and model improvements, their uptake, and public health decisions led to many models being used after the epidemiological environment was no longer reflected in the model's underlying assumptions. Furthermore, in the USA, for example, the fragmented COVID-19 response forced many state and local governments to rely upon a few highly publicized early models, which led to an outsized influence of some early non-peer-reviewed models (e.g. [38]).

Here, we quantitatively evaluate the successes and failures of an early epidemic model by retrospectively analysing an epidemiological compartment model that we developed in March 2020 for COVID-19 dynamics in Santa Clara County, California during the first wave of the US pandemic. While later development improved upon this model (see figure 1; [42]), here we use the early model as a snapshot in time, analysing it as an artefact rather than improving upon it



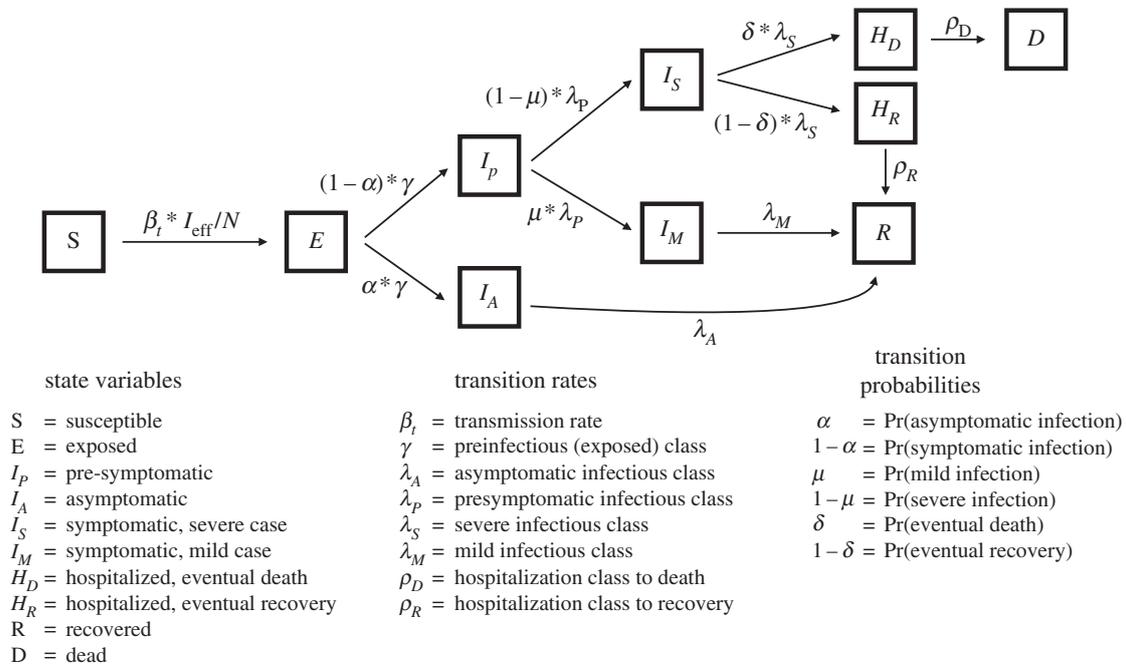
**Figure 1.** Timeline of the early COVID-19 epidemic period in Santa Clara County, California between February and May 2020. Reported COVID-19 death data (top) were used to estimate model parameters. The proportion of devices sheltering in place (bottom) [41] helps to illustrate how behaviour changed along the timeline of public health interventions and epidemic dynamics; these data were not used in our early model. (Online version in colour.)

[43]. That is, we intentionally seek to evaluate our model in light of the limited information and rapid decisions used to build it as well as the sparse data used to fit it. We critically gauge the strengths and limitations of this model by evaluating decisions that benefited and hindered model accuracy and potential model mis-specifications, and by quantifying the accuracy of predictions over time to understand why inaccuracy increased.

We began development on the model on 13 March 2020, 4 days after the first reported death in Santa Clara County and four days before the introduction of the San Francisco Bay Area shelter-in-place orders that applied to this county, which were initially proposed to last three weeks (figure 1). Our first aim was to deploy (within approximately two weeks) a public-facing user-friendly graphical model (<http://covid-measures.stanford.edu/>) that would allow users to adjust intervention parameters to help the public and local decision-makers understand the potential for resurgence if restrictions were lifted or relaxed too early and to evaluate viable exit strategies. Like other early COVID-19 epidemic models that were built for a specific purpose, we designed the model with the following considerations in mind. First, because we sought to compare the effects of various non-pharmaceutical interventions, we chose to build a mechanistic epidemiological compartment (susceptible–exposed–infectious–recovered: SEIR) model. Second, given early work

showing the potential for asymptomatic transmission [44] and our interest in symptomatic isolation as a potential intervention strategy, we broke up the infected classes into multiple compartments (asymptomatic, pre-symptomatic, mildly symptomatic and severely symptomatic). Because we were also interested in interventions triggered by hospital capacity thresholds, we included model compartments (state variables) for hospitalizations. Third, as we were interested in the implications of both initial, short-term interventions and longer-term exit strategies, we used a time-varying transmission parameter,  $\beta_t$ , to encapsulate the impact of non-pharmaceutical interventions on epidemic dynamics and control. Fourth, as case data were significantly biased at the time due to limits in testing capacity and access, we fit the model to local epidemic dynamics using daily reported COVID-19 deaths in Santa Clara County (which we assumed were more reliably reported than cases). We completed our first analyses on 13 April and posted a preprint on 30 April 2020 (figure 1). In the light of the narrow time window and these considerations, like many other early models we made a series of decisions (electronic supplementary material, table S1), some of which we deemed to be sub-optimal and improved in our later model, which we began developing on 1 May 2020 [42].

After developing the public-facing website, we originally used this model to estimate key epidemiological metrics and to evaluate the effectiveness of long-term intervention



**Figure 2.** Epidemiological model structure. English letters in boxes designate state variables (model compartments), while the Greek letters  $\beta$ ,  $\gamma$ ,  $\lambda$  and  $\rho$  refer to transition rates between states; the Greek letters  $\alpha$ ,  $\mu$  and  $\delta$  designate transition probabilities between states. We assume that the *per capita* transmission rate ( $\beta_t$ ) is directly proportional to the effectiveness of social distancing ( $\sigma_t$ ), such that  $\beta_t = \beta_0 \cdot \sigma_t$ , where  $\beta_0$  is the transmission rate prior to any social distancing (i.e.  $\sigma = 1$ ). The transition rate between the susceptible and exposed states is given by the force of infection ( $\beta_t \cdot I_{\text{eff}}/N$ ), where  $I_{\text{eff}}$  is the effective number infectious, which is equal to the weighted sum of infectious classes (in which the weights are the relative infectiousness of the infected classes:  $I_{\text{eff}} = I_A \kappa_A + I_p^* \kappa_p + I_M \kappa_M + I_S \kappa_S$ ). See electronic supplementary material, tables S3 and S4 for details on the relative infectiousness ( $\kappa$ ) parameters.

strategies, specifically focusing on Santa Clara County, California as a case study. We estimated the transmission rate under pre-intervention and shelter-in-place conditions, calculated reproduction numbers before and during interventions, explored the impact of long-term intervention strategies, and investigated counterfactuals to understand the impact of early intervention decisions. We now seek to answer the following retrospective questions: what did the model suggest about epidemic metrics, dynamics, and the impact of non-pharmaceutical interventions, and how did these estimates change over time? How accurately did the model predict epidemic dynamics going forward? For how long was the model accurate enough to be useful, and what limited its longer-term accuracy?

## 2. Methods

### (a) Model structure

We developed a stochastic compartmental model using an SEIR framework. We divided the population into states with respect to SARS-CoV-2 infection: susceptible (S); exposed but not yet infectious (E); infectious and pre-symptomatic ( $I_p$ ), asymptomatic ( $I_A$ ), mildly symptomatic ( $I_M$ ), or severely symptomatic ( $I_S$ ); hospitalized cases that will recover ( $H_R$ ) or die ( $H_D$ ); recovered and immune (R); and dead (D). We assumed an underlying, unobserved process model of SARS-CoV-2 transmission depicted in figure 2. We used a Euler approximation of the continuous time process with a time step of 4 h. We assumed that transitions between states were simulated as binomial or multinomial processes, which treat periods within each state as being geometrically distributed. Given that each period is geometrically distributed with a different rate, transition times through multiple states follow no named distribution but are unimodal (e.g. disease onset-to-death: electronic supplementary material, figure S1). It is

also possible to divide each state (e.g. infectious classes) into multiple sub-stages to produce Erlang-distributed periods within stages [45,46] (a change that we implemented in later iterations of our model [42]); however, here we relied on single compartments for each state for simplicity (electronic supplementary material, table S1). The equations (electronic supplementary material, Eq. S1–Eq. S8) describe in detail the stochastic transitions between states. We assumed that the observed deaths are a Poisson random variable with mean of total new deaths accumulated over the observation period (1 day for this analysis).

The transmission parameter,  $\beta_t$ , describes the average *per capita* contact rate, at time  $t$ , between susceptible and infectious people multiplied by the per-contact transmission probability. We defined  $\beta_t$  as being directly proportional to the impact of social distancing at time  $t$ , which is given by  $\sigma_t$ , such that  $\beta_t = \beta_0 \cdot \sigma_t$ , where  $\beta_0$  is the transmission rate prior to any social distancing (with  $\sigma_t = 1$ ). The degree to which a social distancing intervention reduces the overall population contact rate is  $1 - \sigma_t$ . While any sequence of time-varying transmission rates can be implemented in this framework, given the limited data to inform estimates of different transmission rates, we model  $\beta_t$  in three distinct segments of time that are characterized by different social contact structures: (1) baseline prior to any interventions ( $\beta_0 = \beta_0 \cdot 1$ ), assumed to occur at least until 29 February (2) the San Francisco Bay Area ‘work-from-home’ initiative, which we model as beginning some time between 1 March and 9 March ( $\beta_{\text{WFH}} = \beta_0 \cdot \sigma_{\text{WFH}}$ ); and (3) the San Francisco Bay Area shelter-in-place, which began on 17 March ( $\beta_{\text{SIP}} = \beta_0 \cdot \sigma_{\text{SIP}}$ ). We included  $\sigma_{\text{WFH}}$  and the work-from-home start date as two of the parameters we sampled across a plausible parameter range (see electronic supplementary material, table S4) but allowed  $\sigma_{\text{SIP}}$  to be estimated by the model (see ‘Fitting the Model’ below). We did not fit  $\sigma_{\text{WFH}}$  due to concerns about identifiability as  $\sigma_{\text{WFH}}$  modifies the transmission rate for only a brief period of time prior to the first observed death in Santa Clara County (figure 1).

