Epidemics 34 (2021) 100430

Contents lists available at ScienceDirect

Epidemics

journal homepage: www.elsevier.com/locate/epidemics

Chopping the tail: How preventing superspreading can help to maintain COVID-19 control

Morgan P. Kain [a,b,1,*], Marissa L. Childs [c,1,**], Alexander D. Becker [a], Erin A. Mordecai [a]

[a] Department of Biology, Stanford University, Stanford, CA, 94305, USA
[b] Natural Capital Project, Woods Institute for the Environment, Stanford University, Stanford, CA 94305, USA
[c] Emmett Interdisciplinary Program in Environment and Resources, Stanford University, Stanford, CA, 94305, USA

ARTICLE INFO

Dataset link: https://github.com/nytimes/covid-19-data, https://github.com/morgankain/COVID_interventions

ABSTRACT

Disease transmission is notoriously heterogeneous, and SARS-CoV-2 is no exception. A skewed distribution where few individuals or events are responsible for the majority of transmission can result in explosive, superspreading events, which produce rapid and volatile epidemic dynamics, especially early or late in epidemics. Anticipating and preventing superspreading events can produce large reductions in overall transmission rates. Here, we present a stochastic compartmental (SEIR) epidemiological model framework for estimating transmission parameters from multiple imperfectly observed data streams, including reported cases, deaths, and mobile phone-based mobility that incorporates individual-level heterogeneity in transmission using previous estimates for SARS-CoV-1 and SARS-CoV-2. We parameterize the model for COVID-19 epidemic dynamics by estimating a time-varying transmission rate that incorporates the impact of non-pharmaceutical intervention strategies that change over time, in five epidemiologically distinct settings—Los Angeles and Santa Clara Counties, California; Seattle (King County), Washington; Atlanta (Dekalb and Fulton Counties), Georgia; and Miami (Miami-Dade County), Florida. We find that the effective reproduction number ($\mathcal{R}_E$) dropped below 1 rapidly in all five locations following social distancing orders in mid-March, 2020, but that gradually increasing mobility starting around mid-April led to an $\mathcal{R}_E$ once again above 1 in late May (Los Angeles, Miami, and Atlanta) or early June (Santa Clara County and Seattle). However, we find that increased social distancing starting in mid-July in response to epidemic resurgence once again dropped $\mathcal{R}_E$ below 1 in all locations by August 14. We next used the fitted model to ask: how does truncating the individual-level transmission rate distribution (which removes periods of time with especially high individual transmission rates and thus models superspreading events) affect epidemic dynamics and control? We find that interventions that truncate the transmission rate distribution while partially relaxing social distancing are broadly effective, with impacts on epidemic growth on par with the strongest population-wide social distancing observed in April, 2020. Given that social distancing interventions will be needed to maintain epidemic control until a vaccine becomes widely available, "chopping off the tail" to reduce the probability of superspreading events presents a promising option to alleviate the need for extreme general social distancing.

1. Introduction

In the face of emerging epidemics with limited pharmaceutical options for treatment and prevention of infection, non-pharmaceutical interventions such as social distancing are critical for slowing epidemic growth. Shelter-in-place and other social distancing orders have helped to slow the pace of the COVID-19 pandemic, reducing the effective reproduction number $\mathcal{R}_E$ – or the number of secondary infections produced by each infected person – to one or below in most places. In doing so, social distancing has effectively kept most regional healthcare systems operating under maximum capacity. However, after only a few weeks of declining numbers of daily cases due to an $\mathcal{R}_E$ at or below one in the spring of 2020, most state and county governments in the United States began relaxing social distancing orders, citing their major economic impacts. In order to avoid epidemic resurgence, it is vitally important that governments employ long-term strategies that maintain epidemic control while economic reopening commences.

* Corresponding author at: Department of Biology, Stanford University, Stanford, CA, 94305, USA.
** Corresponding author.
E-mail addresses: morganpkain@gmail.com, kainm@stanford.edu (M.P. Kain), marissac@stanford.edu (M.L. Childs).
[1] Denotes equal authorship.

https://doi.org/10.1016/j.epidem.2020.100430
Received 30 June 2020; Received in revised form 4 December 2020; Accepted 9 December 2020
Available online 21 December 2020

One obstacle to designing effective long-term strategies is a notoriously heterogeneous transmission process. It is now widely recognized that the minority of infections generate the majority of secondary cases, leading to the so-called 20–80 rule in epidemiology (the rule-of-thumb that 20% of people generate 80% of cases) (Woolhouse et al., 1997). Work on SARS-CoV-1, measles, and other respiratory viruses suggests that this skew in secondary cases is even larger (Lloyd-Smith et al., 2005). This heterogeneity gives rise to events in which a single infected person transmits a disease to dozens or hundreds of people – called superspreading events – which have played an important role in the COVID-19 pandemic (Adam et al., 2020; Kumar et al., 2020; Liu et al., 2020; Althouse et al., 2020; Hébert-Dufresne et al., 2020). Indeed, the frequency of asymptomatic and presymptomatic transmission, potential disconnect between infection and clinical presentation (Davies et al., 2020), and potential transmission via direct contact, aerosols, and surfaces (World Health Organization, 2020; Guo et al., 2020) are all features of SARS-CoV-2 that tend to promote superspreading. As local and national governments search for viable alternatives to shelter-in-place, a critical question is how effective curtailing superspreading events could be in controlling epidemic spread.

Practically, one strategy to help prevent superspreading is to prohibit medium to large indoor gatherings such as exercise classes, sporting events, concerts, and weddings for an extended period after allowing smaller and lower-risk activities to resume. From a modeling standpoint, predicting the effects of this straightforward intervention is difficult for two reasons: (1) local epidemiological dynamics are changing with evolving intervention strategies; and (2) information may not be available to parameterize detailed models of disease spread through heterogeneous populations. Despite these difficulties, it is important to consider some individual-level heterogeneity in transmission because model analyses of mean transmission rates alone may over-estimate the effectiveness of interventions, overlook potentially effective interventions that act on the heterogeneity within populations, miss potentially explosive resurgences, and poorly predict the final epidemic size (Lloyd-Smith et al., 2005; Hébert-Dufresne et al., 2020).

Studies of superspreading often empirically estimate secondary case distributions from recorded transmission chains and/or using branching process models (Lloyd-Smith et al., 2005; Liu et al., 2020; Althouse et al., 2020; Zhang et al., 2020b). These studies estimate a dispersion parameter, k, that describes the variance in secondary cases based on a Negative Binomial distribution, where smaller values indicate more heterogeneity and skew and large values approach a Poisson distribution. Estimated k values for SARS-CoV-2 remain uncertain, but are thought to range from 0.04–0.3 (Althouse et al., 2020; Hébert-Dufresne et al., 2020; Zhang et al., 2020b; Endo et al., 2020a), similar to the estimate of 0.16 for SARS-CoV-1 (Lloyd-Smith et al., 2005), which we use for this analysis. These empirical and branching process approaches are ideal for characterizing heterogeneity in secondary cases, but not for projecting epidemic trajectories through time, without being further embedded in a compartmental or network modeling framework.

Here, we present a mechanistic susceptible, exposed, infectious, removed (SEIR) model that uses data on cases, deaths, and mobility for parameter estimation, incorporates heterogeneity in transmission rates, and is realistic enough to be useful for scenario exploration but simple enough to be adapted to a wide range of settings. The key innovation in our model is in using the average of Gamma-distributed *individual* transmission rates at each time step, as supported by previous work on secondary case distributions, to generate the distribution of *population-average* transmission rates. This formulation allows us to both generate more realistic variation in trajectories than SEIR models that assume a single average transmission rate, and explore and quantify the impact of altering individual-level transmission distributions on population-level dynamics without more detailed information on contact networks, age structure, or other social information.

The model, with accompanying open-access code, can be used to fit to any county in the U.S. using publicly available data; here we focus on five contrasting epidemiological settings—Seattle (King County), Washington; Los Angeles (Los Angeles County), California; Santa Clara County, California; Atlanta (Dekalb and Fulton Counties), Georgia; and Miami (Miami-Dade County), Florida. For each location we estimate a time-varying effective reproduction number, $\mathcal{R}_E$, which represents the average number of secondary infections produced by each infected person, and is an important (though imperfect (Hébert-Dufresne et al., 2020)) metric of epidemic control. Using each fitted model, we truncate the individual-level transmission rate distribution and stochastically simulate epidemic dynamics into the future, representing a scenario where high-risk events are eliminated but smaller and lower-risk activities are allowed to resume. This model construction focuses on reducing superspreading events in which a single individual infects many others over a short time period (e.g., due to simultaneous contact with many susceptible people); it does not address situations where individuals act as superspreaders because they have a longer-than-average infected duration. We investigate the absolute impact of this superspreading prevention strategy on epidemic control, and compare its impact on epidemic dynamics (and $\mathcal{R}_E$) to test-and-isolate and strong social distancing (e.g., shelter-in-place) interventions. Using this comparison we highlight non-pharmaceutical intervention strategies that are less extreme than shelter-in-place but still expected to reduce both epidemic growth (i.e., keep $\mathcal{R}_E$ below one) and the probability of explosive resurgence.

