A model for COVID-19 transmission in Connecticut

Olga Morozova[1], Zehang Richard Li[2], and Forrest W. Crawford[3,4,5,6]

1. Program in Public Health and Department of Family, Population and Preventive Medicine, Stony Brook University (SUNY)
2. Department of Statisitcs, University of California, Santa Cruz
3. Department of Biostatistics, Yale School of Public Health
4. Department of Statistics & Data Science, Yale University
5. Department of Ecology & Evolutionary Biology, Yale University
6. Yale School of Management

March 8, 2021

Abstract

To support public health policymakers in Connecticut, we developed a county-structured compartmental SEIR-type model of SARS-CoV-2 transmission and COVID-19 disease progression. We calibrated this model to the local dynamics of deaths and hospitalizations and used a novel measure of close interpersonal contact frequency to approximate changes in infection transmission over time. In addition, we incorporated information on multiple time-varying parameters including the case fatality ratio, severity risk of incident cases, and length of hospital stay. In this paper, we describe the design, implementation, and calibration of a transmission model developed to meet the changing requirements of public health policymakers and officials in Connecticut. We describe methodology for producing short- and long-term projections of the epidemic evolution under uncertain future scenarios, as well as analytical tools for estimating epidemic features that are difficult to measure directly, such as cumulative incidence and effects of non-pharmaceutical interventions. The approach takes advantage of our unique access to Connecticut public health surveillance and hospital data to deliver COVID-19 projections tailored to the local context and responsive to the needs of local decision-makers.

Keywords: SEIR epidemic model, SARS-CoV-2, social distancing

1 Introduction

Epidemiologic models of disease transmission play an important role in supporting public health decision-making. Trajectories from these models can provide insights into historical trends in epidemic dynamics or future outcomes under hypothetical intervention scenarios. Transmission models are especially useful in situations of high uncertainty, offering a structured way to assess the potential effects of interventions given plausible assumptions about disease transmission. Models cannot predict the future with certainty, but they can be helpful for scenario analysis by bounding the range of plausible future trajectories [1].

In the absence of effective pharmaceutical interventions, many countries, including the US, implemented social distancing measures and stay-at-home orders to slow transmission of SARS-CoV-2. As transmission subsided, many states, including Connecticut, began considering phased lifting of social distancing restrictions. Due to unprecedented nature of these events, public health policymakers were faced with many questions: 1) How soon

can interventions like school closures and stay-at-home orders be lifted? 2) How should public health interventions be implemented in the future to minimize the risk of a resurgence? 3) What will be the effect of phased reopening plans? Surveillance data on testing, case counts, hospitalizations, and deaths was useful to assess the past dynamics, however policymakers needed additional analytic tools to help them evaluate the current state of an epidemic, as well as potential future risks to design strategies to address possible resurgence.

Multiple transmission models have been developed recently to help answer these and other policy questions [2–24]. When models are developed with a primary goal to support decision-making, they must balance parsimony and realism. Projection models should be simple enough to fit to data, and should provide clear outputs in a timely manner with assumptions that can be understood by policymakers. At the same time, for such models to be useful, they need to be flexible enough to accommodate realistic epidemiological features and likely policy scenarios identified by stakeholders.

Many of the nation-wide COVID-19 forecasting models that have been developed in the United States either do not use local data, or make simplifying assumptions not appropriate for localities [2]. These models may capture national-level dynamics, but are less useful for supporting decision-making in individual states or counties. Several nation-wide models have been developed with a goal to provide projections at the state level [25–35]. These models employ varying methods, make different structural and parameter value assumptions, and project potential effects of different future policy interventions. They also vary in terms of model outputs and the ways of handling uncertainty. Most of these models use the same set of assumptions and estimates of key model parameters across all US states, and may not be able to capture important local variation.

In this paper, we introduce a county-structured transmission model of SARS-CoV-2 transmission and COVID-19 disease progression in Connecticut. The model was developed with a goal to support intervention planning and decision-making in Connecticut, but could be adapted to other states or regions. This paper provides an in-depth technical description of the model and data calibration approach. Additional COVID-19 reports for Connecticut in this series are available from https://crawford-lab.github.io/covid19_ct/. An earlier version of this model was used in late May of 2020 to produce projections for Connecticut through the end of summer of 2020 [36].

2 Methods

We developed a deterministic compartmental model of SARS-CoV-2 transmission and COVID-19 disease progression. The model is based on the SEIR (Susceptible, Exposed, Infectious, Removed) framework [37], which we extend to accommodate geographical variation in Connecticut and distinct features of COVID-19 disease. We calibrate the model to observed dynamics of deaths and hospitalizations in Connecticut. Similar models have been published recently, offering intervention effect estimates and projections in various locations [2, 3, 6, 7, 9, 13, 14, 24, 38, 39].