By including asymptomatic and pre-symptomatic individuals, we were able to track ‘silent spreaders’ of the disease, both of

which have been shown to contribute to COVID-19 transmission [47] (electronic supplementary material, table S1). Tracking deaths allowed us to compare our simulations to a data source that was likely more reliable than confirmed cases, particularly in the absence of widespread rapid testing and case detection. Mildly symptomatic cases were defined as those people that show symptoms but do not require hospitalization, while we assumed that all severely symptomatic cases would eventually require hospitalization (figure 2). We also assumed that no onward transmission occurred from hospitalized individuals. We further assumed that all individuals not exposed to the virus begin as susceptible to infection, and that all model compartments other than susceptible and exposed began with zero individuals.

## (b) Fitting the model

We fit  $\beta_0$ , which describes the initial transmission rate in the absence of any interventions;  $\sigma_{\text{SIP}}$ , which describes the proportional reduction in  $\beta_0$  under shelter-in-place; and  $E_0$ , with which we drew the initial number of exposed individuals as  $\text{Poisson}(E_0) + 1$ . Fitting more than these parameters with only a few weeks of daily deaths (our first model iteration was hosted online on 22 March 2020: figure 1), was unrealistic because parameters were not identifiable. To estimate these three parameters, we assumed point estimates for parameters for which there was at least some convergence in estimates in the literature (electronic supplementary material, table S3); most notably these parameters include the average time individuals spend in infectious states. We use the inverse of durations as average exit rates, but note the possibility that taking the inverse of durations from individual-based studies (e.g. incubation period [49] and time from symptom onset to death [59,62]) might not scale appropriately for use as rates in a population-level model. For the remaining parameters, we drew 200 Sobol sequences, a more efficient method than Latin hypercube for sampling input parameters [64], across a range of plausible values (electronic supplementary material, table S4) to form 200 plausible parameter sets. While sampling over all non-fitted parameters is possible, we decided against this strategy in an effort to focus computation time on the areas of greatest uncertainty.

We note that we use  $\lambda$  to refer to the exponential rates at which individuals leave infectious classes and not force of infection as is common. The  $\kappa$  parameters (electronic supplementary material, table S3) scale  $\beta_i$  for individual infectious classes, where  $\kappa=1$  indicates no scaling. For simplicity and in the absence of better data, we assumed that only asymptomatic transmission had a scaling factor different from one (specifically,  $\kappa_A < 1$ ; see electronic supplementary material, table S4). For the rates shown in electronic supplementary material, tables S3 and S4, the unimodal distribution for the time from first symptoms to death has a mean of approximately 23.5 days, a median of 20 days, a mode greater than zero and a moderate right-skew (electronic supplementary material, figure S1). This median is between the mean value of 17.8 found by Verity *et al.* [59,72] and the range of 35–44 days observed in Bi *et al.* [73].

Using the pomp (statistical inference for partially observed Markov processes) package [74] (function `mif2`) in the R programming language [75], we fit  $\beta_0$ ,  $\sigma_{\text{SIP}}$  and  $E_0$  to daily deaths for each of the 200 parameter sets using six independent replicate particle filtering runs. For each independent replicate, we perturbed the starting values for fitted parameters among runs using random samples from a lognormal distribution for  $\beta_0$  (meanlog =  $\log(0.7)$ , sdlog = 0.17) and  $\sigma_{\text{SIP}}$  (meanlog =  $\log(0.2)$ , sdlog = 0.2) and a uniform distribution between 0 and 6 for  $E_0$ . Each individual `mif2` replicate run used 300 iterations, 1000 particles, a cooling fraction of 0.50, and a random-walk perturbation for all parameters of 0.02 (using the function `ivp` for  $E_0$  to designate it as an initial value parameter). The optimization was constrained to positive values for  $E_0$  and  $\beta_0$  and between zero and one for  $\sigma_{\text{SIP}}$ . After the

filtering steps were completed, for each `mif2` replicate run, log-likelihoods were calculated using the function `pfiler` 10 times with 10 000 particles each to produce both mean and standard errors for log likelihoods for each parameter set.

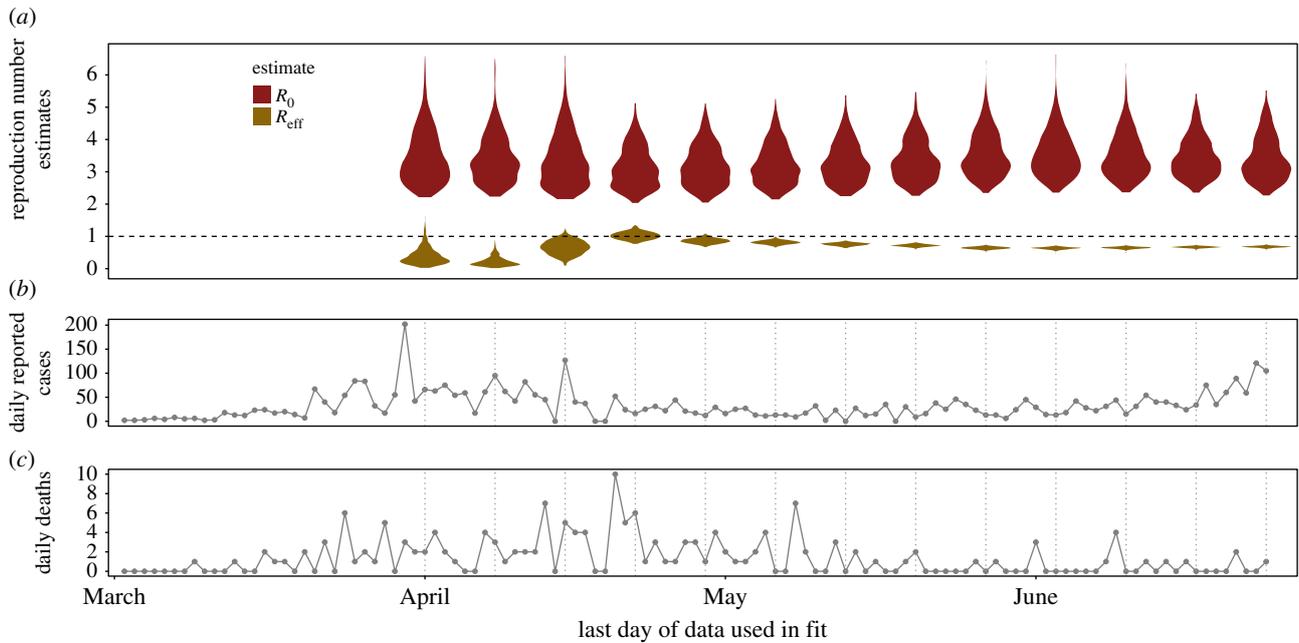
We computed weekly fits from 1 April through 24 June by withholding data reported after the given fit date. We used COVID-19 death data from *The New York Times*, based on reports from state and local health agencies (available at <https://github.com/nytimes/covid-19-data>, figure 3b). Daily deaths were calculated from differences in cumulative death reports. Using these data, which are available for all counties in the USA, our model can be used to fit  $\beta_0$ ,  $\sigma_{\text{SIP}}$  and  $E_0$  in any county.

We calculated  $\mathcal{R}_0$  as estimated  $\beta_0$  times the duration and infectiousness of an average infection (as defined by our model structure) for each of the 1200 parameter sets (using all six estimates from each of the `mif2` replicate runs). For each of the 1200 parameter sets, we calculated the effective reproduction number  $\mathcal{R}_E$  for each weekly fit by modifying the calculation for  $\mathcal{R}_0$  to scale the estimated  $\beta_0$  by both  $\sigma_{\text{SIP}}$  and the estimated median proportion of the population remaining susceptible across 200 simulated epidemics for the given parameter set.

For each fitted model parameter ( $\beta_0$ ,  $\sigma_{\text{SIP}}$  and  $E_0$ ), uncertainty comes from two sources: variation in fitted values among replicate `mif2` iterations (e.g. uncertainty in the value of fitted parameter conditional on a given parameter set—a single given conceivable state of the world), and variation in the estimated parameter value across the 200 parameter sets (uncertainty in the value of the fitted parameter given uncertainty in the state of the world). We computed likelihood profiles for the three fit parameters over 30 uniformly spaced points (hereafter, fixed points): from 0.2 to 1.2 for  $\beta_0$ , 0.01 to 0.6 for  $\sigma_{\text{SIP}}$  and 1 to 30 for  $E_0$ . For each parameter and each fixed point, we refit the model for each of 200 unique Sobol-sequenced parameter combinations with the same `mif2` settings used in other model fitting steps (except with three `mif2` replicates rather than six due to computational cost). We identified the maximum log-likelihood for each fixed point among all 600 fits (200 parameter sets, each fit three times with random starting values). We computed likelihood profiles for three fit dates only (1 April, 13 May and 24 June 2020) because of computational costs (18 000 model fits per profile).

## (c) Simulating epidemics

While a set of parameters can produce many simulations, only some of these trajectories are conceivable given the observed data. To simulate using only trajectories that are plausible conditional on the observed data, we drew trajectories from the smoothing distribution using the `filter.traj` and `pf` functions in `pomp` with 5000 particles for each particle filter. These filtering trajectories can be viewed as weighted samples from the distribution of unobserved state processes given the observed data, where the weights are determined by the likelihoods from the particle filtering [74]. All simulations were run forward in time from filtering trajectories, which constrains forecasts to continue from a present state matched to the observed epidemic dynamics in order to avoid overly large forecast uncertainty. Unless otherwise noted, all simulations used parameter sets within the top two log likelihood units for each fit date. For all simulated trajectories, we used 25 filtered trajectories and 25 forward simulations from each filtered trajectory for a total of 625 total epidemic forecasts for each of the parameter sets within the top two log-likelihood units for each fit date (the number of which varied by fit date). Unless noted otherwise, the uncertainty bands that we display for all simulations prior to the fit date contain the central 95% range of outcomes across parameter variation among the fits within the top two log likelihood units and variation among the 25 filtered trajectories. The uncertainty bands after the fit date (forecasts) contain the same parameter variation



**Figure 3.**  $\mathcal{R}_0$  (red, upper) and  $\mathcal{R}_E$  (gold, lower) estimated using time series death data up to the date on the horizontal axis (violin density plots, *a*). Time series show reported cases (*b*; not used in model fitting) and reported deaths used for model fitting (*c*) in Santa Clara County, with vertical dotted lines marking last data point used in each corresponding model fit (*a*). (Online version in colour.)

and stochastic variation and include additional variation from 25 forward simulations for each of the 25 filtered trajectories. The uncertainty bands on future simulations of deaths also contain additional variation from the Poisson observation process. These are simultaneously wide because of large numbers of stochastic simulations, but narrow because we ignored uncertainty in the parameters listed in electronic supplementary material, table S3 as well as uncertainty in the estimated parameters for each fit, and thus should be interpreted with caution.