2. Methods

2.1. Model structure

We developed a stochastic compartmental model using an SEIR (Susceptible, Exposed, Infectious, Recovered) framework to model COVID-19 transmission, which was first described in Childs et al. (2020). Our model divides the population into the following classes: susceptible (S); exposed but not yet infectious (E); infectious and presymptomatic (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 assume an underlying, unobserved process model of SARS-CoV-2 transmission described by Eqs. (1)–(10) and shown in Figure S1, where each term $d_{X,Y}$ denotes the transition from compartment X to Y. Eqs. (11)–(18) describe in detail the stochastic rates used to approximate the transition terms in Eqs. (1)–(10). Transitions between compartments are simulated as binomial ($\mathcal{B}$) or multinomial ($\mathcal{M}$) processes to account for competing rates (implemented with the `rbinom` and `reulermultinom` functions, respectively, in the `pomp` package (King et al., 2016)). We use an Euler approximation of the continuous time process with a time step of 4 h. To produce more realistic latent and infectious periods we divide each infectious class and the exposed period into multiple sub-stages (not shown in the equations below), which results in Erlang distributed periods within stages (Anderson and Watson, 1980; Lloyd, 2001). Specifically, we use three sub-stages for the exposed class, seven sub-stages for the asymptomatic infectious class, two sub-stages for the presymptomatic infectious class, five sub-stages for the mildly symptomatic infectious class, and five sub-stages for the severely symptomatic class. We translate durations into rates for our model with sub-classes and a Euler approximation using the method described in He et al. (2010).

$$\frac{\Delta S}{\Delta t} = -d_{S,E} \tag{1}$$

$$\frac{\Delta E}{\Delta t} = d_{S,E} - d_{E,I_A} - d_{E,I_P} \tag{2}$$

$$\frac{\Delta I_A}{\Delta t} = d_{E,I_A} - d_{I_A,R} \tag{3}$$

$$\frac{\Delta I_P}{\Delta t} = d_{E,I_P} - d_{I_P,I_S} - d_{I_P,I_M} \tag{4}$$

$$\frac{\Delta I_M}{\Delta t} = d_{I_P,I_M} - d_{I_M,R} \tag{5}$$

$$\frac{\Delta I_S}{\Delta t} = d_{I_P,I_S} - d_{I_S,H_R} - d_{I_S,H_D} \tag{6}$$

$$\frac{\Delta H_R}{\Delta t} = d_{I_S,H_R} - d_{H_R,R} \tag{7}$$

$$\frac{\Delta H_D}{\Delta t} = d_{I_S,H_D} - d_{H_D,D} \tag{8}$$

$$\frac{\Delta R}{\Delta t} = d_{I_A,R} + d_{I_M,R} + d_{H_R,R} \tag{9}$$

$$\frac{\Delta D}{\Delta t} = d_{H_D,D} \tag{10}$$

By including asymptomatic and presymptomatic individuals, we are able to track "silent spreaders" of the disease, which have been shown to contribute to COVID-19 transmission (Ferretti et al., 2020; Li et al., 2020b). Mildly symptomatic cases are defined as those people that show symptoms but do not require hospitalization. We assume that all severely symptomatic cases will eventually require hospitalization and that no onward transmission occurs from hospitalized individuals.

$$d_{S,E} \sim \mathcal{B}\left(S, 1 - \exp\left(-\beta_t \frac{\kappa_A I_A + \kappa_P I_P + \kappa_M I_M + \kappa_S I_S}{N} \Delta t\right)\right) \tag{11}$$

$$\begin{pmatrix} d_{E,I_A} \\ d_{E,I_P} \end{pmatrix} \sim \mathcal{M}\left(E, \begin{pmatrix} \alpha(1 - \exp(-\gamma \Delta t)) \\ (1-\alpha)(1 - \exp(-\gamma \Delta t)) \end{pmatrix}\right) \tag{12}$$

$$d_{I_A,R} \sim \mathcal{B}(I_A, 1 - \exp(-\lambda_A \Delta t)) \tag{13}$$

$$\begin{pmatrix} d_{I_P,I_M} \\ d_{I_P,I_S} \end{pmatrix} \sim \mathcal{M}\left(I_P, \begin{pmatrix} \mu(1 - \exp(-\lambda_P \Delta t)) \\ (1-\mu)(1 - \exp(-\lambda_P \Delta t)) \end{pmatrix}\right) \tag{14}$$

$$d_{I_M,R} \sim \mathcal{B}(I_M, 1 - \exp(-\lambda_M \Delta t)) \tag{15}$$

$$\begin{pmatrix} d_{I_S,H_R} \\ d_{I_S,H_D} \end{pmatrix} \sim \mathcal{M}\left(I_S, \begin{pmatrix} \delta(1 - \exp(-\lambda_S \Delta t)) \\ (1-\delta)(1 - \exp(-\lambda_S \Delta t)) \end{pmatrix}\right) \tag{16}$$

$$d_{H_R,R} \sim \mathcal{B}(H_R, 1 - \exp(-\rho_R \Delta t)) \tag{17}$$

$$d_{H_D,D} \sim \mathcal{B}(H_D, 1 - \exp(-\rho_D \Delta t)) \tag{18}$$

The time-varying transmission parameter, β_t, describes the average per capita rate of contact between susceptible and infectious people at time t, multiplied by the per-contact transmission probability. We modeled β_t as a function of human movement using the scaling function:

$$\beta_t = \beta_0 \beta_m^{\theta}, \tag{19}$$

which treats β_t as an exponentially decreasing function of physical distancing (θ; on a scale of 0–1 where 0 is no physical distancing, and 1 is maximum physical distancing). Here, $\beta_0\beta_m$ is the estimated minimum possible transmission rate given minimal human movement (i.e., maximal physical distancing) and thus contact rate, while β_0 is the estimated transmission rate given maximal human movement (i.e., no physical distancing). To measure physical distancing we use SafeGraph's "Shelter in Place Index" (Shelter in Place Index, 2020), which calculates the proportion of cell phone devices that are staying home.

To model individual heterogeneity in SARS-CoV-2 transmission rate, we allow individuals to vary over time by modeling an individual's transmission rate in each time step as a Gamma distributed random variable with a dispersion such that the sum of an individual's transmission rates over the duration of their infection approximates a Gamma distributed random variable with dispersion equal to previous Negative Binomial parameterizations for SARS-CoV-1 ($k = 0.16$) (Lloyd-Smith et al., 2006), which closely approximates estimates of overdispersion for SARS-CoV-2 (Endo et al., 2020a; Grantz et al., 2020). Because we model the transmission rates as the multiplication of contact rate and infection probability, this heterogeneity implicitly considers both variation among individuals in infectiousness and contact rate, and can be thought of as modeling superspreading *periods* or *events*—windows in time when an infected individual has a particularly high transmission rate. Some other forms of superspreading, such as when individuals have unusually long infectious durations, are not captured by this formulation. Additionally, this formulation assumes that infectious individuals' transmission rates are independent draws, while events that involve large aggregations of people inherently imply dependence. For relatively small superspreading events likely to be driven by one infectious individual, this independence assumption will have no effect. However, our model may underestimate the occurrence of the most extreme superspreading events, which could be characterized by multiple superspreading events occurring in the same 4h time window and would require multiple independent samples in the tail of the individual-level gamma distributions. To incorporate individual variation into an average time step, we model the population level transmission rate β_t as the average of the transmission of all infected individuals at time t. To do so we apply the property of Gamma distributions that the mean and variance of N samples from a Gamma distribution with defined rate and scale is itself a Gamma distribution with mean equal to that of the original Gamma distribution and variance equal to the variance of the original Gamma divided by N. A full derivation of the equivalence between the individual time step transmission rate distributions (which we will hereafter refer to with π), the individual infectious period transmission rate distributions, and the population-level transmission rate distribution is presented in the Appendix A.

We assume that observed deaths are a Negative Binomial random variable with a mean equal to total new deaths accumulated over the observation period (i.e., one day for this analysis), and a dispersion parameter that we estimate. We also assume that daily observed cases are a Negative Binomial random variable, but have a mean equal to the daily accumulated number of new symptomatic infections multiplied by a daily detection probability that we estimate from the data. We model daily detection probability as a monotonically increasing logistic function:

$$\frac{\omega}{1 + e^{-m(t-\phi)}}, \tag{20}$$

where ω is the maximum fraction of symptomatic cases detected, m is the logistic growth rate, and ϕ gives the location of the inflection point (where the probability of detection equals one half of the maximum detection probability, ω). Because ϕ can be estimated to be in the future, the probability of detection of an infected case in the present can be any value between 0 and ω. We estimate newly observed cases to be a fraction of all new symptomatic infections at time t. Though this ignores testing asymptomatic infections, any detection of asymptomatic infections will be captured as a higher estimated detection fraction of symptomatic infections.

2.2. Fitting the model

We used COVID-19 death and case data from The New York Times, based on reports from state and local health agencies (available at https://github.com/nytimes/covid-19-data). Using these data, which are available for all counties in the US, and any form of human movement data that can be scaled to 0–1, our model can be used to fit infection dynamics in any county.

We sampled over uncertainty in a series of parameters (Tables 2, 3) for which there was large uncertainty in the literature. To do so, we drew 800 sobol sequences, an efficient method for sampling input parameters (Burhenne et al., 2011), across a range of plausible values for each parameter in order to form 800 plausible parameter sets. For computational efficiency we assumed point estimates for other parameters (Table 1). These parameters most notably include the average duration individuals spend within infectious categories, for which some consensus has been reached in the COVID-19 modeling literature. While the impact of changes in these assumed means would be somewhat dampened because of our use of Erlang distributed periods (which leads to variation in transition times), more work could be done to precisely identify the sensitivity of model results to these point estimates. We also assumed equivalence in the relative infectiousness of each infected

Table 1
Parameters for which we used point estimates and their values. These parameters are assumed to not vary by location.

Parameter	Value	Description	Estimates and sources
κ_P, κ_M, κ_S	1	Relative infectiousness of presymptomatic, mild symptomatic, and severe symptomatic	Assumed
γ	3.5 days	Preinfectious period	One meta-analysis (McAloon et al., 2020) found a mean incubation period of 5.8 days, another found a median of 5.1 days (Lauer et al., 2020). We use a shortened duration because we assume 2 days of presymptomatic transmission
λ_P	2 days	Presymptomatic duration	Range of 1–3 days (Wei et al., 2020), mean of 3.8 days (Zhang et al., 2020a), viral shedding estimated to begin 2.3 days prior to symptom onset (He et al., 2020a), many articles find presymptomatic infection is likely but do not estimate duration (Zhang, 2020; Gao and Li, 2020; Arons et al., 2020; Lee et al., 2020)
λ_A	7 days	Infectious period for asymptomatic infections	Mean seroconversion after 7 days (Wölfel et al., 2020)
λ_S	5 days	Time from symptom onset to hospitalizations (severe cases)	Mean of 5.5 days (Sanche et al., 2020), median of 4 days (To et al., 2020), mean of 5.6 days (Verity et al., 2020)
λ_M	5.0 days	Time from symptom onset to recovery (mild cases)	Infectiousness based on viral shedding estimated to decline substantially within 7 days (He et al., 2020a; Wölfel et al., 2020) *Note, He et al. (2020a) takes samples from hospitalized patients; we assume similar viral shedding in mild infections
ρ_R	13.3 days	Time from hospitalization to recovery	Mean of 13.3 (Tindale et al., 2020), mean of 20.51 (Verity et al., 2020), highly variable by region (Rees et al., 2020)

Table 2
Parameters for which we used a range of values and those ranges. Parameter ranges are assumed to not vary by location.