Figure 1 shows a schematic representation of the model structure. We categorize infections as asymptomatic, mild symptomatic, and severe. Only severe infections may lead to death. Severe infections are defined as those requiring hospitalization. If hospitalization capacity is overwhelmed, some severe cases in the community are denied hospitalization, and experience a higher probability of death compared to hospitalized cases. Mild symptomatic cases are assumed to self-isolate shortly after they develop symptoms and remain isolated until they recover. During the infectiousness period of symptomatic cases, we assume a period of presymptomatic viral shedding, i.e. the latency period is shorter than incubation period [40–44]. In line with other similar models, we assume that individuals with asymptomatic infection exert a lower force of infection, but remain infectious for a longer period of time, since they are less likely to self-isolate in the absence of widespread comprehensive screening programs [4, 9, 39]. The average time that severe cases spend in the infectious state is approximated by the time between onset of infectiousness and hospitalization (or attempted hospitalization in case of hospital overflow). This model is intended to represent community spread of SARS-CoV-2, and excludes transmission occurring in congregate settings like skilled nursing

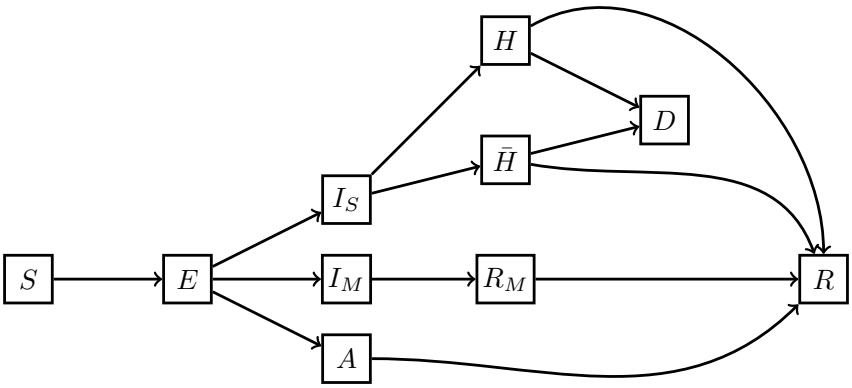

Figure 1: Schematic illustration of the SARS-CoV-2 transmission model and COVID-19 disease progression structure. Individuals begin in the susceptible (S) compartment. Exposed individuals (E) may develop either asymptomatic (A), mild (I_M), or severe (I_S) infection. Asymptomatic and mild infections resolve without hospitalization and do not lead to death. Mild symptomatic cases self-isolate (R_M) shortly after development of symptoms, and transition to recovery (R) when infectiousness ceases. All severe cases require hospitalization (H) unless hospitalization capacity is exhausted, in which case they transition to $\bar{H}$ representing hospital overflow, then to recovery (R) or death (D). The models captures infection transmission in non-congregate settings, and excludes cases and deaths occurring in settings like nursing homes and prisons. It assumes a closed population without births and does not capture non-COVID-19 deaths.

and assisted living facilities, or prisons. Similar to [14], we excluded congregate settings, since transmission in small closed communities violates key mass action-type modeling assumptions. The force of infection from hospitalized patients to unhospitalized susceptible individuals is assumed to be negligible. We further assume that recovered individuals remain immune to reinfection for the duration of the study period. The model is implemented at the level of individual counties in Connecticut assuming that most contacts are happening within a given county. A small proportion of contacts is allowed to happen between adjacent counties. The analysis was performed using the R statistical computing environment [45]. We used package deSolve to perform numerical integration of the system of ordinary differential equations (ODE) [46].

2.1 Compartmental model and parameters

The model divides the population of a given county into 10 compartments: susceptible (S), exposed, latent infections (E), infectious and asymptomatic (A), infectious and mild symptomatic (I_M), infectious and severe (I_S), isolated mild infections removed from the pool of infectious individuals (R_M), hospitalized (H), severe in need of hospitalization, but denied it due to hospital capacity overflow ($\bar{H}$), recovered (R), and died (D). Let N_i be the population size of county i and J_i be the set of counties adjacent to county i. Let $C^{(i)}$ represent hospitalization capacity in county i,