### (i) Model assessment

To assess model performance, we compared model-forecasted deaths—simulated for 14 days forward in time from filtering trajectories assuming that the existing levels of social distancing were maintained—to observed deaths. Specifically, we quantified model forecast performance for deaths using the ‘quadratic score’. The quadratic score is a commonly used strictly proper scoring rule (a forecasting evaluation metric with a unique maximum that is reached by increasing both the accuracy and the concentration of the predictive distribution around the true value) for a predictive model with a discrete (e.g. Poisson) error distribution [76–79]. We calculated the quadratic score for each simulation over the 14 days following the fit date as

$$\frac{1}{N} \sum_{i=1}^N [-2(\Pr(y_i|\hat{\mu}_i)) + \sum_{k=0}^{\infty} (\Pr(k|\hat{\mu}_i))^2],$$

where  $i$  indexes the days since the fit date (from 1 to  $N=14$ ),  $y_i$  is the observed new daily reported deaths,  $\hat{\mu}_i$  is the daily prediction from the simulation, and  $\Pr(y_i|\hat{\mu}_i)$  is the probability mass on  $y_i$  given the prediction  $\hat{\mu}_i$ . The sum  $\sum_{k=0}^{\infty} (\Pr(k|\hat{\mu}_i))^2$  runs over all positive integers ( $k$ ) to measure the dispersion of probability mass given the prediction  $\hat{\mu}_i$ ; for a Poisson distribution, we calculate this sum analytically based on the following closed form expression for any  $x \geq 0$ :  $\sum_{k=0}^{\infty} (\Pr(k|x))^2 = e^{-2x} I_0(x)$ , where  $I_0$  is the Bessel function. For the model predictions  $\hat{\mu}_i$ , we used the underlying new daily deaths from the simulation’s trajectory ( $D_{\text{new}}$ ), rather than the deaths arising from the Poisson observation process, given that the model-predicted distribution of deaths for a given simulation is best characterized by  $\text{Poisson}(D_{\text{new}})$ . For each fit date, we show the distribution of

quadratic scores across simulations from fitted parameter sets within the top two log-likelihood units (one score per simulation and 625 simulations—25 forecasts for each of 25 filtered trajectories—per fitted parameter set).

We refrained from calculating a quadratic score comparing the model’s predicted cases and the reported cases given that we did not model incomplete case detection. Instead, we simply relied on a visual comparison between the trajectory (curvature) of the predicted cases and reported cases to qualitatively assess the model’s predictive accuracy. Specifically, we compared predicted new symptomatic infections to the cases reported one week later in order to account for a week of reporting lag.

### (ii) Future interventions

Our modelling framework allows for different types, intensities and durations of interventions, and thereby illustrates how these interventions impact dynamics and the resulting number of COVID-19 cases and fatalities through time. Here, we use fits generated from deaths reported prior to 22 April to consider three possible interventions that can be implemented at different times during the simulation:

1. *Social distancing for a set duration* applied as a scaling of the transmission rate for all individuals;
2. *Isolation of symptomatic individuals* applied as a scaling of the transmission rate for only symptomatic individuals  $I_S$  and  $I_M$ ; we assume isolation paired with partially relaxed social distancing;
3. *Adaptive triggering* applied as a tightening or relaxing of social distancing, triggered by hospitalizations crossing a defined threshold.

Other scenarios that can be modelled as a time-varying reduction in  $\beta_0$  (such as contact tracing and quarantine, which we do not include here) can be explored using the open-source code (available at [https://github.com/marissachilds/COVID19\\_early\\_model](https://github.com/marissachilds/COVID19_early_model)) and saved model fit data files (available on Dryad: <https://doi.org/10.5061/dryad.cvdncjt4t>).

To visualize the dynamics of a single intervention scenario, we simulated 25 epidemics from each of 25 filtered trajectories

(for a total of 625 total simulations) from the parameter set with the best likelihood using data up to 22 April. To quantify the effectiveness of each intervention scenario, we estimated cumulative deaths and the number of new symptomatic infections over time from the simulated epidemics for parameters sets within 2 log-likelihood units of the best for each intervention (e.g. the effectiveness of infected isolation).

### (iii) Counterfactuals

In addition to making forward projections under different intervention scenarios, models can help compare past actions taken (e.g. public health interventions) to alternative hypothetical scenarios (e.g. alternative types and timings of non-pharmaceutical interventions). Such comparisons can help to highlight which actions were the most helpful and which could have been improved. Early in an epidemic, a counterfactual analysis is particularly useful to assist local policymakers and the public contextualize the impact of early decisions relative to other possible decisions that could have been made. To assess the impact that existing county orders and resources had on the epidemic trajectory, we limited filtering trajectories to the date in which the counterfactual scenarios diverged (i.e. 17 March when the county shelter-in-place order went into effect), then simulated forward assuming: (1) shelter-in-place orders went into effect on 17 March; (2) shelter-in-place orders went into effect one week later on 24 March and (3) testing and isolation of infected individuals began in addition to the shelter-in-place orders on 17 March (in reality, testing remained limited in Santa Clara County and throughout the USA through the end of April [80]). In particular, we assumed that testing and isolation of symptomatic individuals further reduced their infectious contacts by 80% for severely symptomatic individuals and 70% for mildly symptomatic individuals.

## 3. Results

### (a) Model estimates and performance over time

We iteratively fit our model each week from 1 April through 24 June 2020 using data on daily reported COVID-19 deaths up to that date. For each fitted model, we first estimated  $\mathcal{R}_0$  and  $\mathcal{R}_E$  to investigate how our understanding of epidemic dynamics changed with increasing data availability, then compared the fit of simulations to out-of-sample data to evaluate how model performance changed over time.

The model consistently estimated that  $\mathcal{R}_0$  was between 3 and 4 (with a median among all fits of 3.76), though  $\mathcal{R}_E$  varied considerably, especially among the fits throughout April (figure 3). For example,  $\mathcal{R}_E$  jumped from a confident estimate below one on the 15 April fit to an uncertain estimate spanning one on the 22 April fit, after a week of higher deaths was included. After the April volatile period,  $\mathcal{R}_E$  estimates stabilized near 0.69 by mid-late May with very narrow confidence intervals (e.g. on 27 May the model estimated a median  $\mathcal{R}_E$  of 0.642 with a 95% CI of 0.571–0.708 among fits within two units of the top log likelihood) and continued to vary little throughout June. Even as cases increased in June (figure 3), deaths remained low, leading to little change in  $\mathcal{R}_E$  estimates. The estimated impact of social distancing ( $\sigma_{\text{SIP}}$ ) and transmission rate in the absence of non-pharmaceutical interventions ( $\beta_0$ ) followed a similar pattern of increased confidence over time, while the initially exposed class ( $E_0$ ) was estimated with large uncertainty in all fits (electronic supplementary material, figures S2 top panel and S3). Profiles over the fitted parameters reflect these patterns: the  $\beta_0$  and  $\sigma_{\text{SIP}}$  profiles showed clear peaks, the  $E_0$  profile was flat across

the range of values examined, and the  $\sigma_{\text{SIP}}$  profile began showing jaggedness for the last fit date (24 June 2020) when the model was no longer suited for the changing epidemiological situation (electronic supplementary material, figure S2 bottom panel). The difficulty in identifying the initial value parameter  $E_0$  is not surprising given that only early time points are expected to inform the initial state in a stochastic process [81]; in the case of this dataset, the time between the initial conditions (the start of the epidemic) and the first observation reached up to 68 days for some parameter sets. We ran additional diagnostics to understand the effect of the model's difficulty in identifying  $E_0$  (electronic supplementary material, figure S3) on the other focal parameters and quantities of interest. We found that altered *mif2* settings (an expanded range of starting values and random walk standard deviations; see electronic supplementary material, figure S3): (1) permitted larger  $E_0$  values which allowed for overly late, biologically implausible, start dates (electronic supplementary material, figure S3A); (2) did not lead to much better convergence (electronic supplementary material, figure S3B) and (3) despite allowing for much larger estimates of  $E_0$ , did not meaningfully impact the values of other fitted parameters (electronic supplementary material, figure S3C).

Despite the uncertainty in initial conditions, strong convergence of the 10 replicate log-likelihood estimates (as measured by small standard errors among them, electronic supplementary material, figure S4), the six replicate *mif2* runs in terms of log-likelihood (electronic supplementary material, figure S5), and good convergence in estimates for  $\sigma_{\text{SIP}}$  and  $\beta_0$  (electronic supplementary material, figure S6) indicates that much of this uncertainty is due to the inability of our model to differentiate among alternative parameter sets (parameter identifiability issues) (electronic supplementary material, figures S7–S10) and not a misuse or failure of the fitting algorithm or log-likelihood calculations.

Near-term forecasts and model performance also varied substantially over time. Simulations based on the model fit to deaths through 1 April show that while the bulk of simulations predicted declining deaths (figure 4, corresponding to the majority of the  $\mathcal{R}_E$  density being below 1, see figure 3), some simulations show rapidly increasing daily deaths over time (corresponding to an  $\mathcal{R}_E > 1$ ). With the 22 April model fit, uncertainty in whether  $\mathcal{R}_E$  was above or below 1 (figure 3) resulted in very large uncertainty in epidemic trajectories (figure 4). Simulations from later model fits (e.g. those from late May through early June) projected a decline in deaths through the end of June; model fits in June consistently suggested epidemic fade-out by the end of the month, assuming that the existing non-pharmaceutical intervention regime had remained in place. Weekly model fits (electronic supplementary material, figure S11) show that the model tended to under-predict deaths early in the study period (early to mid April) when a period of low daily deaths led to low estimates of  $\mathcal{R}_E$  (figure 3, gold shaded violin plots). The model then equally under- and over-predicted deaths in the middle of the study period (late April to mid-May), when sufficient data had likely allowed for more accurate estimates of the shelter-in-place effectiveness. Finally, the model under-predicted deaths again at the end of the study period (mid-May to late June) after shelter-in-place orders were relaxed on 4 May (figure 1) and the single estimated value for the effectiveness of social distancing was no longer realistic.