Parameter	Lower bound	Upper bound	Description	Estimates and sources
κ_A	0.4	0.8	Relative infectiousness of asymptomatic infections	0.6 (Perkins et al., 2020), few direct estimates, but many examples of asymptomatic transmission potential less than, but potentially close to that of, symptomatic infected individuals (He et al., 2020b; Ye et al., 2020; Hu et al., 2020)
α	0.3	0.5	Proportion of infections that are asymptomatic	Mean of 43.3% (Lavezzo et al., 2020), 44% (Gudbjartsson et al., 2020)
δ	0.1	0.3	Fatality rate among hospitalizations	Demographic weighted average that will vary by location, see Verity et al. (2020), 5% (Li et al., 2020a)
ρ_D	13 days	20 days	Time from hospitalization to death	Mean of 16 days (Gaythorpe et al., 2020), mean of 17.8 days (Verity et al., 2020), highly variable by region (Rees et al., 2020)
$1-\mu$	0.025	0.075	Proportion of symptomatic infections that require hospitalization	Demographic weighted average that will vary by location, see Verity et al. (2020)

Table 3
Location-specific parameter range estimates. Population sizes obtained from the US Census Bereau (County Population Totals 2010-2019, 2019).

Parameter	Santa Clara county, CA	Los Angeles County, CA	Miami-Dade County, FL	King County, WA	Fulton and DeKalb Counties, GA
Epidemic start date	01-Jan–05-Feb	01-Jan–31-Jan	01-Jan–16-Mar	01-Jan–04-Mar	01-Jan–Mar-07
Population size	1,927,852	10,039,107	2,716,940	2,252,782	1,755,830

category except completely asymptomatic infections, given a lack of any published estimates for these parameters. Our range of plausible epidemic start dates (Table 3) spanned a period from the first of the year to the last day before the first recorded case in each location.

For each parameter set we used the package `pomp` (King et al., 2016) in the statistical programming language R (R Core Team, 2020) to estimate the following parameters: β_0 and β_m, transmission rate parameters (see Eq. (19)); E_0, number of exposed individuals that initiate the epidemic; ω, m, and ϕ, maximum, slope, and inflection point day of the sigmoidal case detection function; θ_d, Negative Binomial dispersion parameter for deaths; and θ_c, Negative Binomial dispersion parameter for cases. We assume that all individuals are initially susceptible except for those who are exposed (E_0). We fit all parameters to daily deaths, cases, and mobility in four steps. First, for each of the 800 parameter sets from the sobol sequences, we used the `mif2` function in `pomp` to perform iterated filtering, beginning with random starting conditions for the remaining parameters that were fit and using 90 filtering iterations with 4000 particles. We then continued to fit the top 20% (i.e., 160) parameter sets with the highest log likelihoods for an additional 200 iterations using 4000 particles. Finally the top 10% of those parameter sets were filtered twice more using 200 iterations at 4000 particles. At the end of each of the four filtering steps, log-likelihoods were calculated ten times using 5000 particles to produce standard errors. From these four filtering sets, the top ten parameter sets by log likelihood were selected. In total, each county took approximately 24 h to fit using twenty cores.

We focus on $\mathcal{R}_E$ and refrain from drawing conclusions directly about, for example, the estimated case detection probability (ω), the effectiveness of physical distancing (θ), or infection fatality ratio (IFR), because of some identifiability issues due to correlated parameters (Figures S2–S6). We calculated $\mathcal{R}_E$ at each time t as estimated β_t times the median proportion of the population remaining susceptible on each day times the average infectiousness over an infection (as defined by our model structure). For model fits and parameter estimates (e.g., $\mathcal{R}_E$) shown in results figures, we combine two sources of uncertainty: parameter uncertainty, which we include by using the 10 parameter sets with the largest log likelihoods; and stochastic epidemic trajectory variation, which we include by simulating 500 epidemics for each parameter set. Taken together, this uncertainty captures the widest plausible range of epidemic dynamics given our model and data. To avoid implausibly wide variation we discard simulated epidemics that did not reach at least a total of 500 infected individuals.

2.3. Simulating epidemics under interventions

Any intervention type, intensity, or duration can be modeled using this framework and open-source code (available at https://github.com/morgankain/COVID_interventions) given that it can be written as a function that modifies either human movement or β_t (e.g., social distancing or a pharmaceutical intervention that reduces the probability of infection). Here we consider interventions that reduce the skew of the individual time step transmission rate distribution (π), and thus

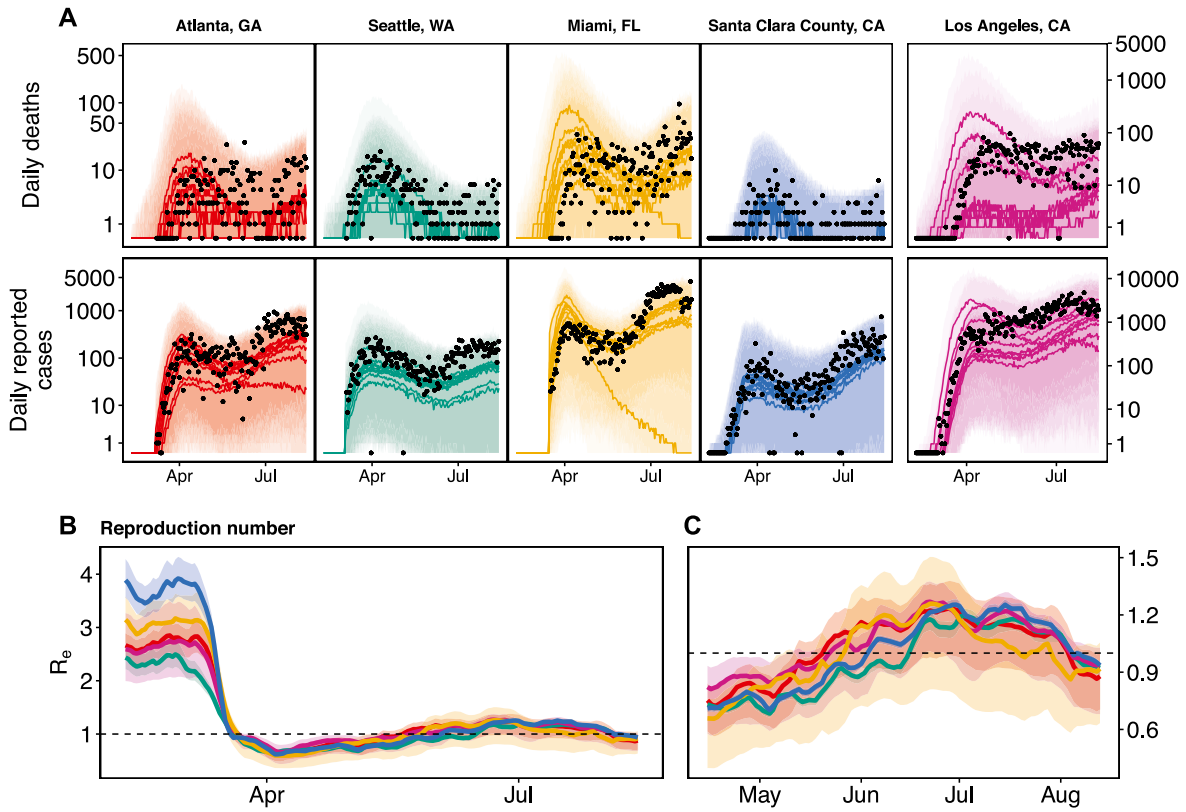

Fig. 1. Model estimated daily deaths and reported cases (A), and reproduction number (B, C) for five locations: Atlanta (red), Seattle (green), Miami (gold), Santa Clara County (blue), and Los Angeles (purple). Los Angeles is displayed on a different *y*-axis due to differences in magnitude of reported deaths and cases. For each county, we show the 10 model fits with the best log likelihoods. Panel C show the same results pictured in B, but are zoomed in to April 15–Aug 14 to better show the dynamics around $\mathcal{R}_E = 1$. Black points are observed daily deaths (A, top row) and reported cases (A, bottom row) in each county. Solid lines display median of model simulated trajectories (A) and mean of 7-day smoothed $\mathcal{R}_E$ (B, C). Ribbons overlay the central 95% of all stochastic simulations for each of the 10 best-fit parameter sets for epidemic dynamics (A) and full range of estimated $\mathcal{R}_E$ (B, C) over the 10 best-fit parameter sets. We highlight the large differences in stochastic epidemic trajectories for the single best parameter set in Figure S8. Vertical axes in panel A are pseudo-log transformed for visibility. Goodness of fit (R^2 values) and log likelihoods for each of these fits are given in Figure S9 and Table S1 respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the average time-varying transmission rate β_t; this is our mathematical representation of reducing highly infectious contact periods or events, which for COVID-19 tend to occur in crowded enclosed environments (e.g., church choirs and exercise classes). Specifically, we model truncation of the π distribution by assuming that all samples within the top X% of the π distribution are resampled. In particular, we consider truncation strategies that act in addition to moderate background social distancing measures. In doing so, we model the effects of social distancing as reducing the mean transmission rate (Eq. (19)) which still allows for individual heterogeneity in the resultant transmission rate. Truncation then operates to remove the highest transmission rates after considering background social distancing. Because of the skew of this distribution, truncating just the upper 0.1% of the β_t distribution can have a profound effect on the mean transmission rate (Figure S7). To visualize the dynamics of interventions, for each location we simulate 500 epidemics from the maximum likelihood estimate across the 800 parameter sets. The uncertainty band we plot represents the central 95% range of outcomes seen across all stochastic realizations that resulted in epidemics for this parameter set, and thus should not be taken as representation of uncertainty in parameter values or model structure.

3. Results

3.1. Epidemic trajectories

The model produced realistic fits to five contrasting epidemiological settings—King County, Washington; Los Angeles County, California; Santa Clara County, California; Fulton and Dekalb Counties, Georgia; and Miami-Dade County, Florida (hereafter, Seattle, Los Angeles, Santa Clara County, Atlanta, and Miami; Fig. 1A). Among these locations, we estimated that prior to interventions, $\mathcal{R}_0$ ranged between approximately 2 and 4 (Fig. 1B). We also estimated that $\mathcal{R}_E$ dropped below one following initial shelter-in-place orders in all counties, though only briefly in some locations. In particular, in Miami, Los Angeles, and Atlanta $\mathcal{R}_E$ climbed above 1 by mid-May, while in Seattle and Santa Clara County $\mathcal{R}_E$ remained below 1 until early June. After approximately another month of rising $\mathcal{R}_E$, increased social distancing led to a second decline in $\mathcal{R}_E$, which once again dropped below one in all locations by mid August (Fig. 1C).

From the ten best fitting parameter sets in each location, our model estimated moderate variation in infection fatality rate (IFR), case detection probability, and other model parameters among locations (Figure S10, Figure S11); however we caution against interpreting these parameter estimates in isolation because of their correlation with other parameters (Figures S2–S6). While our model is able to recreate plausible epidemic dynamics and sensible estimates for $\mathcal{R}_E$ (Fig. 1), model fitting indicated that there are a number of plausible parameter combinations that produce epidemic dynamics that match the observable data. As data continues to increase and more is learned about the disease, continued model fitting may be able to eliminate some of these alternative possibilities.