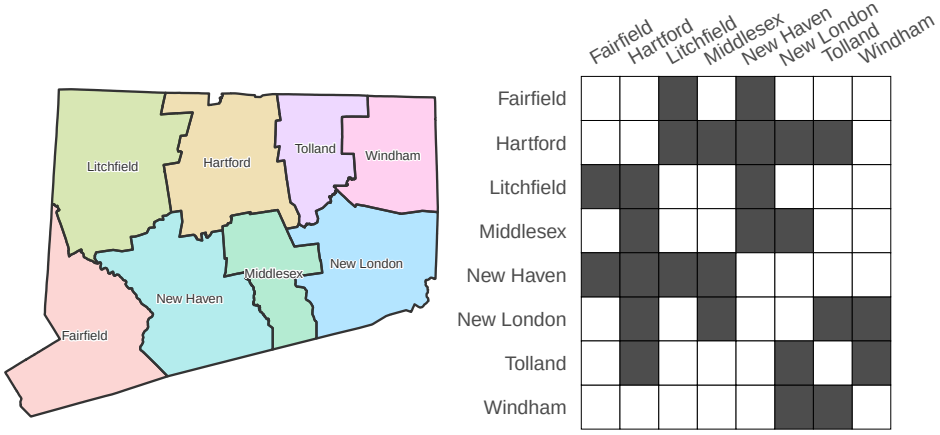

Figure 2: County map of Connecticut and county adjacency matrix. The dark gray cells correspond to counties that are adjacent. Contacts between adjacent counties are included in the model in addition to contacts within counties.

which may vary over time. Transmission dynamics for county i are given by the following ODE system:

$$
\begin{aligned}
\frac{\mathrm{d}S^{(i)}}{\mathrm{d}t} &= -\beta S^{(i)}\left[(1-k_n)\frac{I_M^{(i)}+I_S^{(i)}+k_A A^{(i)}}{N_i}+\frac{k_n}{|J_i|}\sum_{j\in J_i}\frac{I_M^{(j)}+I_S^{(j)}+k_A A^{(j)}}{N_j}\right]\\
\frac{\mathrm{d}E^{(i)}}{\mathrm{d}t} &= \beta S^{(i)}\left[(1-k_n)\frac{I_M^{(i)}+I_S^{(i)}+k_A A^{(i)}}{N_i}+\frac{k_n}{|J_i|}\sum_{j\in J_i}\frac{I_M^{(j)}+I_S^{(j)}+k_A A^{(j)}}{N_j}\right]-\delta E^{(i)}\\
\frac{\mathrm{d}A^{(i)}}{\mathrm{d}t} &= q_A\delta E^{(i)}-\alpha_A A^{(i)}\\
\frac{\mathrm{d}I_M^{(i)}}{\mathrm{d}t} &= q_{I_M}\delta E^{(i)}-\alpha_{I_M} I_M^{(i)}\\
\frac{\mathrm{d}R_M^{(i)}}{\mathrm{d}t} &= \alpha_{I_M} I_M^{(i)}-\gamma_{R_M} R_M^{(i)}\\
\frac{\mathrm{d}I_S^{(i)}}{\mathrm{d}t} &= q_{I_S}\delta E^{(i)}-\alpha_{I_S} I_S^{(i)}\\
\frac{\mathrm{d}H^{(i)}}{\mathrm{d}t} &= (1-\eta^{(i)})\alpha_{I_S} I_S^{(i)}-\gamma_H H^{(i)}\\
\frac{\mathrm{d}\bar{H}^{(i)}}{\mathrm{d}t} &= \eta^{(i)}\alpha_{I_S} I_S^{(i)}-\gamma_{\bar{H}}\bar{H}^{(i)}\\
\frac{\mathrm{d}D^{(i)}}{\mathrm{d}t} &= \gamma_H m_H H^{(i)}+\gamma_{\bar{H}} m_{\bar{H}}\bar{H}^{(i)}\\
\frac{\mathrm{d}R^{(i)}}{\mathrm{d}t} &= \alpha_A A^{(i)}+\gamma_{R_M} R_M^{(i)}+\gamma_H(1-m_H)H^{(i)}+\gamma_{\bar{H}}(1-m_{\bar{H}})\bar{H}^{(i)},
\end{aligned}
\tag{1}
$$

where $q_{I_M} = 1 - q_A - q_{I_S}$. The function $\eta^{(i)} = \left[1+\exp(0.5(C^{(i)}-H^{(i)}))\right]^{-1}$ is a "soft" hospitalization capacity overflow function.

Table 1 lists model parameters and their definitions. Recognizing that many of the model parameters are unlikely to be constant over time, we allow the most critical parameters to vary over time. In particular, transmission parameter β is expected to change substantially in response to social distancing measures. The description and assumptions regarding time-varying model parameters follows below.

Table 1: Transmission model parameters.

Notation	Definition
β	Transmission parameter per susceptible - infectious pair
δ	1 / Latency period (days^{-1})
q_A, q_{I_M}, q_{I_S}	Proportions of infections that are asymptomatic, mild symptomatic, and severe, $q_A + q_{I_M} + q_{I_S} = 1$
α_A	1 / Duration of infectiousness among asymptomatic cases (days^{-1})
k_A	Relative infectiousness of asymptomatic cases compared to symptomatic
α_{I_M}	1 / Duration of infectiousness among mild symptomatic cases, time until isolation (days^{-1})
γ_{R_M}	1 / Duration of isolation among mild symptomatic cases, remaining time to recovery (days^{-1})
α_{I_S}	1 / Duration of infectiousness among severe cases, time to hospitalization (days^{-1})
γ_H	1 / Length of hospital stay (time until recovery or death) (days^{-1})
$\gamma_{\bar{H}}$	1 / Remaining time until recovery or death among hospital overflow patients^{-1} (days^{-1})
m_H	Case fatality ratio among hospitalized cases
$m_{\bar{H}}$	Case fatality ratio among hospital overflow patients
k_n	Proportion of all contacts that happen with individuals from adjacent counties (as opposed to within the county)
C	Hospitalization capacity, may be constant or vary over time representing capacity increase intervention
E_0	Number of exposed individuals statewide at the time of epidemic onset