The model qualitatively matched the curvature in the cases reported within the time window used to fit the model for fits through 13 May (figure 5, blue points), although, as expected, the model predicted far more new symptomatic infections than reported cases (figure 5, difference in left and right vertical axes). Estimates of daily cases through the date of fitting produced realistic estimates of the proportion of the population remaining susceptible (electronic supplementary material, figure S12). However, all future projections failed to capture the trajectory of future, out-of-sample, reported cases; in most instances the model predicted declining cases despite the increase in cases observed starting in late May. By 3 June, the model began to fail to capture the increasing cases even within the observed time period for which death data were used to fit the model (figure 5, blue points). This inaccuracy in predictions was likely due to both the changing epidemiological environment (figure 1) that made model assumptions unrealistic (i.e. a constant effectiveness of social distancing,  $\sigma_{SIP}$ , and constant mortality rate) and the fact that changes in epidemic dynamics will be apparent in cases prior to deaths.

### (b) Scenario analysis

We originally designed and fitted our model (and accompanying interactive website) in part to communicate to local policymakers and the public the impact of the early social distancing interventions in Santa Clara County and the importance of continuing strong non-pharmaceutical interventions for saving lives and preventing an epidemic resurgence. To achieve these goals, we used counterfactual analyses to compare what transpired to alternative unrealized scenarios and to forecast the epidemic under alternative future scenarios with different non-pharmaceutical interventions. We revisit these analysis here, in brief, to illustrate this use of our model.

We estimated that a second peak would have been inevitable in the absence of any non-pharmaceutical interventions even if shelter-in-place had been maintained until 1 June 2020, as illustrated here for the single best-fitting parameter set (figure 6, red lines). Across all parameter sets within 2 log-likelihood units of the MLE and stochastic epidemic simulations, we estimated that in this scenario Santa Clara County would have had a median of 6140 deaths (95% CI: 546–19 494) and a peak number of daily new symptomatic infections of 33 193 (95% CI: 12 536–58 259) occurring on 5 July 2020 (95% CI: 26 June 2020–18 July 2020).

Maintaining shelter-in-place until 1 June, followed by less stringent social distancing (50% of baseline contacts), combined with strong symptomatic case isolation (removing an additional 80% and 70% of contacts from severe and mild symptomatic infections, respectively), would have allowed for higher background contact rates (e.g. more businesses reopening) and yet fewer deaths, as predicted by our single best model fit (figure 6). For a range of possible combinations of symptomatic case isolation efficiencies and background social distancing (electronic supplementary material, figure S13), we found an overlap in confidence intervals for deaths, but higher median estimated deaths at the weakest levels of social distancing in the general population. For reference, the median number of estimated deaths under maintained shelter-in-place is shown by the horizontal black line, with 80% and 95% CI in dashed and dotted lines, respectively (electronic supplementary material, figure S13). These confidence intervals span a wide range because our estimated  $\mathcal{R}_E$  values

as of 22 April ranged from 0.76–1.34, which led to some epidemics growing and some declining through time.

We proposed that, without widespread testing availability before the end of shelter-in-place, a hypothetical alternative strategy would have been adaptive triggering, in which social distancing orders are intensified and relaxed as hospitalizations exceed and fall below critical thresholds, respectively. However, because the estimated  $\mathcal{R}_E$  for Santa Clara County was approximately one (and the confidence interval included one on 22 April), a strategy that periodically reduces the strength of social distancing may have led to an overall increase in cases that would not be reversed when the shelter-in-place was reinstated. We found that an adaptive triggering strategy that alternates between social distancing that reduces transmission to 20% and 80% of baseline could be effective in keeping cases and deaths low (electronic supplementary material, figure S14). This method would have kept the epidemic within the capacity of the healthcare system, but resulted in prolonged cycles of epidemic resurgence and control, continuing until herd immunity was reached through recovery of infected individuals or vaccination.

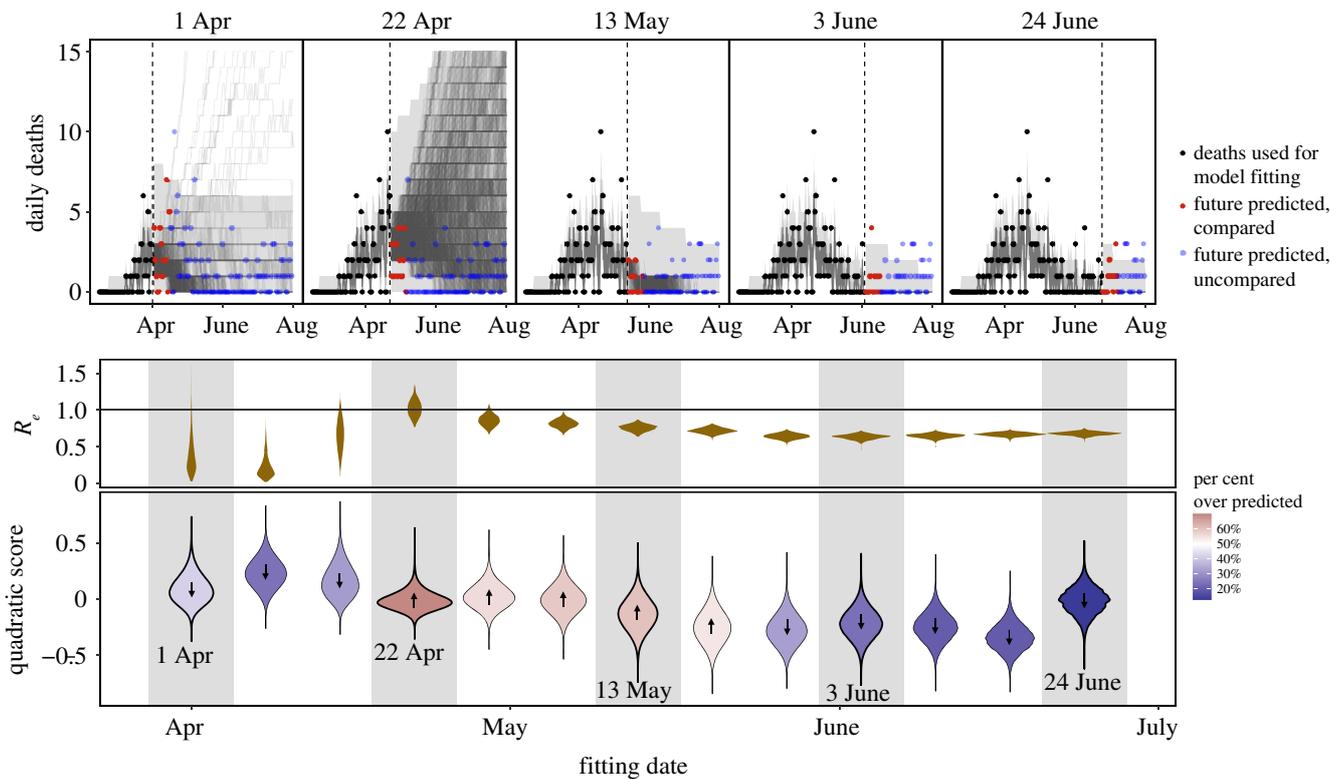
In simulations of counterfactual scenarios, we found that an additional 57 (95% CI: 10–143) lives would have been lost if shelter-in-place orders had been delayed even a week, and 26 (95% CI: 3–51) deaths could have been averted if testing and isolation of symptomatic individuals was available from the time of the shelter-in-place (electronic supplementary material, figure S15).

## 4. Discussion

During an unfolding pandemic, modelling is an essential tool for tactical decision-making, strategic planning and communicating qualitative scenarios [17]. Many early COVID-19 models played a critical role in highlighting the importance of social distancing to governments and to the public (e.g. [2–5]). Models like our own helped communicate to the public that ‘flattening the curve’ and slowing transmission was not a short-term endeavour, but also that early and sustained interventions had major benefits for local public health. Despite their importance, it is often unclear how quickly and for what reasons early models become obsolete given that retrospective analyses are not usually conducted on the first iterations of models. Here, we presented an example retrospective analysis on our early COVID-19 model to ask the following questions: (1) what did the model suggest about epidemic metrics, dynamics, and the impact of non-pharmaceutical interventions, and how did these estimates change over time? (2) How accurately did the model predict epidemic dynamics going forward? (3) For how long was the model accurate enough to be useful, and what limited its longer-term accuracy?