3.2. Interventions

As a basis for comparison, focusing on just three locations – Los Angeles, Miami, and Seattle – if all social distancing was simply to

end, a rapid increase in cases and deaths would be inevitable (Fig. 2, blue). However, non-pharmaceutical interventions, including continuing strong social distancing, infected isolation with intermediate levels of social distancing, or averting superspreading with intermediate levels of social distancing are capable of limiting epidemic growth (Fig. 2) and keeping $\mathcal{R}_E$ near or under 1. Here, we consider intermediate levels of social distancing that correspond to mobility levels that are an average of baseline mobility prior to the first wave of shelter-in-place orders and final mobility levels observed in the last week of data. Either an infected isolation strategy that reduces to intermediate levels of social distancing and catches 90% of all mild and severe cases of COVID-19 before they are able to transmit (Fig. 2, green), or a truncation strategy that similarly reduces to intermediate levels of social distancing and further removes the top 1% of the individual time step transmission rate distribution (π) with 75% efficiency (Fig. 2, purple) are able to suppress epidemic growth (and reduce $\mathcal{R}_E$ to below one) in all locations.

3.3. Curtailing superspreading

Limiting opportunities for superspreading by "chopping off the tail" of the contact rate or infectiousness distributions can be highly effective at epidemic control (Fig. 2), driving epidemic growth to be negative and bringing the average number of secondary cases ($\mathcal{R}_E$) below 1. An example truncation intervention is illustrated in Fig. 3: because the individual transmission rate distribution, π, over a 4-hour time period is so skewed (Fig. 3A; see Appendix A for derivation), truncating the upper 0.1% yields a large reduction in the mean and a moderate reduction in the variance of the population-level average transmission rate (Fig. 3B, shifting from red to blue distribution). A variety of possible truncation strategies exist, including eliminating varying proportions of π (e.g., upper 0.5%, upper 1%) with varying levels of efficiency (ranging from 50%–100%) (Figure S12).

An alternative measure of the impact of averting superspreading (i.e., truncation interventions) is how much general social distancing can be avoided by instead truncating the transmission rate distribution. Prior to social distancing orders, the estimated proportion sheltering in place (the proportion of cell phone devices staying home; hereafter, pSIP) ranged from ~20%–25% across our focal locations (Fig. 3C, triangles), indicating the baseline level of mobility. By increasing pSIP in combination with truncation interventions, a variety of combinations are predicted to provide epidemic control (for example, by reducing transmission rates such that $\mathcal{R}_E$ in a fully susceptible population would be 1; Fig. 3C). If the truncation intervention is 100% effective, truncating only approximately the upper 0.15% of individual transmission rates, π, (Fig. 3A) is effective enough to maintain transmission rates such that $\mathcal{R}_E$ would be 1 in a fully susceptible population, while allowing mobility levels to return to baseline (Fig. 3C). Alternatively, if truncation interventions are only half as effective, the same 0.15% truncation intervention would require moderate-strong social distancing (pSIP from ~30%–45%; Fig. 3C). The nonlinear effects of social distancing and truncation on transmission make the combination of interventions needed to maintain epidemic control sensitive to the efficiency and strength of each intervention mechanism.

3.4. Superspreading and epidemic resurgence

Even if the epidemic is brought almost entirely under control (e.g., to within 1–5 infected individuals remaining in the population), epidemic resurgence remains a possibility if interventions wane, allowing $\mathcal{R}_E$ to increase above one. As we show in (Fig. 3), many different combinations of pSIP and truncation can be used to produce the same $\mathcal{R}_E$ (in Fig. 3, an $\mathcal{R}_E$ of 1); however, epidemic dynamics will vary by combination because of the variation in individual time-step transmission rates, π. If $\mathcal{R}_E$ rises above one because interventions are relaxed, the specific combination of pSIP and truncation that remains in place will determine the resulting dynamics. Here we examine how different truncation interventions will affect epidemic extinction probability and the size of epidemic resurgence when it does not go extinct. We compare the full effect of truncation interventions (which influence both the mean of the transmission rate distribution and its shape) to the effects of truncation when $\mathcal{R}_E$ is held constant by scaling pSIP, reflecting only truncation effects on the shape of the transmission rate distribution (variance, skew, etc.).

With few infected individuals and $\mathcal{R}_E > 1$, stochasticity and heterogeneity in β_t can either lead to extinction, moderate resurgence, or explosive resurgence. Keeping interventions in place that remove even a tiny percent of the largest β_t values can help to avoid the more explosive events (Fig. 4). Truncation markedly reduces the probability of explosive epidemic resurgence (Fig. 4A) both by increasing the extinction probability (Fig. 4B) and by reducing the magnitude of resurgent epidemics when they do occur (Fig. 4C). While epidemic size was less sensitive to the number initially infected when resurgences do occur (Fig. 4C), the stochastic extinction probability was extremely sensitive to the difference between even one, three, or five remaining infections (Fig. 4B). Much, but not all, of the benefit of truncation comes from changing the mean transmission rate (and therefore $\mathcal{R}_E$). When $\mathcal{R}_E$ is held constant by adjusting pSIP, effects of truncation are more moderate. An increase in efficiency at truncating the top 0.1% of the β_t distribution noticeably decreases the number of infected 42 days after interventions are relaxed (Fig. 4D,F). However, because of the need to slightly reduce pSIP to hold $\mathcal{R}_E$ constant under truncation, truncation of π marginally decreases the extinction probability of the epidemic, which remains much more sensitive to the number initially infected (Fig. 4E). The highly stochastic nature of epidemic growth when cases are rare, combined with the fact that each truncation leaves behind highly skewed distributions regardless of the truncation parameters, results in even 10,000 epidemic simulations producing noisy patterns across intervention scenarios. Similar patterns are seen as more of the π distribution is truncated (Figure S13).

4. Discussion

Understanding local epidemiological dynamics of COVID-19 – and the impact of heterogeneity on those dynamics – remains a challenge due to both limited and imperfect data in most regions and ever evolving interventions and adherence. Reported cases are only a small fraction of all infections, and the proportion of symptomatic cases that are detected remains highly uncertain and variable over space and time. Our approach takes an important step toward capturing locally-specific epidemic dynamics and the impact of heterogeneity across settings by providing a platform (including a mathematical model and open access code) for estimating time-varying transmission rates (β_t) from death, mobility, and imperfectly observed case report data, all of which are publicly available. The model can estimate epidemic dynamics and transmission rates over time across epidemiological settings that vary in population size, demography, and control. By incorporating individual variation in contact rates (or, equivalently, infectiousness) into time step transmission rate distributions, we incorporate some of the known effects of heterogeneity without requiring detailed information on population mixing, structure, social networks, or movement patterns. We find that initial shelter-in-place measures in March of 2020 rapidly brought the average reproduction number—$\mathcal{R}_E$—from ~2–4 to below 1 in all locations we considered in early April. However, from early April through early July, relaxed social distancing led to a steady increase in $\mathcal{R}_E$ in all locations, which we estimate to have peaked around 1.2, with some variation among the locations we considered. In response to the second wave brought on by $\mathcal{R}_E > 1$, increased social distancing in all locations starting in early July once again successfully reduced $\mathcal{R}_E$ to just below one by mid August.

Non-pharmaceutical interventions have been and will continue to be necessary to control COVID-19 in all settings until better pharmaceutical options (in particular, effective vaccines) are widely available.

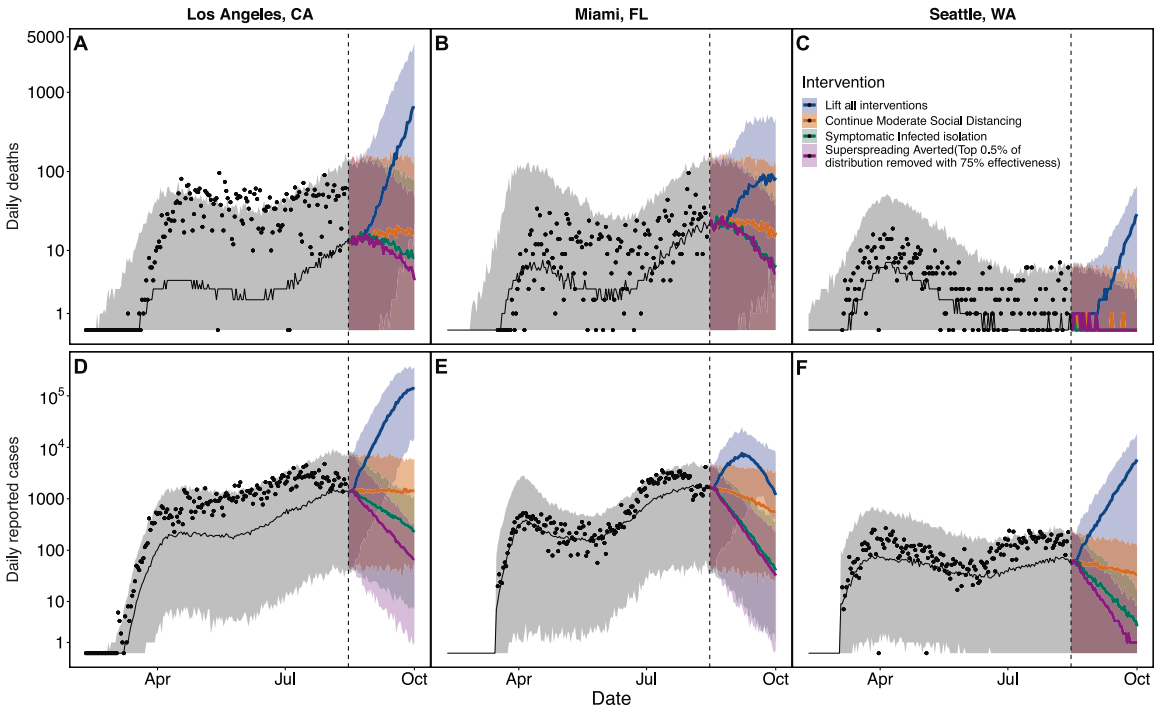

Fig. 2. Maintaining strong social distancing (orange), test-and-isolate with intermediate social distancing (green), or superspreading aversion with intermediate social distancing (purple) are necessary over long periods to prevent a major epidemic resurgence (blue) in each location where we fit our model. Daily deaths are shown in (A - C) and daily reported cases are shown in (D - F). Continuing social distancing at the levels seen as of August 14 (orange) will lead to nearly steady daily reported cases and deaths in Los Angeles, CA (A, D), and slow declines in cases and deaths in Miami, FL (B, E) and Seattle, WA (C, F) due to estimated R_e very near one. For both test-and-isolate and truncation interventions we assume an intermediate level of mobility (an average of baseline mobility prior to the first shelter-in-place orders and final mobility levels observed in the last week of data). Bands show the central 95% of stochastic simulations of daily cases and deaths for the single maximum likelihood parameter set and solid lines show the median among simulations. Dates range from February through October of 2020. Vertical axes are pseudo-log transformed for visibility. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