Figure 2 shows the county map of Connecticut along with the county adjacency matrix. The geographic boundary files were obtained from the Connecticut Department of Environmental Protection [47]. We assume that a fraction $(1 - k_n)$ of all contacts happen within a given county, and the remaining k_n contacts happen between individuals residing in adjacent counties.

2.2 Time-varying model parameters

2.2.1 Transmission parameter β

Social distancing measures and practices reduce the value of transmission parameter β. We use data from [48] to approximate changes in close interpersonal contact among Connecticut residents between February 1, 2020 and the end of the modeling period, and assume the following functional form of transmission parameter β:

$$\beta(t) = \beta_0 M_{\text{contact}}(t) \exp[B(t)],$$

where $M_{\text{contact}}(t)$ is a normalized measure of close interpersonal contact at time t relative to the pre-epidemic level, and $\exp[B(t)]$ is a function that approximates residual changes in transmission parameter β that are not explained by changes in close contact and other time-varying parameters. $B(t)$ is a smooth function obtained by applying spline smoothing on a piecewise linear function $B^*(t)$, where $B^*(t)$ is modeled with $B^*(w) = \epsilon_{[(w-t_0)/14]}$ defined on bi-weekly knots $w = \{t_0, t_0 + 14, t_0 + 28, \ldots\}$ over the observation period and linearly imputed between the knots. We model the vector of random effects $\boldsymbol{\epsilon}$ using a random walk of order one:

$$\epsilon_0 = 0, \quad \epsilon_i | \epsilon_{i-1} \sim \mathcal{N}(\epsilon_{i-1}, \sigma_\epsilon^2).$$

For the hyperparameter σ_ϵ^2, we use Inverse-Gamma(a_ϵ, b_ϵ) prior with a shape parameter $a_\epsilon = 2.5$ and a rate parameter $b_\epsilon = 0.1$.

The function $B(t)$ can also be used to impose arbitrary changes in $\beta(t)$ in the future to test different scenarios and potential intervention effects.

2.2.2 Rates of isolation and recovery: α_{I_M} and α_A

Wide-spread testing and contact tracing efforts can potentially reduce duration of infectiousness. Effectiveness of testing and contact tracing strategies depends on many factors. While there is no information about reduction in duration of infectiousness in response to specific testing efforts implemented in Connecticut, our model accommodates the possibility of such reduction as a function of daily testing volume.

$$\alpha(t) = \alpha_0(1 + M_{\text{testing}}(t)\tau),$$

where τ is the size of testing effect per unit increase in testing volume measure $M_{\text{testing}}(t)$ modeled as:

$$M_{\text{testing}}(t) = \begin{cases} \log\left(v_{\text{testing}}(t)\right) - \log\left(v_{\text{testing}}(t^*)\right), & t > t^* \text{ and } v_{\text{testing}}(t) \geq v_{\text{testing}}(t^*) \\ 0, & \text{otherwise.} \end{cases}$$

$v_{\text{testing}}(t)$ is a spline-smoothed measure of testing volume at time t. Testing efforts early in the epidemic were primarily used to confirm severe and highly symptomatic infections, and were unlikely to have any appreciable impact on overall duration of infectiousness. Early response daily testing volume is denoted by $v_{\text{testing}}(t^*)$.

This approach is used to model time-varying rates $\alpha_{I_M}(t)$ and $\alpha_A(t)$ with $\tau_{I_M} = \tau$ and $\tau_A = 0.5\tau$. Rate α_{I_S} is assumed to remain constant over time.

2.2.3 Severe fraction q_{I_S}

Probability of severe infection increases with age. Compliance with social distancing recommendations and other behavioral changes aimed to reduce the chances of infection are more likely among older people. Indeed, age distribution of confirmed cases in the U.S. has shifted toward younger people in the summer compared to spring [49]. At the same time, as community transmission increases, it becomes more difficult to protect the most vulnerable people from infection, even if this group continues to comply with social distancing measures. We model the proportion of infections that are severe as:

$$q_{I_S}(t) = q_{I_S,0} M_{\text{severity}}(t),$$

where measure of severity $M_{\text{severity}}(t)$ is a normalized spline-smoothed proportion of cases 60+ years old among all cases detected at time t relative to a baseline level. Since testing availability affects this proportion, we assume that $M_{\text{severity}}(t) = 1$ for all $t < t^*$, where t^* denotes the time when testing became widely available.