### (a) Epidemic dynamics over time (Q1)

Our model stably estimated  $\mathcal{R}_0$  between April and June with an overall median of 3.76 (figure 3), comparable to values estimated elsewhere (e.g. [82]), and identified that  $\mathcal{R}_E$  declined substantially after the shelter-in-place order was enacted, to near or below one. However, predicted future epidemic dynamics were highly uncertain given that many model parameters had large uncertainty (especially for the 1 April model fit, see electronic supplementary material, figures S2 and S7), estimated credible intervals on  $\mathcal{R}_E$  spanned one in



**Figure 4.** Model accuracy at predicting deaths over time peaked in mid-May to early June, as  $\mathcal{R}_E$  stabilized near 0.69. Top panels are simulations of the model fitted to reported deaths (black points) up to dates spanning from 1 April (left) to 24 June (right) and compared to future reported deaths (red points for deaths within two weeks, which were used for quadratic scoring, blue points for future deaths beyond two weeks). For future simulations (right of dashed line), medians of simulated observed deaths for each parameter set are shown as lines and a 95% confidence interval over all simulations from all parameter sets are shown in grey ribbons. For past filtered trajectories (left of dashed line), grey lines show medians of new deaths based on the underlying state process filtering trajectories (i.e. without a Poisson observation process) for each parameter set and grey ribbons show 95% central intervals among all trajectories for all parameter sets. Gold violin plots (middle) show the distribution of  $\mathcal{R}_E$  values on each fit date; violins on a grey background correspond to the trajectories in the top panels (matched by date). Blue-to-red coloured violin plots (bottom) illustrate the distribution of the quadratic score across simulations for the two weeks following each model fit. A lower quadratic score reflects better model predictions. Red shading and an up arrow indicate a higher percentage of overestimates (above 50%) and blue shading and a down arrow indicate a higher percentage of underestimates. All parameter sets within two log-likelihood units are included for each each fitting date, with 25 filtered trajectories for each parameter set and 25 forward simulations for each filtering trajectory resulting in 625 simulations for each parameter set. (Online version in colour.)

model fits until early-mid-May (figure 3), and  $\mathcal{R}_E$  estimates were variable from week to week when data were sparse (figure 3). For example, the inclusion of a week with higher deaths (most of which occurred in long-term care facilities [83], a distinction that is not captured in our model that assumes a homogeneous population), led estimated  $\mathcal{R}_E$  to jump to span one on 22 April (figure 3). The volatility and uncertainty in  $\mathcal{R}_E$  estimates highlight the difficulty in inferring epidemic metrics from deaths alone, which are a noisy and lagged indicator of the underlying epidemic dynamics. Despite these limitations, early models like ours play a critical role in estimating coarse epidemiological metrics and dynamics (e.g.  $\mathcal{R}_0$  and  $\mathcal{R}_E$ ), which are fundamental for quantifying and comparing the efficacy of various intervention scenarios and predicting the future course of the epidemic.

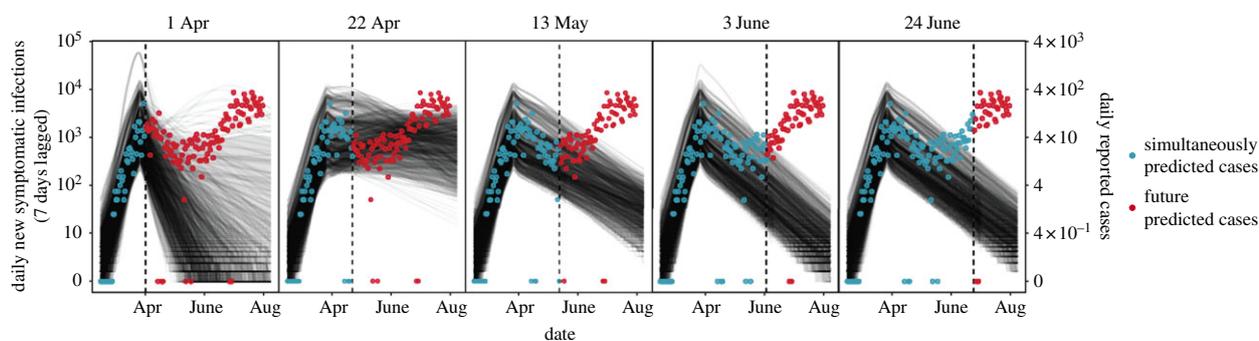
### (b) Forecasting accuracy (Q2)

Our model initially gained accuracy in predicting epidemic dynamics as additional data increased parameter identifiability (April to mid-May 2020, figure 4, bottom panel) but then began to decline in performance as the model assumptions became too simplistic to capture the changing epidemiological context (late May to June 2020). Out-of-sample predictions of deaths suffered from high uncertainty during the early period (e.g. on 22 April)

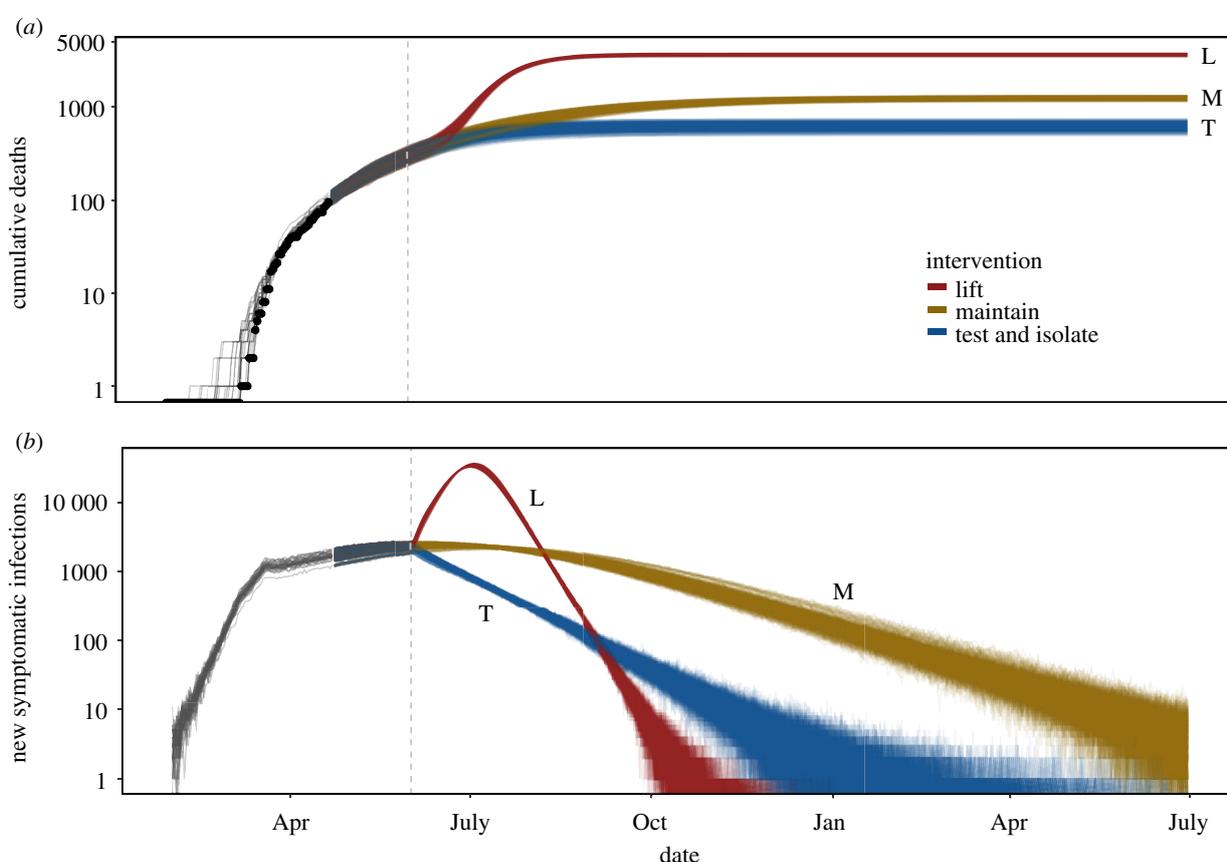
while predictions during the end of the study window (e.g. on 3 June) became overly confident that the epidemic would die out. Because of limited and variable testing capacity, we relied on a visual comparison of the curvature in model predictions of new symptomatic infections to observed daily cases (lagged one week for a plausible reporting lag). Model fits through 13 May qualitatively match the curvature in the reported cases within the time window used to fit the model, but by 3 June the model failed to capture even the increasing cases within the observed time period (figure 5, blue points). Thus, our model illustrates that trajectories of cases can be captured in relatively simple mechanistic models based on only reported deaths, but only for a limited period of time until epidemiological conditions change.

### (c) Limitations to long-term accuracy (Q3)

Predictions deteriorated as the epidemiological environment (figure 1) began to deviate further from our model assumptions. We fit a simple step function for the impact of non-pharmaceutical interventions with the aim of balancing identifiability in the face of limited data and accuracy of early predictions. However, this assumption restricted the utility of our model once Santa Clara County began to relax social distancing orders (figure 1). Additionally, the



**Figure 5.** The model accurately predicted cases until early June, when social distancing orders were relaxed. Panels show simulations of the model (which is fitted to reported death data; not shown) from 1 April (left) to 24 June (right), on a log scale. Model simulations of daily new symptomatic infections (with a one-week lag to approximate a plausible time from symptom onset to case reporting) are shown with lines and correspond to the left  $y$ -axis. Note that reported cases and predicted symptomatic infections differ by approximately 2 orders of magnitude, as expected due to under-reporting. As in figure 4, each black line represents the median among simulations for a single parameter set and the grey shaded region represents the 95% central interval among all simulations. Reported cases are shown as points and correspond to the right  $y$ -axis: blue points are cases reported during the time period used in model fits, while red points (those to the right of the vertical dotted line) are cases reported after the time period associated with model fitting. The vertical dashed line indicates the date of model fitting. (Online version in colour.)



**Figure 6.** Maintaining non-pharmaceutical interventions after 1 June, when early shelter-in-place restrictions were relaxed, is critical for preventing a devastating resurgence. Maintaining shelter-in-place at the  $\sigma_{SIP}$  value predicted by the model (here  $\sigma_{SIP} = 0.40$ ) (gold) or test-and-isolate ( $\sigma_{SIP} = 0.50$  plus an additional 80% effectiveness of test and isolate for severe infections and 70% for mild infections) (blue) strategies over long periods are necessary to prevent a major epidemic resurgence (red) following the end of the initial shelter-in-place order on 1 June (dashed vertical line) in Santa Clara County. Grey lines prior to 22 April show 25 filtered trajectories; coloured lines after 22 April show 25 simulated trajectories from each of the 25 filtered trajectories for each of three intervention scenarios. Black points in  $a$  show observed data. Cumulative deaths are plotted for visualization only; model is fit using raw death counts. For ease of visibility, only simulations from the parameter set with the maximum likelihood from the 22 April fits are shown here. (Online version in colour.)

relationship between cases and deaths fundamentally changed between the first and second waves. COVID-19 deaths remained relatively flat through July and August even while reported cases surged above 300 per day by the third week in July [83]. This pattern occurred in many places in

the US during the summer resurgence [84] and may reflect some combination of differences in personal protective behaviours and social distancing adherence across disease severity risk groups, resulting in a larger share of cases occurring in people less vulnerable to severe disease and death. Improved

standards of care, safety protocols and the availability of personal protective equipment in long-term care facilities, as well as increased testing to improve the detection of asymptomatic or mildly symptomatic cases all likely played a role as well.