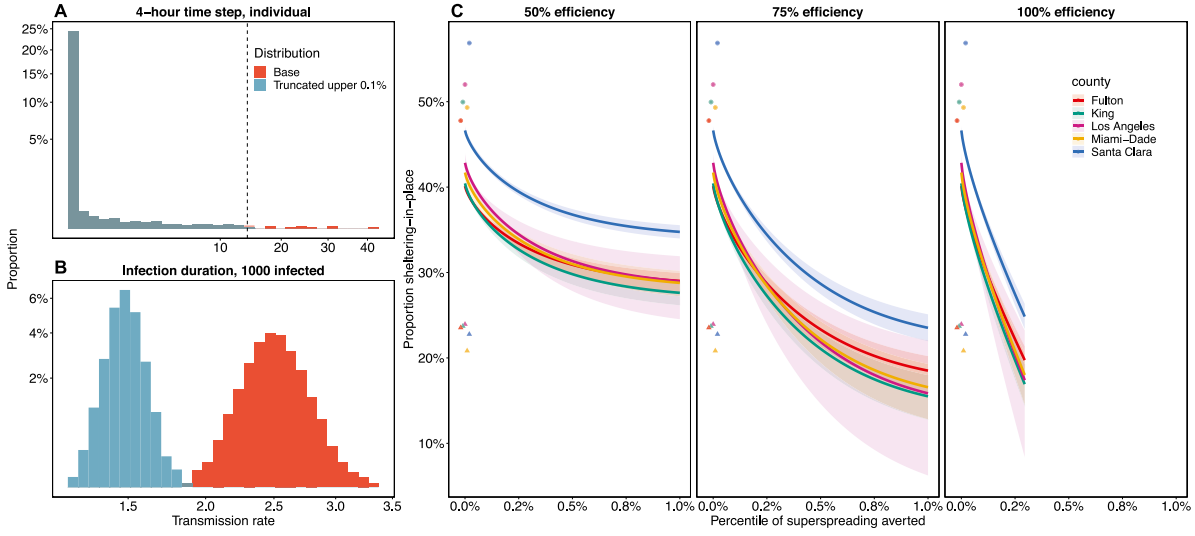

Fig. 3. Example of how truncating the individual-level transmission rate distribution, π, (A) affects the population-average transmission rate (B), and combinations of pSIP and truncation strategies that reduce $\mathcal{R}_E$ to one in a fully susceptible population (C). The three panels in C show the combinations of truncation and pSIP that produce an $\mathcal{R}_E$ of one for three levels of truncation efficiency. (A) Truncation at the upper 0.1% of π (sampled over a 4-hour time step), in which truncation occurs with 100% efficiency for values above the dashed line. (B) Resulting effect on the population-level average infection rate when there are 1000 infected people currently in the population, where the original distribution is in red and the truncated distribution is in blue. The distribution is shown over 10,000 simulations for a population characterized by an individual reproduction number distribution with mean of 2.5 and overdispersion parameter, $k = 0.16$. Horizontal and vertical axes in A and B are square root transformed for visibility. In C, the triangles show baseline pSIP in each location and circles show max pSIP reached during social distancing. Solid lines indicate the mean over the ten best fits, and the ribbon is the full range of estimates from these fits. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Social distancing in the general population is effective but costly: it is a blunt and imprecise tool. The social and economic necessity of relaxing social distancing demands safe long-term non-pharmaceutical intervention strategies based on more precise, targeted interventions to reduce transmission. Testing and isolating symptomatic people, combined with contact tracing, remains the gold standard intervention for limiting onward transmission as social distancing is lifted, but it is expensive and capacity remains limited in many settings. Our model

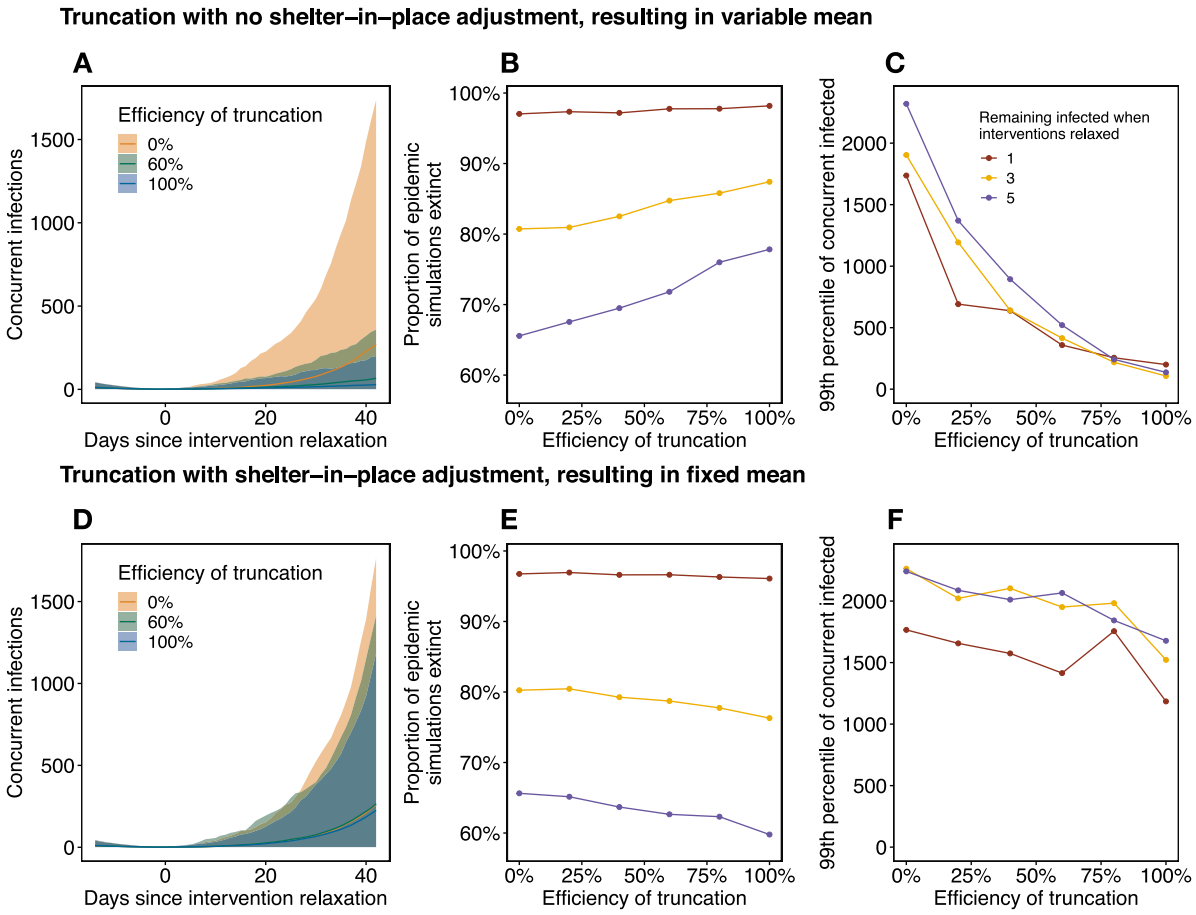

Fig. 4. Effects of transmission rate truncation on epidemic die-out and explosive resurgence. With skewed individual variation in transmission rate, relaxing social distancing interventions when infections become rare (allowing $\mathcal{R}_E$ to increase above one) may lead to explosive stochastic epidemic resurgence. Top panels (A–C) show the overall effect of truncation interventions, including effects on both the mean and shape of the transmission rate distribution, and resulting $\mathcal{R}_E$. Bottom panels (D–F) show the effect of truncation when $\mathcal{R}_E$ is held constant by rescaling pSIP at the time of intervention relaxation. Specifically, for a 0% truncation efficiency we simulate epidemic resurgence assuming $\mathcal{R}_0 = 2$, which results in an $\mathcal{R}_E = 2 \cdot S/N$ at the time of resurgence, which will vary by simulation (where S is the number of susceptible individuals and N is the total population size). In panels (A–C) as truncation efficiency increases $\mathcal{R}_E$ decreases; in panels (D–E) we scale pSIP to retain an average $\mathcal{R}_E = 2 \cdot S/N$ across truncations. Simulations are performed with varying efficiencies of truncation of the top 0.1% of the π distribution. Envelopes in (A) and (D) show the central 98% of resurgent simulations (across 10,000 total simulations) for three efficiencies of truncation (0% in orange, 60% in green, 100% in blue). The proportion of epidemic simulations that go extinct within 42 days of intervention relaxation for thresholds of 1 (red), 3 (gold), and 5 (blue) infected individuals is shown in (B) and (E). The upper 99th percentile of concurrent infections 42 days after intervention relaxation in resurgent simulations for the same thresholds is shown in (C) and (F).

shows that it is possible to target interventions even without precise information on specific population mobility, mixing, and infectiousness patterns, by limiting just the most high-risk activities. Specifically, we find that our method of truncating the upper percentiles of individual-level transmission rates, which models a restriction in the ability of individuals to contact a large number of susceptible people in a short period of time (such as by restricting their access to large gatherings and indoor events that have many close contacts), and thereby reduces population-level transmission rates, can be highly effective at maintaining epidemic control (Fig. 2), particularly when combined with mild to moderate social distancing (Fig. 3). Importantly, even after epidemic control is achieved and case numbers drop very low, "chopping off the tail" can provide powerful insurance against explosive resurgence after social distancing interventions are otherwise lifted (Fig. 4).

Five types of factors tend to promote superspreading: (1) high rates or intensity of contact between people or with surfaces; (2) large aggregations of people; (3) poorly ventilated physical environments, especially indoors (Lu et al., 2020); (4) highly infectious individuals; (5) highly susceptible recipient population (Kumar et al., 2020; Althouse et al., 2020; Frieden and Lee, 2020). Many settings where SARS-CoV-2 superspreading has occurred—including nursing homes (Arons et al., 2020), exercise classes, bars and restaurants (Lu et al., 2020), funerals, churches (Shim et al., 2020), meat-packing plants (Dyal, 2020)—combine multiple risk factors. For example, choir practices combine high densities of people, a high-risk activity (singing) (Stadnytskyi et al., 2020), and potentially poorly ventilated indoor spaces; long-term care facilities combine mobile, high-contact caregivers with highly vulnerable residents, often in high-density indoor spaces. In practical terms, "chopping off the tail" as used here aggregates these individual risk factors into the single metric of transmission rate, and models the reduction in transmission rate over four hour periods which we describe as superspreading events—short periods of time characterized by very high rates of transmission.