2.2.4 Rate of hospital discharge γ_H

We compute the rate of hospital discharge $\gamma_H(t)$ (including deaths and alive discharges) as a reciprocal of the average length of hospital stay at time t, which has been inversely correlated with the incidence of COVID-19. The average length of hospital stay at time t is approximated using a spline-smoothed monthly averages of this quantity.

2.2.5 Hospital case fatality ratio m_H

Due to potentially time-varying criteria for hospitalization, risk profile of incident cases, and changing clinical management practices, hospital case fatality ratio (hCFR) has been changing over time. We model hCFR as:

$$m_H(t) = m_{H,0} M_{\text{hCFR}}(t),$$

where $M_{\text{hCFR}}(t)$ is a normalized spline-smoothed hCFR at time t relative to the baseline hCFR = $m_{H,0}$. hCFR is calculated as a ratio of hospital deaths at time t to hospital admissions at time $(t - \text{hlos}(t))$, where $\text{hlos}(t)$ is an average hospital length of stay at time t.

2.3 Calibration data and estimation of hospitalizations coming from congregate and non-congregate settings

We calibrate the model to the observed dynamics of confirmed COVID-19 hospitalizations census, cumulative COVID-19 hospitalizations, and cumulative number of deaths among hospitalized cases. These data were obtained from the Connecticut Hospital Association (CHA) [50]. Data used to approximate time-varying model parameters were obtained from the following sources: mobile devices data used to estimate the frequency of close interpersonal contact were obtained from X-Mode (estimation details are given in [48]); daily testing volume and age distribution of confirmed cases were obtained from Connecticut Department of Public Health (CT DPH) daily reports [51]; monthly average hospital length of stay among COVID-19 patients was obtained from the CHA. We calculated non-institutionalized county-level population and age structure in Connecticut based on the American Community Survey [52]. Daily total available hospital beds (including occupied) in each county were obtained from the CHA/CHIMEData [50, 53] and used as hospitalization capacity values on a given date.

Available hospitalizations data does not disaggregate by the patient's place of residence at the time of diagnosis or hospitalization. According to CT DPH, as of October 30, 2020, about 73% of all deaths have occurred among residents of congregate settings, primarily nursing homes. Given that infection spread in congregate settings does not follow the mass action assumptions underlying our transmission model, we estimate the time series of the number of hospitalizations (census and cumulative) and hospital deaths coming from non-congregate settings and use these estimated counts in the model calibration.

We received the data on daily COVID-19 death counts in hospitals disaggregated by the type of residence at the time of diagnosis or hospitalization (congregate vs. non-congregate) from the CT DPH. Based on these data, the time-varying proportion of hospitalizations census and cumulative hospitalizations coming from congregate settings was estimated as follows:

1. estimate the number of cumulative hospitalizations coming from congregate settings as of the most recent death data date as the total number of hospital deaths among residents of congregate settings divided by the hospital CFR among this population, adjusting for an average hospital length of stay. The estimate of hospital CFR among residents of congregate settings in Connecticut is 0.38 and was obtained from the CT DPH based on a survey of a sample of 50 nursing homes, representing approximately half of all nursing homes in Connecticut;

2. estimate the cumulative proportion of hospitalizations coming from congregate settings as an estimate of the number of cumulative hospitalizations coming from congregate settings divided by the total cumulative number of hospitalizations as of the same date;

3. to estimate the instantaneous value of this proportion at each time point, approximate its temporal dynamics using the lagged temporal dynamics of the proportion of all hospital deaths coming from congregate settings, which can be estimated directly from the available death counts data.

This time-varying proportion is then applied to the CHA-reported daily hospital census and daily hospital admissions to estimate the number of current hospitalizations, daily hospital admissions, and cumulative hospitalizations coming from congregate settings, and respective time series coming from non-congregate settings follow directly.

2.4 Model calibration and Bayesian posterior inference

We calibrate the posterior distribution of model parameters to the estimated hospitalizations and hospital deaths coming from non-congregate settings using a Bayesian approach with a Gaussian likelihood. Model-based estimates of observed quantities are adjusted for reporting lags. The distributions of hospitalizations census, cumulative hospitalizations, and hospital deaths are given by:

$$h(t) \sim \mathcal{N}\left(H(t, L_H, \boldsymbol{\theta}), \sigma_h^2\right), \tag{2}$$

$$u(t) \sim \mathcal{N}\left(U(t, L_H, \boldsymbol{\theta}), \sigma_u^2\right), \tag{3}$$

$$d(t) \sim \mathcal{N}\left(D(t, L_D, \boldsymbol{\theta}), \sigma_d^2\right), \tag{4}$$