#### (d) Early modelling and lessons for future epidemic control

Generally, real-time models face opposing forces: additional time means additional data to aid in parameter estimation; however, as more time passes the epidemiological environment deviates further from that which the early model was built to address, eventually reducing model accuracy [17,19]. Our case study and others (e.g. [85]) clearly demonstrate these opposing forces. Rapidly changing US policy forced many COVID-19 models to consider continuous-time estimates of movement (e.g. [42,86–89]) instead of constant intervention strengths. Yet, even these more detailed models faced a changing epidemiological context, which included the implementation of mask mandates and other behavioural changes (e.g. [90,91]), seasonality modifying viral kinetics and behavioural contact patterns, holidays altering mixing and travel patterns, and emerging virus variants with new epidemiological properties. In general, changes such as these cause unexpected variation in disease dynamics across time and space, limiting the accuracy of long-term forecasts [13,17,19,20] and potentially reducing the accuracy of our model, which relied partially on fixed parameters derived from early outbreaks in other locations (e.g. China [51,57,61,66], Italy [44] and Singapore [50]).

While simplifying assumptions are necessary in early and real-time models, those assumptions must be frequently re-evaluated to ensure continuing accuracy [11,19,22,24,27,32,33]. We made a series of simplifying assumptions for our early model in an attempt to overcome a lack of data (electronic supplementary material, table S1). While we relaxed some of these in a follow-up model [42], several alternative technical decisions could have improved our model from the beginning. For example, we used single compartments for each state rather than subdividing the infectious states into multiple compartments, resulting in geometrically distributed transition periods instead of more realistic Erlang-distributed periods [45,46]. Though it saved us only one parameter, we also assumed a Poisson observation process rather than the more flexible negative binomial observation process which is often estimable without too much difficulty (e.g. [25,72]). Finally, we did not consider uncertainty in the estimated parameter values, for example by using importance sampling [92].

Our case study highlights key lessons for the practice of early, real-time modelling of emerging epidemics. In our analysis, we focused on quantifying near-term forecast accuracy, a common strategy for evaluating model performance [85,93], which could be used more frequently from the outset to monitor changes in model performance over time. We also advocate tracking variation in parameters and signs of declining parameter identifiability. Together, systematically applying these approaches will provide warning signs of model inaccuracies sooner and help to ensure that hidden mis-specifications are identified and corrected promptly. Early models will inevitably be imperfect reflections of reality constrained by limited data; using them responsibly requires considering and communicating uncertainty, a benefit of stochastic frameworks like that employed here [25]. Our model

struggled to fit a relatively complex structure with a short time series of available data, particularly early in the epidemic, which resulted in some implausible epidemic dynamics (for example, unrealistically rapid depletion of susceptible individuals for some parameter sets from 1 April fits, electronic supplementary material, figure S12). This further highlights the difficulty of fitting complex models early in an epidemic and reinforces the importance of thorough evaluation of early predictions in order to avoid making biologically implausible claims. Our model was also particularly limited in its ability to estimate the initial number of infected individuals (electronic supplementary material, figures S2 and S3). In light of our and others' struggles with estimating  $E_0$  for COVID-19 and other emerging infectious diseases [24,29,65], surveillance programmes (including genomic surveillance [94]) may assist with both disease control and modelling in early stages of an epidemic.

Regularly updated, centralized databases for parameter estimates (e.g. [95]) and relevant time series data on epidemiology and mobility (e.g. [41,96,97]) are major assets for modelling in emerging epidemics. These tools facilitate more rapid development of early models and streamline comparisons among models [16]. Changes to any model, especially those addressing flaws in previous versions, should be clearly communicated and shared with adequate documentation to ensure that outdated versions of the model are not used and do not guide others' model development [17]. However, especially in early stages of a pandemic, a relatively common and simple mechanistic modelling framework with a long and robust history, as outlined here, may be able to provide quicker and more reliable insight into disease dynamics than developing new model structures from scratch. For example, SIR-type models can do surprisingly well predicting epidemic metrics even with limited data [98–100]. Furthermore, simple transmission functions can be an effective alternative to more complicated functional forms, as illustrated here for the period of strong social distancing between approximately mid-March and mid-May 2020 (figure 1). Simple mechanistic models, unlike phenomenological models (e.g. statistical curve fits), also allow for scenario analysis through alterations in inputs and parameters, which allows for longer-term forecasts comparing alternative interventions.

Given the similarity between our model and others developed concurrently (e.g. [91]), we recommend developing infrastructure to facilitate collaboration, rapid communication, and workflows to minimize duplication of effort, facilitate troubleshooting, and aggregate and analyse projections across sets of models [101]. Further, we identified human behavioural changes as a key source of inaccuracy in our model predictions, suggesting the importance of collaboration between disease modellers and behavioural scientists, as well as guidelines for proper incorporation of mobility data [88]. Greater engagement between policymakers and scientists, particularly to clarify types and timings of interventions being considered, the importance of key modelling decisions, and the differences between early models considered in policymaking (e.g. the strengths and weaknesses of both phenomenological and mechanistic models) would ensure that a model is appropriately designed (e.g. distinguishing between symptomatic and asymptomatic infections to capture the dynamical implications of testing and isolating only symptomatic infections) and applied to relevant scenarios.

## (e) Conclusion

Despite all that has been learned about the impact of non-pharmaceutical interventions such as social distancing and mask wearing, and the approval of effective clinical therapies and vaccines, the US experienced two additional major epidemic waves within a year that each dwarfed the one in the early epidemic and control period we studied here. Given the order-of-magnitude difference in deaths and over three orders-of-magnitude difference in cases observed between the spring 2020 and winter 2021 periods [83], it may be tempting to conclude that non-pharmaceutical interventions and public health orders did not work, or were too economically and socially costly to justify their use. However, this is a dangerous conclusion. Mechanistic models like those we present here make it clear that, however imperfect, these interventions saved large numbers of lives: as of 1 July 2021, Santa Clara County has seen 2201 total deaths [83], a terrible toll, but one that is only 30% of our median prediction occurring from an unmitigated epidemic. As of July 2021, vaccination coverage in adults in Santa Clara County has reached approximately 75%, which, combined with the anticipated eligibility for younger children [102], portends an end to the epidemic locally in the coming months. Even during the peak of the winter surge, the county saw just over 700 concurrent hospitalizations, far shy of our median estimated value of 12 975 (95% CI: 760–28 927) that could have occurred without control measures in place. Although the US COVID-19 response clearly could have been better at controlling transmission, illness, and death, mechanistic models make it clear that the situation also could have been much worse without the control measures that remained in place, which were at least in part motivated by early models. Moving forward with COVID-19 and in future epidemics, models that incorporate changes in contact behaviour, population immunity derived from natural infection and vaccination, population heterogeneity in behaviour and immunity, and changes in

immunity over time due to natural waning and emerging immune-evading variants will be critical for determining how to safely transition between initial and long-term interventions.

**Data accessibility.** Data used in this study are available at: <https://github.com/nytimes/covid-19-data>. Code used to produce the results in this study are available at: [https://github.com/marissachilds/COVID19\\_early\\_model](https://github.com/marissachilds/COVID19_early_model). Output from model fits and simulations are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.cvdncjt4t> [103].

**Authors' contributions.** M.C.: conceptualization, data curation, formal analysis, investigation, methodology, software, visualization, writing-original draft, writing-review and editing; M.K.: conceptualization, data curation, formal analysis, investigation, methodology, software, visualization, writing-original draft, writing-review and editing; M.J.H.: conceptualization, data curation, formal analysis, investigation, methodology, software, visualization, writing-original draft, writing-review and editing; D.K.: conceptualization, data curation, writing-original draft; L.I.C.: conceptualization, data curation, writing-original draft; N.N.: conceptualization, data curation, writing-original draft; I.D.: conceptualization, data curation, writing-original draft; J.R.: conceptualization, software, visualization; A.D.B.: investigation, methodology, software, visualization, writing-review and editing; E.M.: conceptualization, funding acquisition, methodology, project administration, supervision, writing-original draft, writing-review and editing.

**Competing interests.** We declare we no competing interests.

**Funding.** Funding provided by: the National Science Foundation (DEB-1518681 and DEB-2011147, along with the Fogarty International Center); the National Institute of General Medical Sciences (R35GM133439); the Natural Capital Project; the Helman Scholarship; the Terman Award; and seed grants from the Stanford King Center for Global Development, Center for Innovation in Global Health, and Woods Institute for the Environment. M.L.C. was supported by the Illich-Sadowsky Fellowship through the Stanford Interdisciplinary Graduate Fellowship programme at Stanford University. M.P.K. was supported by the Natural Capital Project. N.N. was supported by the Stanford Data Science Scholars programme. J.R. was supported by the Terry Winograd Fellowship. M.J.H. was supported by the Knight-Hennessy Scholars Program.