Some superspreading events may be easier to eliminate than others. Clearly, healthcare and long-term care facilities serve critical functions despite their high-risk nature, and taking all possible steps for decontamination and personal protection in these facilities is critical to mitigate this risk (Frieden and Lee, 2020). On the other end of the spectrum, voluntary, large, indoor events that are mainly for entertainment and could be postponed – gyms, clubs, sporting events, concerts, large lectures – may be the most viable option to reduce superspreading and "chop off the tail" of the contact rate distribution (Benzell et al., 2020). While these common sense interventions are not novel suggestions (Althouse et al., 2020), and are already part of reopening plans in almost

all locations, our work allows a direct comparison of how much general social distancing is avoided by eliminating a fraction of these high-risk events (Fig. 3). Truncation strategies are even more desirable in light of their effectiveness at preventing explosive resurgence after controls are otherwise lifted (Fig. 4). Mapping actual event types onto the contact rate distribution to determine how particular superspreading reduction policies would affect control remains an important next step. Importantly, associating superspreading with events and locations, rather than specific people, can avoid the stigma sometimes associated with being identified as a superspreader (Kumar et al., 2020). While superspreading *individuals* (e.g., those with consistently high viremia, contact rates, or duration of shedding) likely contribute to the spread of COVID-19, our model cannot compare the relative importance of superspreading events and superspreading individuals nor discount the importance of superspreading individuals. We emphasize instead that removing just the highest transmission rates by all individuals can have a large positive impact on the population level transmission rate without the great difficulty and cost of identifying and sequestering specific individuals.

The impact of truncation interventions is two-fold. First, removing the upper tail of the individual transmission rate distribution reduces the population-level mean, often dramatically (Fig. 3A,B). If the mean transmission rate already placed $\mathcal{R}_E$ near 1 (for example, due to other interventions), then additional truncation could be enough to cross this critical threshold. However, most intervention strategies that bring $\mathcal{R}_E$ to 1 already include prohibiting large gatherings, especially indoors, so additional truncation may not be possible within the context of first-wave interventions. However, truncation also acts on the variance and skew of the transmission rate distribution, though these effects are smaller than the effect on the mean (Fig. 4D–F compared to A–C). Given that super-spreading events are particularly dangerous when cases are few (in the early or late phases of the epidemic) (Lloyd-Smith et al., 2005), sustained truncation interventions could be extremely important for preventing explosive stochastic re-emergence when low case numbers allow general social distancing to be lifted (Fig. 4). In this scenario, resurgence remains rare (Fig. 4B) but possible because individual variation in transmission rates is large; most of the time infectious people transmit to few others, but occasionally someone infects dozens (Fig. 4A), quickly overwhelming testing, contact tracing, and isolation efforts. Sustained truncation dramatically reduces the probability of explosive resurgence, and constrains incipient transmission chains to be smaller and more manageable.

One limitation on understanding the effect of heterogeneity in transmission in particular locations is the challenge of estimating epidemiological parameters from noisy and imperfect data: necessarily a balancing act between model simplicity and complexity. Here, we rely on metrics of heterogeneity previously estimated for SARS-CoV-1 and SARS-CoV-2 (Lloyd-Smith et al., 2005; Endo et al., 2020a; Grantz et al., 2020) instead of estimating them directly from data; we focus our parameter estimation on the mean of the transmission rate distribution. Heterogeneity in contact rates or infectiousness, and the resulting distributional variance and skew, may vary based on local patterns of movement, contact, behavior, and population demography. This heterogeneity can have important consequences: in some cases epidemics with low mean R_0 can actually infect a larger proportion of the population than epidemics with higher mean R_0 – as was the case for the 1918 influenza pandemic as compared to the 2014 Ebola outbreak – due to the heterogeneity in transmission rates, as described by higher moments of the secondary case distribution (Hébert-Dufresne et al., 2020).

Despite our focus solely on transmission rate heterogeneity, many sources of individual-level heterogeneity are likely to affect transmission, including variation in susceptibility, which lowers the herd immunity threshold and therefore produces a smaller epidemic than would be predicted in a homogeneously susceptible population (Britton et al., 2020; Gomes et al., 2020). Differences in the length of the viral shedding period, where individuals are infectious for a long period of time, could also create superspreading individuals and impact transmission, and possibly impact the efficacy of the truncation intervention we propose here. For example, because we assume that truncation is only partially effective, if that effectiveness is heterogeneous across individual types (i.e., worse at reducing spread by individuals with long-term infectiousness), then this extra layer of heterogeneity could cause us to over-estimate the effectiveness of our modeled interventions. Finally, it is also possible that we overstate the effectiveness of tail truncation because infectious individuals behave differently than how we model them. In particular, we truncate by re-sampling values from the individual-level transmission rate distribution for any values that fall above a given threshold. This re-sampling can be interpreted as individuals substituting the highest transmission rate activities (e.g., a concert) with activities that are represented by the remaining transmission rate distribution (e.g., staying at home, or meeting a couple of friends). If instead individuals substitute this unavailable activity with the highest transmission rate activity that is available, our estimates for the effectiveness of truncation will be overstated.

The true epidemiological parameters in any given location, and the extent of our uncertainty in these parameters, also remain unknown because of the computational challenges of parameter estimation given the limited information contained in noisy case, death, and mobility data. For example, depending on how a particular candidate parameter combination weights the noisiness of cases and deaths and estimates initial conditions, transmission rate estimates can vary substantially (Fig. 1). Fully characterizing uncertainty in model structure and parameter values proved difficult; some pairs of parameters were highly correlated (e.g., β_0 and β_{min}: see Figures S2–S6), and the estimated values for many parameters from the best fitting parameter sets were indistinguishable from the full range of parameter values sampled over to fit the model (Figure S10). Future work that directly estimates case ascertainment rates (e.g., through metrics of percentage of tests that are positive, age distributions of positive tests, epidemiological contact information on cases, and analysis of viral genome sequences (Miller et al., 2020), as well as more detailed mobility and contact network information (Ferretti et al., 2020) could help to improve the model fit to the full shape of the transmission rate distribution.

First-wave interventions that eliminated large social gatherings and indoor activities and mandated mask-wearing and physical distancing have likely already affected the heterogeneity in transmission rates, by eliminating many of the high-risk events likely to fall into the upper tail of the distribution. It is important to recognize that as social distancing interventions relax, sustaining such truncation interventions may be critical for keeping transmission down to levels manageable through testing, contact tracing, and isolation. This truncation strategy can potentially reduce the social and economic costs of non-pharmaceutical interventions on the general populace, and facilitate sustained adherence by allowing lower-risk activities to resume while insuring against a resurgence. Ultimately, an unmitigated epidemic, whether as a first, second, or third wave, would kill thousands to tens of thousands of people in each of the locations we studied, reinforcing the point that aiming for population herd immunity through naturally acquired infections is not a viable public health strategy. Instead, exit strategies that can sustain epidemic control after shelter-in-place orders end, including truncating the transmission rate distribution, will be necessary until an effective vaccine can be widely distributed.

CRediT authorship contribution statement

Morgan P. Kain: Conceptualization, Methodology, Software, Investigation, Writing - original draft, Writing - review & editing. **Marissa L. Childs:** Conceptualization, Methodology, Software, Investigation, Writing - original draft, Writing - review & editing. **Alexander D. Becker:** Software, Investigation, Writing - review & editing. **Erin A. Mordecai:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data and code availability

Data used in this study are available at: https://github.com/nytimes/covid-19-data. Code used to produce the results (and complete raw output from fits for all five locations) in this study are available at: https://github.com/morgankain/COVID_interventions.

Acknowledgments

We thank the members of the Mordecai Lab at Stanford University for feedback on our model. Funding provided by: the National Science Foundation, USA (DEB-1518681 and DEB-2011147); the National Institutes of Health, USA (National Institute of General Medical Sciences: R35GM133439); the Natural Capital Project, USA; the Helman Scholarship, USA; the King Center for Global Development, USA; and the Terman Award, USA. Morgan Kain was supported by the Natural Capital Project, USA. Marissa Childs was supported by the Illich-Sadowsky Fellowship through the Stanford Interdisciplinary Graduate Fellowship program at Stanford University, USA.

Appendix A

A.1. Derivation of the population-level β_t distribution

In heterogeneous populations, the expected number of secondary infections caused by a particular individual (or the individual reproductive number, ν) can be modeled as a negative binomial random variable with mean R_0 and overdispersion parameter k (Lloyd-Smith et al., 2005; Althouse et al., 2020; Endo et al., 2020b), i.e. $\nu \sim NB(k, \frac{k}{R_0+k})$. This is equivalent to modeling ν as a Poisson random variable whose mean is itself a random gamma variable with shape k and scale R_0/k,

$$\nu \sim \text{Poisson}(\theta)$$
$$\theta \sim \Gamma(k, \frac{R_0}{k})$$

Now let d be the duration of infection for an individual and τ be a time step. Using the fact that $k = \sum_M k/M$, we have

$$\begin{aligned}\theta &\sim \Gamma\left(k, \frac{R_0}{k}\right)\\ &\sim \Gamma\left(\sum_{d/\tau} \frac{k\tau}{d}, \frac{R_0}{k}\right)\\ &\sim \sum_{d/\tau} \Gamma\left(\frac{k\tau}{d}, \frac{R_0}{k}\right).\end{aligned}$$

Thus for a constant duration of infection d, we have the individual infection rate over a time step

$$\pi \sim \Gamma\left(\frac{k\tau}{d}, \frac{R_0}{k}\right).$$

When there are N infected individuals, the average infection rate β_t over a time step is

$$\begin{aligned}\beta_t &= \frac{1}{N}\sum_{i=1}^{N} \pi\\ &\sim \frac{1}{N}\sum_{i=1}^{N} \Gamma\left(\frac{k\tau}{d}, \frac{R_0}{k}\right)\\ &\sim \Gamma\left(\frac{Nk\tau}{d}, \frac{R_0}{Nk}\right),\end{aligned}$$

which will have mean $R_0\tau/d$ and variance $\frac{R_0^2\tau}{Nkd}$. Notably, this behaves well with scalings on R_0 as a function of interventions: Let θ be the amount of physical distancing occurring in the population on a scale of 0–1 where 0 is no physical distancing, and 1 is maximum physical distancing, and f be a function mapping θ to a scaling on R_0. Now $\beta_t \sim f(\theta)\Gamma\left(\frac{Nk\tau}{d}, \frac{R_0}{Nk}\right) = \Gamma\left(\frac{Nk\tau}{d}, \frac{f(\theta)R_0}{Nk}\right)$, which means that properties of the distribution are preserved with $f(\theta)R_0$. Specified as a gamma white noise process $\Gamma_{WN}(\sigma, \mu)$ which has mean μ and variance $\sigma^2\mu$, this is equivalently

$$\beta_t \sim \Gamma_{WN}\left(\sqrt{\frac{R_0}{Nk}}, \frac{R_0\tau}{d}\right).$$