where $H(t, L_H, \boldsymbol{\theta})$, $U(t, L_H, \boldsymbol{\theta})$, and $D(t, L_D, \boldsymbol{\theta})$ are model-projected hospitalizations census (lagged by L_H), cumulative hospitalizations (lagged by L_H), and cumulative deaths (lagged by L_D) at time t under parameter values $\boldsymbol{\theta}$. Prior distributions imposed on calibrated model parameters ensure that all compartments remain non-negative during the modeling period. We put uniform priors on L_H and L_D over a range of plausible integer values. Reporting lags are correlated with other unknown parameters, including latency period, time between infection and hospitalization, time between infection and death, and length of hospital stay, therefore L_H and L_D should not be strictly interpreted as reporting lags. We put the same independent Inv-Gamma(a, b) prior on all three hyperparameters σ_h^2, σ_u^2, and σ_d^2. The prior was gradually relaxed as number of observations increased.

We construct the posterior distribution over unknown parameters $(\boldsymbol{\theta}, \boldsymbol{\sigma})$ as:

$$\begin{aligned} p(\boldsymbol{\theta}, \boldsymbol{\sigma} | h(t), u(t), d(t)) \propto p(\boldsymbol{\theta}) p(\boldsymbol{\sigma}) & \prod_{t \in t_H} \left[p\left(h(t) | H(t, L_H, \boldsymbol{\theta}), \sigma_h\right)\right]^{w_h z(t)} \\ & \prod_{t \in t_U} \left[p\left(u(t) | U(t, L_H, \boldsymbol{\theta}), \sigma_u\right)\right]^{w_u z(t)} \\ & \prod_{t \in t_D} \left[p\left(d(t) | D(t, L_D, \boldsymbol{\theta}), \sigma_d\right)\right]^{w_d z(t)}, \end{aligned} \tag{5}$$

where $\boldsymbol{\theta} = (\beta_0, q_A, \alpha_{I_S}, m_{H,0}, E_0, L_H, L_D, \tau, \boldsymbol{\epsilon})$ and $\boldsymbol{\sigma} = (\sigma_h, \sigma_u, \sigma_d, \sigma_\epsilon)$.

We assume the date of epidemic onset to be February 16th, 2020 - 21 days before the first case was officially confirmed in Connecticut on March 8th, 2020, and initialize the model with E_0 exposed individuals at the time of epidemic onset, setting the size of all downstream compartments to be zero. County-level distribution of E_0 is fixed and was estimated based on the county population size and dates of first registered case and death in each county.

Each likelihood term is weighted by the time-dependent weight $z(t)$ times the weight assigned to a respective time series. We let the weight function $z(t)$ take the following form,

$$z(t) = \frac{1}{1 + \exp(-k_z t)}, \tag{6}$$

where $z(t)$ is the weight assigned to an observation at time t, $t \in \{t_0, ..., t_{\max}\}$, and the correspondence between $\{t_0, ..., t_{\max}\}$ and calendar time is set such that $t = 0$ corresponds to 90 days prior to the most recent observation. Parameter k_z controls the smoothness of logistic function. We set $k_z = 0.01$, resulting in a range of weights between $0.06 - 0.7$ for observations between March 1, 2020 - February 24, 2021.

We set $w_h = 0.89$, $w_u = 0.01$, and $w_d = 0.1$. We place a large weight on the hospitalizations census, since this time series is most sensitive to changes in epidemic dynamics, and a small weight on cumulative hospitalizations, since it measures a feature that is related to hospitalizations census. The range of observation times differ for different time series. For hospitalizations census and deaths, observation times start with the first non-zero observation. For cumulative hospitalizations, observation times start on May 29, 2020 when this indicator started being reported routinely. The last observation in all three time series is from February 24, 2021.

Sampling from the joint posterior distribution of $(\boldsymbol{\theta}, \boldsymbol{\sigma})$ given in (5) is performed using Markov Chain Monte Carlo (MCMC). We employ a hybrid algorithm that combines elliptical slice sampling (ESS) [54], Gibbs sampling, and Metropolis-Hastings sampling with random walk proposals. We first provide an overview of the steps in drawing samples from the full posterior distribution and then describe each steps in more details. Let $\boldsymbol{\theta} = (\boldsymbol{\theta}_{\text{MH}}, \boldsymbol{\theta}_{\text{ESS}})$, where $\boldsymbol{\theta}_{\text{MH}} = (\beta_0, q_A, \alpha_{I_S}, m_{H,0}, E_0, L_H, L_D)$ and $\boldsymbol{\theta}_{\text{ESS}} = (\tau, \boldsymbol{\epsilon})$. The sampler proceeds with the following steps:

1. Update $\boldsymbol{\theta}_{\text{ESS}}|\boldsymbol{\theta}_{\text{MH}}, \boldsymbol{\sigma}$ using ESS.
2. Update $\boldsymbol{\theta}_{\text{MH}}|\boldsymbol{\theta}_{\text{ESS}}, \boldsymbol{\sigma}$ with a Metropolis-Hastings step.
3. Update hyperparameters $\boldsymbol{\sigma}|\boldsymbol{\theta}$ with a Gibbs sampler step.