## References

- COVID-19 Dashboard. Johns Hopkins University; 2021. See <https://coronavirus.jhu.edu/map.html>.
- Ferguson N *et al.* 2020 Report 9: impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Accepted: 2020-03-17T09:57:15Z Publication Title: 20. See <http://spiral.imperial.ac.uk/handle/10044/1/77482>.
- Hellewell J *et al.* 2020 Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Global Health* **8**, e488–e496. (doi:10.1016/S2214-109X(20)30074-7)
- Hao X, Cheng S, Wu D, Wu T, Lin X, Wang C. 2020 Reconstruction of the full transmission dynamics of COVID-19 in Wuhan. *Nature* **584**, 420–424. (doi:10.1038/s41586-020-2554-8)
- Yang P, Qi J, Zhang S, Wang X, Bi G, Yang Y, Sheng B, Yang G. 2020 Feasibility study of mitigation and suppression strategies for controlling COVID-19 outbreaks in London and Wuhan. *PLoS ONE* **15**, e0236857. (doi:10.1371/journal.pone.0236857)
- Chowell G, Sattenspiel L, Bansal S, Viboud C. 2016 Mathematical models to characterize early epidemic growth: a review. *Phys. Life Rev.* **18**, 66–97. (doi:10.1016/j.plrev.2016.07.005)
- Pandey A, Atkins KE, Medlock J, Wenzel N, Townsend JP, Childs JE, Nyenswah TG, Ndeffo-Mbah ML, Galvani AP. 2014 Strategies for containing Ebola in West Africa. *Science* **346**, 991–995. (doi:10.1126/science.1260612)
- Yan Y, Shin WI, Pang YX, Meng Y, Lai J, You C, Zhao H, Lester E, Wu T, Pang CH. 2020 The first 75 days of novel coronavirus (SARS-CoV-2) outbreak: recent advances, prevention, and treatment. *Int. J. Environ. Res. Public Health* **17**, 2323. (doi:10.3390/ijerph17072323)
- Bedford J *et al.* 2020 COVID-19: towards controlling of a pandemic. *Lancet* **395**, 1015–1018. (doi:10.1016/S0140-6736(20)30673-5)
- Elder BD, Dukic VM, Dwyer G. 2006 Uncertainty in predictions of disease spread and public health responses to bioterrorism and emerging diseases. *Proc. Natl. Acad. Sci. USA* **103**, 15 693–15 697. (doi:10.1073/pnas.0600816103)
- Dye C, Gay N. 2003 Modeling the SARS epidemic. *Science* **300**, 1884–1885. (doi:10.1126/science.1086925)
- Galvani AP. 2004 Emerging infections: what have we learned from SARS? *Emerg. Infect. Dis.* **10**, 1351–1352. (doi:10.3201/eid1007.040166)
- Bauch CT, Lloyd-Smith JO, Coffee MP, Galvani AP. 2005 Dynamically modeling SARS and other newly emerging respiratory illnesses: past, present, and future. *Epidemiology* **16**, 791–801. (doi:10.1097/01.ede.0000181633.80269.4c)
- Kerkhove MDV *et al.* 2010 Studies needed to address public health challenges of the 2009 H1N1 influenza pandemic: insights from modeling. *PLoS Med.* **7**, e1000275. (doi:10.1371/journal.pmed.1000275)
- Lee BY, Waring AE. 2011 The 2009 H1N1 influenza pandemic. *Hum. Vaccin.* **7**, 115–119. (doi:10.4161/hv.7.1.13740)

16. Van Kerkhove MD, Ferguson NM. 2012 Epidemic and intervention modelling: a scientific rationale for policy decisions? Lessons from the 2009 influenza pandemic. *Bull. World Health Organ.* **90**, 306–310. (doi:10.2471/BLT.11.097949)
17. Lee BY, Haidari LA, Lee MS. 2013 Modelling during an emergency: the 2009 H1N1 influenza pandemic. *Clin. Microbiol. Infect.* **19**, 1014–1022. (doi:10.1111/1469-0691.12284)
18. Bajardi P, Poletto C, Ramasco JJ, Tizzoni M, Colizza V, Vespignani A. 2011 Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. *PLoS ONE* **6**, e16591. (doi:10.1371/journal.pone.0016591)
19. Poletti P, Ajelli M, Merler S. 2011 The effect of risk perception on the 2009 H1N1 pandemic influenza dynamics. *PLoS ONE* **6**, e16460. (doi:10.1371/journal.pone.0016460)
20. Opatowski L, Fraser C, Griffin J, De Silva E, Van Kerkhove MD, Lyons EJ, Cauchemez S, Ferguson NM. 2011 Transmission characteristics of the 2009 H1N1 influenza pandemic: comparison of 8 southern hemisphere countries. *PLoS Pathog.* **7**, e1002225. (doi:10.1371/journal.ppat.1002225)
21. Apolloni A, Poletto C, Colizza V. 2013 Age-specific contacts and travel patterns in the spatial spread of 2009 H1N1 influenza pandemic. *BMC Infect. Dis.* **13**, 176. (doi:10.1186/1471-2334-13-176)
22. Drake JM, Kaul RB, Alexander LW, O'Regan SM, Kramer AM, Pulliam JT, Ferrari MJ, Park AW. 2015 Ebola cases and health system demand in Liberia. *PLoS Biol.* **13**, e1002056. (doi:10.1371/journal.pbio.1002056)
23. Drake JM, Bakach I, Just MR, O'Regan SM, Gambhir M, Fung ICH. 2015 Transmission models of historical Ebola outbreaks. *Emerg. Infect. Dis.* **21**, 1447. (doi:10.3201/eid2108.141613)
24. Scarpino SV *et al.* 2015 Epidemiological and viral genomic sequence analysis of the 2014 Ebola outbreak reveals clustered transmission. *Clin. Infect. Dis.* **60**, 1079–1082. (doi:10.1093/cid/ciu1131)
25. King AA, Rohani P. 2015 Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to Ebola. *Proc. R. Soc. B* **282**, 20150347. (doi:10.1098/rspb.2015.0347)
26. Chretien JP, Riley S, George DB. 2015 Mathematical modeling of the West Africa Ebola epidemic. *eLife* **4**, e09186. (doi:10.7554/eLife.09186)
27. Burghardt K, Verzijl C, Huang J, Ingram M, Song B, Hasne MP. 2016 Testing modeling assumptions in the West Africa Ebola Outbreak. *Sci. Rep.* **6**, 34598. (doi:10.1038/srep34598)
28. Chowell G, Viboud C, Simonsen L, Merler S, Vespignani A. 2017 Perspectives on model forecasts of the 2014–2015 Ebola epidemic in West Africa: lessons and the way forward. *BMC Med.* **15**, 42. (doi:10.1186/s12916-017-0811-y)
29. Faria NR *et al.* 2017 Establishment and cryptic transmission of Zika virus in Brazil and the Americas. *Nature* **546**, 406–410. (doi:10.1038/nature22401)
30. Grubaugh ND *et al.* 2017 Genomic epidemiology reveals multiple introductions of Zika virus into the United States. *Nature* **546**, 401–405. (doi:10.1038/nature22400)
31. Metsky HC *et al.* 2017 Zika virus evolution and spread in the Americas. *Nature* **546**, 411–415. (doi:10.1038/nature22402)
32. Wiratsudakul A, Suparit P, Modchang C. 2018 Dynamics of Zika virus outbreaks: an overview of mathematical modeling approaches. *PeerJ* **6**, e4526. (doi:10.7717/peerj.4526)
33. Fox SJ, Bellan SE, Perkins TA, Johansson MA, Meyers LA. 2019 Downgrading disease transmission risk estimates using terminal importations. *PLOS Negl. Trop. Dis.* **13**, e0007395. (doi:10.1371/journal.pntd.0007395)
34. Kucharski AJ *et al.* 2020 Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect. Dis.* **20**, 553–558. (doi:10.1016/S1473-3099(20)30144-4)
35. Sanche S, Lin YT, Xu C, Romero-Severson E, Hengartner N, Ke R. 2020 High contagiousness and rapid spread of severe acute respiratory syndrome Coronavirus 2. *Emerg. Infect. Dis.* **26**, 1470–1477. (doi:10.3201/eid2607.200282)
36. Van Holm EJ, Wyczalkowski CK, Dantzler PA. 2021 Neighborhood conditions and the initial outbreak of COVID-19: the case of Louisiana. *J. Public Health* **43**, 219–224. (doi:10.1093/pubmed/fdaa147)
37. Keeling MJ, Rohani P. 2011 *Modeling infectious diseases in humans and animals*. Princeton, NJ: Princeton University Press.
38. University of Washington. IHME COVID-19 Projections. Library Catalog. See <https://covid19.healthdata.org/>.
39. Los Alamos National Laboratory. COVID-19 Science Updates; 2020. See <https://www.lanl.gov/updates/covid-19-science.php>.
40. Xiong C, Hu S, Yang M, Younes H, Luo W, Ghader S, Zhang L. 2020 Mobile device location data reveal human mobility response to state-level stay-at-home orders during the COVID-19 pandemic in the USA. *J. R. Soc. Interface* **17**, 20200344. (doi:10.1098/rsif.2020.0344)
41. SafeGraph Inc. Shelter in place index: the impact of coronavirus on human movement. 2020. See <https://www.safegraph.com/dashboard/covid19-shelter-in-place?s=US&d=06-04-2020&t=counties&m=index>.
42. Kain MP, Childs ML, Becker AD, Mordecai EA. 2021 Chopping the tail: how preventing superspreading can help to maintain COVID-19 control. *Epidemics* **34**, 100430. (doi:10.1016/j.epidem.2020.100430)
43. Jewell NP, Lewnard JA, Jewell BL. 2020 Predictive mathematical models of the COVID-19 pandemic: underlying principles and value of projections. *JAMA* **323**, 1893. (doi:10.1001/jama.2020.6585)
44. Lavezzo E *et al.* 2020 Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo? *Nature* **584**, 425–429. (doi:10.1038/s41586-020-2488-1)
45. Anderson D, Watson R. 1980 On the spread of a disease with gamma distributed latent and infectious periods. *Biometrika* **67**, 191–198. (doi:10.1093/biomet/67.1.191)
46. Lloyd AL. 2001 Realistic distributions of infectious periods in epidemic models: changing patterns of persistence and dynamics. *Theor. Popul. Biol.* **60**, 59–71. (doi:10.1006/tpbi.2001.1525)
47. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, Parker M, Bonsall D, Fraser C. 2020 Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368**, eabb6936. (doi:10.1126/science.abb6936)
48. McAloon C *et al.* 2020 Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research. *BMJ Open* **10**, e039652. (doi:10.1136/bmjopen-2020-039652)
49. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, Azman AS, Reich NG, Lessler J. 2020 The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Ann. Intern. Med.* **172**, 577–582. (doi:10.7326/M20-0504)
50. Wei WE, Li Z, Chiew CJ, Yong SE, Toh MP, Lee VJ. 2020 Presymptomatic Transmission of SARS-CoV-2? Singapore, January 23–March 16, 2020. *Morb. Mortal. Wkly Rep.* **69**, 411. (doi:10.15585/mmwr.mm6914e1)
51. Zhang W *et al.* 2020 Secondary transmission of coronavirus disease from presymptomatic persons, China. *Emerg. Infect. Dis.* **26**, 1924. (doi:10.3201/eid2608.201142)
52. He X *et al.* 2020 Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.* **26**, 672–675. (doi:10.1038/s41591-020-0869-5)
53. Zhang W. 2020 Estimating the presymptomatic transmission of COVID 19 using incubation period and serial interval data. medRxiv. See <https://www.medrxiv.org/content/medrxiv/early/2020/04/06/2020.04.02.20051318>.
54. Gao W, Li L. 2020 Advances on presymptomatic or asymptomatic carrier transmission of COVID-19. *Zhonghua Liu Xing Bing Xue Za Zhi* **41**, 485–488.
55. Arons MM *et al.* 2020 Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. *N. Engl. J. Med.* **382**, 2081–2090. (doi:10.1056/NEJMoa2008457)
56. Lee S, Meyler P, Mozel M, Tauh T, Merchant R. 2020 Asymptomatic carriage and transmission of SARS-CoV-2: what do we know? *Can. J. Anaesth.* **67**, 1–7. (doi:10.1007/s12630-020-01729-x)
57. Wölfel R *et al.* 2020 Virological assessment of hospitalized patients with COVID-2019. *Nature* **581**, 465–469. (doi:10.1038/s41586-020-2196-x)
58. To KK *et al.* 2020 Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect. Dis.* **20**, 565–574. (doi:10.1016/S1473-3099(20)30196-1)
59. Verity R *et al.* 2020 Estimates of the severity of coronavirus disease 2019: a model-based analysis.