There are two main differences between the above derivation and our model formulation (note that for the following we assume $f(\theta) = 1$):

1. above, the number of new infections in a time step should be

   $$\sum_{i=1}^{N} \text{Poisson}(\pi_i),$$

   where each π_i is i.i.d. as $\Gamma\left(\frac{k\tau}{d}, \frac{R_0}{k}\right)$. In our model the number of new infections in a time step is

   $$\mathcal{B}(S, 1 - \exp(-\beta_t(C_a I_a/N + C_p I_p/N + C_m I_m/N + C_s I_s/N))),$$

   where β_t is the average transmission rate over all individuals infectious during that time step and $\mathcal{B}$ is a binomial process. For large S and small $I_a + I_p + I_m + I_s$, this approximates a Poisson distribution for the number of secondary cases from each infected individual in each time step and the total secondary infections caused by an individual over their infectious period.
2. above, we assume a constant duration of infection. In our model periods are Erlang distributed given our division of stages (e.g. I_a) into n sub stages, each with the same period (Anderson and Watson, 1980; Lloyd, 2001). This marginally increases the variance in the per-infectious period distribution as we show in Figure S14.

A.2. Derivation of the relationship between R, gamma truncation, and the fraction of individuals sheltering in place

Let d be the average duration of infection, τ be a time step, θ be the proportion of the population sheltering in place, and $X_{p,\eta}$ be a random variable X with right truncation, where truncation occurs at the pth percentile, with probability η. Suppose that in a given time step we truncate the individual infection rate (π) over a time step at the pth percentile with probability η. Then the reproduction number is

$$\begin{aligned}R &= \mathbf{E}\left[\frac{d}{\tau}\beta_t\right]\\ R &= \mathbf{E}\left[\frac{d}{\tau N}\sum_{i=1}^{N} \pi_{p,\eta} \exp\left(\log(\beta_{min})\theta\right)\right]\\ R &= \frac{\beta_{min}^{\theta} d}{\tau N}\sum_{i=1}^{N} \mathbf{E}[\pi_{p,\eta}]\\ R &= \frac{\beta_{min}^{\theta} d}{\tau}\mathbf{E}[\pi_{p,\eta}]\\ R &= \frac{\beta_{min}^{\theta} d}{\tau}\left(\eta\mathbf{E}[\pi_{p,1}] + (1-\eta)\mathbf{E}[\pi_{1,1}]\right)\\ \beta_{min}^{-\theta} &= \frac{d}{R\tau}\left(\eta\mathbf{E}[\pi_{p,1}] + (1-\eta)\mathbf{E}[\pi_{1,1}]\right)\\ -\theta\log(\beta_{min}) &= log(d) - \log(R) - \log(\tau) + \log\left(\eta\mathbf{E}[\pi_{p,1}] + (1-\eta)\mathbf{E}[\pi_{1,1}]\right)\\ \theta &= \frac{log(d) - \log(R) - \log(\tau) + \log\left(\eta\mathbf{E}[\pi_{p,1}] + (1-\eta)\mathbf{E}[\pi_{1,1}]\right)}{-\log(\beta_{min})}\end{aligned}$$

For a truncated gamma distribution with shape a and scale b with upper truncation at u, the expected value is

$$\mathbf{E}[\Gamma(a,b,u)] = \frac{b\left[\Gamma(a+1,0) - \Gamma(a+1,u/b)\right]}{\Gamma(a,0) - \Gamma(a,u/b)},$$

where Γ is the upper incomplete gamma function. See Okasha and Alqanoo (2014) [Eq. (29)](Okasha and Alqanoo, 2014) for the full derivation.

Letting γ be the lower incomplete gamma function, it follows that

$$\begin{aligned}\mathbf{E}[\pi_{p,1}] &= \frac{R_0/k\left[\Gamma(k\tau/d+1,0) - \Gamma(k\tau/d+1,uk/R_0)\right]}{\Gamma(k\tau/d,0) - \Gamma(k\tau/d,uk/R_0)} \\ &= \frac{R_0/k\left[\gamma(k\tau/d+1,uk/R_0)\right]}{\gamma(k\tau/d,uk/R_0)}\end{aligned}$$

where u is the pth percentile of π (Okasha and Alqanoo, 2014). Then

$$\theta = \frac{log(d) - \log(R) - \log(\tau) + \log\left(\eta\frac{R_0/k\left[\gamma(k\tau/d+1,uk/R_0)\right]}{\gamma(k\tau/d,uk/R_0)} + (1-\eta)\frac{R_0\tau}{d}\right)}{-\log(\beta_{min})}.$$

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.epidem.2020.100430.

References

Adam, D., Wu, P., Wong, J., Lau, E., Tsang, T., Cauchemez, S., Leung, G., Cowling, B., 2020. Clustering and superspreading potential of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections in Hong Kong. Res. Square http://dx.doi.org/10.21203/rs.3.rs-29548/v1, Available from: https://doi.org/10.21203/rs.3.rs-29548/v1.

Althouse, B.M., Wenger, E.A., Miller, J.C., Scarpino, S.V., Allard, A., Hé-Dufresne, L., Hu, H., 2020. Stochasticity and heterogeneity in the transmission dynamics of SARS-CoV-2. Available from: https://covid.idmod.org/data/Stochasticity_heterogeneity_transmission_dynamics_SARS-CoV-2.pdf.

Anderson, D., Watson, R., 1980. On the spread of a disease with gamma distributed latent and infectious periods. Biometrika 67 (1), 191–198.

Arons, M.M., Hatfield, K.M., Reddy, S.C., Kimball, A., James, A., Jacobs, J.R., Taylor, J., Spicer, K., Bardossy, A.C., Oakley, L.P., et al., 2020. Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. New England J. Med. 382, 2081–2090.

Benzell, S.G., Collis, A., Nicolaides, C., 2020. Rationing social contact during the COVID-19 pandemic: Transmission risk and social benefits of US locations. Proc. Natl. Acad. Sci. Available from: https://www.pnas.org/content/early/2020/06/09/2008025117.

Britton, T., Ball, F., Trapman, P., 2020. A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. Science 369 (6505), 846–849.

Burhenne, S., Jacob, D., Henze, G.P., 2011. Sampling based on Sobol'sequences for Monte Carlo techniques applied to building simulations. In: Proc. Int. Conf. Build. Simulat, pp. 1816–1823.

Childs, M.L., Kain, M.P., Kirk, D., Harris, M., Couper, L., Nova, N., Delwel, I., Ritchie, J., Mordecai, E.A., 2020. The impact of long-term non-pharmaceutical interventions on COVID-19 epidemic dynamics and control. medRxiv, Available from: https://www.medrxiv.org/content/medrxiv/early/2020/05/06/2020.05.03.20089078.

County Population Totals 2010-2019, 2019. United states census bureau. Available at: https://www.census.gov/data/datasets/time-series/demo/popest/2010s-counties-total.html.

Davies, N.G., Klepac, P., Liu, Y., Prem, K., Jit, M., Eggo, R.M., working group, C.C.-., et al., 2020. Age-dependent effects in the transmission and control of COVID-19 epidemics. medRxiv. Available from: https://www.medrxiv.org/content/medrxiv/early/2020/03/27/2020.03.24.20043018.

Dyal, J.W., 2020. COVID-19 Among Workers in Meat and Poultry Processing Facilities–19 States, April 2020, Vol. 69. CDC: (MMWR) Morbidity and Mortality Weekly Report, Available from: https://www.cdc.gov/mmwr/volumes/69/wr/mm6918e3.htm.

Endo, A., Abbott, S., Kucharski, A.J., Funk, S., et al., 2020a. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. Wellcome Open Res. 5 (67), 67.

Endo, A., for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, C., Abbott, S., Kucharski, A.J., Funk, S., 2020b. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. Wellcome Open Res. 5.

Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Döner, L., Parker, M., Bonsall, D., Fraser, C., 2020. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science (ISSN: 1095-9203) 368 (6491), http://dx.doi.org/10.1126/science.abb6936.

Frieden, T.R., Lee, C.T., 2020. Identifying and interrupting superspreading events—implications for control of severe acute respiratory syndrome coronavirus 2. CDC Stacks Available from: https://stacks.cdc.gov/view/cdc/88905.

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.

Gaythorpe, K., Imai, N., Cuomo-Dannenburg, G., Baguelin, M., Bhatia, S., Boonyasiri, A., Cori, A., Cucunuba Perez, Z., Dighe, A., Dorigatti, I., Fitzjohn, R., Fu, H., Green, W., Griffin, J., Hamlet, A., Hinsley, W., Hong, N., Kwun, M., Laydon, D., Nedjati Gilani, G., Okell, L., Riley, S., Thompson, H., Van Elsland, S., Verity, R., Volz, E., Walker, P., Wang, H., Wang, Y., Walters, C., Whittaker, C., Winskill, P., Xi, X., Donnelly, C., Ghani, A., Ferguson, N., 2020. Report 8: Symptom progression of COVID-19. Imperial College London, http://spiral.imperial.ac.uk/handle/10044/1/77344.

Gomes, M.G.M., Aguas, R., Corder, R.M., King, J.G., Langwig, K.E., Souto-Maior, C., Carneiro, J., Ferreira, M.U., Penha-Goncalves, C., 2020. Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold. medRxiv, Available from: https://www.medrxiv.org/content/10.1101/2020.04.27.20081893v3.

Grantz, K., Metcalf, C.J.E., Lessler, J., 2020. Dispersion vs. control. Available at: https://hopkinsidd.github.io/nCoV-Sandbox/DispersionExploration.html.

Gudbjartsson, D.F., Helgason, A., Jonsson, H., Magnusson, O.T., Melsted, P., Norddahl, G.L., Saemundsdottir, J., Sigurdsson, A., Sulem, P., Agustsdottir, A.B., et al., 2020. Spread of SARS-CoV-2 in the icelandic population. New Engl. J. Med. 382, 2302–2315.

Guo, Z.-D., Wang, Z.-Y., Zhang, S.-F., Li, X., Li, L., Li, C., Cui, Y., Fu, R.-B., Dong, Y.-Z., Chi, X.-Y., et al., 2020. Aerosol and surface distribution of severe acute respiratory syndrome coronavirus 2 in hospital wards, wuhan, China, 2020. Emerg. Infect Dis. 26 (7), 10–3201.