Update $\boldsymbol{\theta}_{\text{ESS}}|\boldsymbol{\theta}_{\text{MH}}, \boldsymbol{\sigma}$: We use a rejection-free sampler (ESS) to sample a vector of random effects $\boldsymbol{\epsilon}$ and a testing effect τ. The ESS operates by drawing samples from the ellipse defined by a Gaussian prior, and then accept or reject the samples by evaluating the likelihood component. Within the slice sampling step, the sampler moves along the generated ellipse and always accepts a new set of parameters.

Update $\boldsymbol{\theta}_{\text{MH}}|\boldsymbol{\theta}_{\text{ESS}}, \boldsymbol{\sigma}$: We implement a Metropolis-Hastings algorithm with random walk proposals for $\boldsymbol{\theta}_{\text{MH}}$. Proposals for L_H and L_D are made on a subset of integers bounded by a prior distribution on lags. All other elements of $\boldsymbol{\theta}_{\text{MH}}$ are continuous.

Update hyperparameters $\boldsymbol{\sigma}|\boldsymbol{\theta}$: The hyperparameters σ_h^2, σ_u^2, σ_d^2, σ_ϵ^2 are updated with Gibbs sampler steps:

$$\frac{1}{\sigma_h^2}|h(t), H(t) \sim \text{Gamma}(a + \frac{w_h \sum_{t\in t_H} z(t)}{2}, b + \frac{w_h \sum_{t\in t_H} z(t)(h(t) - H(t))^2}{2})$$
$$\frac{1}{\sigma_u^2}|u(t), U(t) \sim \text{Gamma}(a + \frac{w_u \sum_{t\in t_U} z(t)}{2}, b + \frac{w_u \sum_{t\in t_U} z(t)(u(t) - U(t))^2}{2})$$
$$\frac{1}{\sigma_d^2}|d(t), D(t) \sim \text{Gamma}(a + \frac{w_d \sum_{t\in t_D} z(t)}{2}, b + \frac{w_d \sum_{t\in t_D} z(t)(d(t) - D(t))^2}{2})$$
$$\frac{1}{\sigma_\epsilon^2}|\epsilon_1, ..., \epsilon_K \sim \text{Gamma}(a_\epsilon + \frac{K}{2}, b_\epsilon + \frac{\sum_{i=1}^K (\epsilon_i - \epsilon_{i-1})^2}{2})$$

We run 6 chains of the MCMC sampler for 20,000 iterations each, discard the first 2,000 draws from each chain, thin each chain by a factor of 20, and combine the resulting chains in a single posterior sample. Based on the visual inspection of individual parameter trace plots, we found that 10,000 iterations is sufficient for the chain to converge in practice. To generate uncertainty intervals of model projections, we sample from the joint posterior over estimated parameters, and find pointwise 80%, 90% or 95% posterior predictive intervals for each time point.

2.5 Parameter values and prior distributions

Let $\boldsymbol{\theta}_{\text{MH,CONT}}$ denote the subset of continuous transmission model parameters whose joint distribution is calibrated to observed data and updated with Metropolis-Hastings step: $\boldsymbol{\theta}_{\text{MH, CONT}} = (\beta_0, q_A, \alpha_{I_S}, m_{H,0}, E_0)$. For each individual parameter $\theta \in \boldsymbol{\theta}_{\text{MH,CONT}}$, we specify a fixed support $[\theta_{\min}, \theta_{\max}]$, and put independent beta priors on the transformed parameter, i.e.,

$$\frac{\theta - \theta_{\min}}{\theta_{\max} - \theta_{\min}} \sim \text{Beta}(a_\theta, b_\theta), \tag{7}$$

where the shape parameters a_θ and b_θ are set to let θ have mean μ_θ and standard deviation σ_θ. Table 2 provides the summary of $(\mu_\theta, \sigma_\theta, \theta_{\min}, \theta_{\max})$ for all parameters in $\boldsymbol{\theta}_{\text{MH,CONT}}$ along with data sources. For parameters whose values were fixed, only the mean is given.