- Lancet Infect. Dis.* **20**, 669–677. (doi:10.1016/S1473-3099(20)30243-7)
60. Tindale LC *et al.* 2020 Evidence for transmission of COVID-19 prior to symptom onset. *eLife* **9**, e57149. (doi:10.7554/eLife.57149)
  61. Rees EM *et al.* 2020 COVID-19 length of hospital stay: a systematic review and data synthesis. *BMC Med.* **18**, 1–22. (doi:10.1186/s12916-019-1443-1)
  62. Gaythorpe K, Imai N, Cuomo-Dannenburg G, Baguelin M, Bhatia S, Boonyasiri A 2020 *Report 8: Symptom progression of COVID-19*. London, UK: Imperial College London.
  63. U.S. Census Bureau. Small Area Health Insurance Estimates (SAHIE) Program. See <https://www.census.gov/programs-surveys/sahie.html>.
  64. Burhenne S, Jacob D, Henze GP. 2011 Sampling based on Sobol sequences for Monte Carlo techniques applied to building simulations. In *Proc. of Building Simulation 2011: 12th Conf of Int. Building Performance Simulation Assoc., Sydney, Australia, 14–16 November*, pp. 1816–1823. See [http://www.ibpsa.org/proceedings/BS2011/P\\_1590.pdf](http://www.ibpsa.org/proceedings/BS2011/P_1590.pdf).
  65. Perkins TA, Cavany SM, Moore SM, Oidtmann RJ, Lerch A, Poterek M. 2020 Estimating unobserved SARS-CoV-2 infections in the United States. *Proc. Natl Acad. Sci. USA* **117**, 22 597–22 602. (doi:10.1073/pnas.2005476117)
  66. Hu Z *et al.* 2020 Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China. *Sci. China Life Sci.* **63**, 706–711. (doi:10.1007/s11427-020-1661-4)
  67. He D, Zhao S, Lin Q, Zhuang Z, Cao P, Wang MH, Yang L. 2020 The relative transmissibility of asymptomatic cases among close contacts. *Int. J. Infect. Dis.* **63**, 706–711. (doi:10.2139/ssrn.3565270)
  68. Ye F, Xu S, Rong Z, Xu R, Liu X, Deng P, Liu H, Xu X. 2020 Delivery of infection from asymptomatic carriers of COVID-19 in a familial cluster. *Int. J. Infect. Dis.* **94**, 133–138. (doi:10.1016/j.ijid.2020.03.042)
  69. Gudbjartsson DF *et al.* 2020 Spread of SARS-CoV-2 in the Icelandic population. *N. Engl. J. Med.* **382**, 2302–2315. (doi:10.1056/NEJMoa2006100)
  70. Li LQ, Huang T, Wang YQ, Wang ZP, Liang Y, Huang TB, Zhang HY, Sun W, Wang Y. 2020 COVID-19 patients' clinical characteristics, discharge rate, and fatality rate of meta-analysis. *J. Med. Virol.* **92**, 577–583. (doi:10.1002/jmv.25757)
  71. Descartes Labs. Mobility changes in response to COVID-19. See <https://github.com/descarteslabs/DL-COVID-19>.
  72. Flaxman S *et al.* 2020 Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 257–261. (doi:10.1038/s41586-020-2405-7)
  73. Bi Q *et al.* 2020 Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *Lancet Infect. Dis.* **20**, 911–919. (doi:10.1016/S1473-3099(20)30287-5)
  74. King A, Ionides E, Breto C, Ellner S, Ferrari M, Kendall B *et al.* 2020 pomp: statistical inference for partially observed Markov Processes. See <https://kingaa.github.io/pomp/>.
  75. R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2020. See <https://www.R-project.org>.
  76. Wecker WE. 1989 Comment: assessing the accuracy of time series model forecasts of count observations. *J. Bus. Econ. Stat.* **7**, 418–419. (doi:10.1080/07350015.1989.10509751)
  77. Selten R. 1998 Axiomatic characterization of the quadratic scoring rule. *Exp. Econ.* **1**, 43–61. (doi:10.1023/A:1009957816843)
  78. Gneiting T, Raftery AE. 2007 Strictly proper scoring rules, prediction, and estimation. *J. Am. Stat. Assoc.* **102**, 359–378. (doi:10.1198/01621450600001437)
  79. Czado C, Gneiting T, Held L. 2009 Predictive model assessment for count data. *Biometrics* **65**, 1254–1261. (doi:10.1111/j.1541-0420.2009.01191.x)
  80. County of Santa Clara. COVID-19 testing dashboards. See <https://covid19.sccgov.org/dashboard-testing>.
  81. King AA, Nguyen D, Ionides EL. 2020 Statistical inference for partially observed Markov processes via the R Package pomp. *J. Stat. Softw.* **69**, 1–43.
  82. Li Q *et al.* 2020 Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **382**, 1199–1207. (doi:10.1056/NEJMoa2001316)
  83. Santa Clara Emergency Operations Center. COVID-19 cases and deaths; 2020. See <https://www.sccgov.org/sites/covid19/Pages/dashboard-cases.aspx>.
  84. James N, Menzies M, Radchenko P. 2021 COVID-19 second wave mortality in Europe and the United States. *Chaos* **31**, 031105. (doi:10.1063/5.0041569)
  85. Doornik JA, Castle JL, Hendry DF. In press. Short-term forecasting of the coronavirus pandemic. *Int. J. Forecast.* (doi:10.1016/j.ijforecast.2020.09.003)
  86. Kraemer MU *et al.* 2020 The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497. (doi:10.1126/science.abb4218)
  87. Glaeser EL, Gorbach C, Redding SJ. 2020 How much does COVID-19 increase with mobility? Evidence from New York and four other U.S. cities. National Bureau of Economic Research. 27519. See <http://www.nber.org/papers/w27519>.
  88. Grantz KH *et al.* 2020 The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology. *Nat. Commun.* **11**, 1–8. (doi:10.1038/s41467-020-18190-5)
  89. Zhou Y, Xu R, Hu D, Yue Y, Li Q, Xia J. 2020 Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: a modelling study using mobile phone data. *Lancet Digital Health* **2**, e417–e424. (doi:10.1016/S2589-7500(20)30165-5)
  90. Weitz JS, Park SW, Eksin C, Dushoff J. 2020 Awareness-driven behavior changes can shift the shape of epidemics away from peaks and toward plateaus, shoulders, and oscillations. *Proc. Natl Acad. Sci. USA* **117**, 32 764–32 771. (doi:10.1073/pnas.2009911117)
  91. Drake JM, Dahlin K, Rohani P, Handel A. 2021 Five approaches to the suppression of SARS-CoV-2 without intensive social distancing. *Proc. R. Soc. B* **288**, 20203074. (doi:10.1098/rspb.2020.3074)
  92. Park SW, Pons-Salort M, Messacar K, Cook C, Meyers L, Farrar J, Grenfell BT. 2021 Epidemiological dynamics of enterovirus D68 in the United States and implications for acute flaccid myelitis. *Sci. Transl. Med.* **13**, eabd2400. (doi:10.1126/scitranslmed.abd2400)
  93. Ben-Nun M, Riley P, Turtle J, Bacon DP, Riley S. 2019 Forecasting national and regional influenza-like illness for the USA. *PLoS Comput. Biol.* **15**, 1–20. (doi:10.1371/journal.pcbi.1007013)
  94. Hadfield J, Bedford T, Neher R, Hodcroft E. 2018 Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123. (doi:10.1093/bioinformatics/bty407)
  95. MIDAS Coordination Center. 2020 MIDAS Online Portal for COVID-19 Modeling Research. See <https://midasnetwork.us/covid-19/>.
  96. Johns Hopkins University Center for Systems Science and Engineering. 2019 COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. See <https://github.com/CSSEGISandData/COVID-19>.
  97. NY Times. Coronavirus (Covid-19) Data in the United States. See <https://github.com/nytimes/covid-19-data>.
  98. Bjørnstad ON, Finkenstädt BF, Grenfell BT. 2002 Dynamics of measles epidemics: estimating scaling of transmission rates using a time series SIR model. *Ecol. Monogr.* **72**, 169–184. (doi:10.1890/0012-9615(2002)072[0169:DOMEES]2.0.CO;2)
  99. Yang W, Cowling BJ, Lau EH, Shaman J. 2015 Forecasting influenza epidemics in Hong Kong. *PLoS Comput. Biol.* **11**, e1004383. (doi:10.1371/journal.pcbi.1004383)
  100. Gnanvi JE, Salako KV, Kotanmi GB, Glèlè Kakai R. 2021 On the reliability of predictions on COVID-19 dynamics: a systematic and critical review of modelling techniques. *Infect. Dis. Model.* **6**, 258–272. (doi:10.1016/j.idm.2020.12.008)
  101. Borchering RK *et al.* 2021 Modeling of future COVID-19 cases, hospitalizations, and deaths, by vaccination rates and nonpharmaceutical intervention scenarios—United States, April–September 2021. *Morb. Mortal. Wkly Rep.* **70**, 719. (doi:10.15585/mmwr.mm7019e3)
  102. Santa Clara Emergency Operations Center. COVID-19 Vaccination Coverage; 2021. See <https://www.sccgov.org/sites/covid19/Pages/dashboard-vaccine-CAIR2.aspx>.
  103. Childs ML *et al.* 2021 Data from: The impact of long-term non-pharmaceutical interventions on COVID-19 epidemic dynamics and control: the value and limitations of early models. Dryad Digital Repository. (doi:10.5061/dryad.cvdncjt4t)