He, D., Ionides, E.L., King, A.A., 2010. Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. J. R. Soc. Interface 7 (43), 271–283.

He, X., Lau, E.H., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y.C., Wong, J.Y., Guan, Y., Tan, X., et al., 2020a. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat. Med, 26 (5), 672–675.

He, D., Zhao, S., Lin, Q., Zhuang, Z., Cao, P., Wang, M.H., Yang, L., 2020b. The relative transmissibility of asymptomatic cases among close contacts. Int. J. Infec. Dis. 63 (5), 706–711.

Hébert-Dufresne, L., Althouse, B.M., Scarpino, S.V., Allard, A., 2020. Beyond R0: Heterogeneity in secondary infections and probabilistic epidemic forecasting. Available from: https://www.medrxiv.org/content/10.1101/2020.02.10.20021725v2.

Hu, Z., Song, C., Xu, C., Jin, G., Chen, Y., Xu, X., Ma, H., Chen, W., Lin, Y., Zheng, Y., et al., 2020. Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in nanjing, China. Sci. China Life Sci. 63 (5), 706–711.

King, A.A., Nguyen, D., Ionides, E.L., 2016. Statistical inference for partially observed Markov processes via the r package pomp. J. Stat. Softw. 69 (12), 1–43. http://dx.doi.org/10.18637/jss.v069.i12.

Kumar, S., Jha, S., Rai, S.K., 2020. Significance of super spreader events in COVID-19. Indian J. Public Health 64 (6), 139.

Lauer, S.A., Grantz, K.H., Bi, Q., Jones, F.K., Zheng, Q., Meredith, H.R., Azman, A.S., Reich, N.G., Lessler, J., 2020. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann. Intern. Med. 172 (9), 577–582.

Lavezzo, E., Franchin, E., Ciavarella, C., Cuomo-Dannenburg, G., Barzon, L., Vecchio, C.D., Rossi, L., Manganelli, R., Loregian, A., Navarin, N., Abate, D., Sciro, M., Merigliano, S., Decanale, E., Vanuzzo, M.C., Saluzzo, F., Onelia, F., Pacenti, M., Parisi, S., Carretta, G., Donato, D., Flor, L., Cocchio, S., Masi, G., Sperduti, A., Cattarino, L., Salvador, R., Gaythorpe, K.A.M., Team, I.C.L.C.-.R., Brazzale, A.R., Toppo, S., Trevisan, M., Baldo, V., Donnelly, C.A., Ferguson, N.M., Dorigatti, I., Crisanti, A., 2020. Suppression of COVID-19 outbreak in the municipality of vo, Italy. medRxiv, Available at: https://www.medrxiv.org/content/10.1101/2020.04.17.20053157v1.

Lee, S., Meyler, P., Mozel, M., Tauh, T., Merchant, R., 2020. Asymptomatic carriage and transmission of SARS-CoV-2: What do we know?. Canad. J. Anaesth. 1–7.

Li, L.-q., Huang, T., Wang, Y.-q., Wang, Z.-p., Liang, Y., Huang, T.-b., Zhang, H.-y., Sun, W., Wang, Y., 2020a. COVID-19 patients' clinical characteristics, discharge rate, and fatality rate of meta-analysis. J. Med. Virol. 92 (6), 577–583.

Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., Shaman, J., 2020b. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). Science 368 (6490), 489–493.

Liu, Y., Eggo, R.M., Kucharski, A.J., 2020. Secondary attack rate and superspreading events for SARS-CoV-2. Lancet 395 (10227), e47.

Lloyd, A.L., 2001. Realistic distributions of infectious periods in epidemic models: changing patterns of persistence and dynamics. Theor. Popul. Biol. 60 (1), 59–71.

Lloyd-Smith, J.O., Schreiber, S.J., Getz, W.M., 2006. Moving beyond averages: Individual-level variation in disease transmission. In: Mathematical Studies on Human Disease Dynamics: Emerging Paradigms and Challenges: AMS-IMS-SIAM Joint Summer Research Conference on Modeling the Dynamics of Human Diseases: Emerging Paradigms and Challenges, July 17-21, 2005, Vol. 410. American Mathematical Soc., Snowbird, Utah, p. 235.

Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E., Getz, W.M., 2005. Superspreading and the effect of individual variation on disease emergence. Nature 438 (7066), 355–359.

Lu, J., Gu, J., Li, K., Xu, C., Su, W., Lai, Z., Zhou, D., Yu, C., Xu, B., Yang, Z., 2020. COVID-19 outbreak associated with air conditioning in restaurant, guangzhou, China, 2020.. Emerg. Infec. Dis. 26 (7).

McAloon, C.G., Collins, A., Hunt, K., Barber, A., Byrne, A., Butler, F., Casey, M., Griffin, J.M., Lane, E., McEvoy, D., et al., 2020. The incubation period of COVID-19: A rapid systematic review and meta-analysis of observational research. medRxiv, Available from: https://www.medrxiv.org/content/medrxiv/early/2020/04/28/2020.04.24.20073957.

Miller, D., Martin, M.A., Harel, N., Kustin, T., Tirosh, O., Meir, M., Sorek, N., Gefen-Halevi, S., Amit, S., Vorontsov, O., et al., 2020. Full genome viral sequences inform patterns of SARS-CoV-2 spread into and within Israel. medRxiv, Available from: https://www.medrxiv.org/content/medrxiv/early/2020/05/22/2020.05.21.20104521.

Okasha, M.K., Alqanoo, I., 2014. Inference on the doubly truncated gamma distribution for lifetime data. Int. J. Math. Stat. Invent. 2, 1–17.

Perkins, A., Cavany, S.M., Moore, S.M., Oidtman, R.J., Lerch, A., Poterek, M., 2020. Estimating unobserved SARS-CoV-2 infections in the united states. medRxiv, Available from: https://www.medrxiv.org/content/10.1101/2020.03.15.20036582v2.

R Core Team, 2020. R: A language and environment for statistical computing. https://www.R-project.org/.

Rees, E.M., Nightingale, E.S., Jafari, Y., Waterlow, N.R., Clifford, S., Pearson, C.A., Jombart, T., Procter, S.R., Knight, G.M., Group, C.W., et al., 2020. COVID-19 length of hospital stay: a systematic review and data synthesis. medRxiv, Available from: https://www.medrxiv.org/content/10.1101/2020.04.30.20084780v3.

Sanche, S., Lin, Y.T., Xu, C., Romero-Severson, E., Hengartner, N., Ke, R., 2020. High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. Emerg. Infect. Diseases 26 (7).

Shelter in Place Index, 2020. The impact of coronavirus on human movement. Available at: https://www.safegraph.com/dashboard/covid19-shelter-in-place?s=US&d=06-04-2020&t=counties&m=index.

Shim, E., Tariq, A., Choi, W., Lee, Y., Chowell, G., 2020. Transmission potential and severity of COVID-19 in South Korea. Int. J. Infec. Dis. 93, 339–344.

Stadnytskyi, V., Bax, C.E., Bax, A., Anfinrud, P., 2020. The airborne lifetime of small speech droplets and their potential importance in SARS-CoV-2 transmission. Proc. Natl. Acad. Sci. 117 (22), 11875–11877.

Tindale, L., Coombe, M., Stockdale, J., Garlock, E., Lau, W.Y.V., 2020. Transmission interval estimates suggest pre-symptomatic spread of COVID-19. medRxiv, Available from: https://www.medrxiv.org/content/10.1101/2020.03.03.20029983v1.

To, K.K.-W., Tsang, O.T.-Y., Leung, W.-S., Tam, A.R., Wu, T.-C., Lung, D.C., Yip, C.C.-Y., Cai, J.-P., Chan, J.M.-C., Chik, T.S.-H., 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 Infec. Dis. 20, 565–574.

Verity, R., Okell, L.C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., Cuomo-Dannenburg, G., Thompson, H., Walker, P.G.T., Fu, H., Dighe, A., Griffin, J.T., Baguelin, M., Bhatia, S., Boonyasiri, A., Cori, A., Cucunub, Z., FitzJohn, R., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Laydon, D., Nedjati-Gilani, G., Riley, S., Elsland, S.v., Volz, E., Wang, H., Wang, Y., Xi, X., Donnelly, C.A., Ghani, A.C., Ferguson, N.M., 2020. Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet Infec. Dis. 20, 569–677. http://dx.doi.org/10.1016/S1473-3099(20)30243-7.

Wei, W.E., Li, Z., Chiew, C.J., Yong, S.E., Toh, M.P., Lee, V.J., 2020. Presymptomatic Transmission of SARS-CoV-2—Singapore, January 23–March 16, 2020. CDC: (MMWR) Morbidity and Mortality Weekly Report, 69, (14), Centers for Disease Control and Prevention, p. 411.

Wölfel, R., Corman, V.M., Guggemos, W., Seilmaier, M., Zange, S., Müller, M.A., Niemeyer, D., Jones, T.C., Vollmar, P., Rothe, C., et al., 2020. Virological assessment of hospitalized patients with COVID-2019. Nature 581 (7809), 465–469. http://dx.doi.org/10.1038/s41586-020-2196-x.

Woolhouse, M.E., Dye, C., Etard, J.-F., Smith, T., Charlwood, J., Garnett, G., Hagan, P., Hii, J., Ndhlovu, P., Quinnell, R., et al., 1997. Heterogeneities in the transmission of infectious agents: implications for the design of control programs. Proc. Natl. Acad. Sci. 94 (1), 338–342.

World Health Organization, 2020. Modes of transmission of virus causing COVID-19: implications for IPC precaution recommendations: scientific brief, 27 March 2020. World Health Organization, Available from: https://apps.who.int/iris/bitstream/handle/10665/331601/WHO-2019-nCoV-Sci_Brief-Transmission_modes-2020.1-eng.pdf.

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. Infec. Dis. 94, 133–138.

Zhang, W., 2020. Estimating the presymptomatic transmission of COVID19 using incubation period and serial interval data. medRxiv, Available from: https://www.medrxiv.org/content/medrxiv/early/2020/04/06/2020.04.02.20051318.

Zhang, W., Cheng, W., Luo, L., Ma, Y., Xu, C., Qin, P., Zhang, Z., 2020a. Secondary transmission of coronavirus disease from presymptomatic persons, China. Emerg. Infect. Diseases 26 (8).

Zhang, Y., Li, Y., Wang, L., Li, M., Zhou, X., 2020b. Evaluating transmission heterogeneity and super-spreading event of COVID-19 in a metropolis of China. Int. J. Environ. Res. Public Health 17 (10), 3705.