Table 2: Values and prior distributions of model parameters

Parameter	Mean	SD	Lower	Upper	Source
q_A	0.4	0.02	0.3	0.5	[55–60]
$q_{I_S,0}/(1-q_A)$	0.065	-	-	-	6.5% of symptomatic infections are severe [3, 14, 61]
β_0	1	0.5	0.01	2.00	A diffuse prior assumed
δ	1/4	-	-	-	[8, 14, 38, 42, 43, 57, 62–68]
$\alpha_{A,0}$	1/7	-	-	-	[42, 69–74]
k_A	0.5	-	-	-	[8, 56, 71, 75, 76]
$\alpha_{I_M,0}$	1/4	-	-	-	[3, 14, 38]; of 4 days, 2 are assumed to be presymptomatic [8, 42, 43]
γ_{R_M}	1/7	-	-	-	[69, 73, 74, 77]
α_{I_S}	1/9	0.03	0.02	0.2	[8, 42, 43, 78, 79]
$m_{H,0}$	0.13	0.05	0.01	0.25	Estimated from Connecticut COVID-19 hospitalization data [50]
$m_{\bar{H}}/m_H$	1.5	0.25	1	2	Assumed
τ	0.08	0.015	0.01	0.15	Assumed
k_n	0.015	-	-	-	Assumed
E_0	150	5	100	200	Assumed
L_H	-	-	5	14	Assumed, uniform prior between lower and upper values
L_D	-	-	5	14	Assumed, uniform prior between lower and upper values

Asymptomatic infections play an important role in transmission of SARS-CoV-2 [40, 41, 44, 80], but estimates of the proportion of infections that do not exhibit symptoms vary substantially [55, 56, 81]. The true proportion of asymptomatic infections is important for projections and policy planning due to its relationship to evolving herd immunity. Reported estimates of asymptomatic proportion range between 6 – 96%, and the authors of a recent review recommend a range between 40 – 45% [55]. Another review reported an overall asymptomatic proportion estimate of 20%, and 31% among the studies that included follow-up [56]. Large population-based studies conducted in Spain and in the U.K. provide estimates of asymptomatic proportion between 22-36% [82, 83]. We calculated age-adjusted weighted average of several estimates available from the literature: Nishiura et al. [58] estimated 30.8% among Japanese citizens evacuated from Wuhan, China. We applied this estimate to age group 20-64 years old. Mizumoto et al. [59] estimated 17.9% among infections on the Diamond Princess cruise ship. We applied this estimate to age

group 65 plus years old, which is the age group, in which most infections occurred. We assumed 65% for age group 0-19 years old, consistent with findings reported by Russell et al. [60], where 4 out of 6 infections in this age group were asymptomatic among passengers of the Diamond Princess. The average was weighted by the age distribution of Connecticut population, resulting in the estimate of $q_A = 0.37$, which is consistent with systematic reviews and population-based studies. At the time of this writing, the best estimate of asymptomatic proportion recommended by the CDC was 40% [57]. We center the prior distribution of q_A at 0.4 and allow it to vary between 0.3 and 0.5. The proportion of severe cases early in the epidemic ($q_{I_S,0}$) is assumed to be 6.5% of symptomatic infections in line with [3, 14, 61]. Combined with the estimates of asymptomatic proportion, this translates into an estimate of severe proportion of 3.9% (3.3-4.6%) of all infections.

Duration of latency ($1/\delta$) was fixed, since consistent estimates of this parameter are available from the literature, and because it is highly correlated with several calibrated parameters, including reporting lags and parameters that determine duration of infectiousness and time between infection and recovery or death. We assume an average of 4 days of latency [8, 14, 38] and 2 days of presymptomatic infectiousness [8, 42, 43] resulting in an average incubation period of 6 days consistent with [57, 62–68].

Parameters (α_A, k_A, α_{I_M}, α_{I_S}) collectively determine the force of infection at any given time. Force of infection, transmission parameter β and initial number of exposed individuals E_0 together determine the early growth of the epidemic. Without additional data, all of these parameters cannot be simultaneously identified. We therefore fixed the values for a subset of these parameters based on available estimates and assumed a diffuse prior on β, which absorbs additional variation of parameters that determine the force of infection.

Duration of infectiousness of asymptomatic individuals is unknown, but is likely shorter than that of symptomatic individuals [69–72]. While several studies estimated that viral RNA could be detected in upper respiratory tract for 2-3 weeks [84, 85], findings from [69, 73, 74] show that in mild-to-moderate symptomatic patients live virus could be isolated for a substantially shorter time period: up to 7-10 days from the day of symptom onset, suggesting the duration of infectiousness of up to 12 days. Based on these estimates and [42], we assume the duration of infectiousness of 7 days among asymptomatic individuals. Although multiple studies have shown similar viral loads among symptomatic, presymptomatic and asymptomatic cases [70, 86, 87], evidence suggests that asymptomatic individuals are less infectious that symptomatic, likely due to higher viral shedding while coughing and longer duration of infectiousness among symptomatic individuals [56, 69, 71, 88, 89]. Multiple estimates of relative infectiousness of asymptomatic individuals compared to symptomatic are available and range from close to zero to above one [8, 56, 71, 75, 76]. Generally, estimates that are based on attack rates are lower than those based on viral shedding or time until the first negative PCR test. In our analysis, we set relative infectiousness of asymptomatic cases to be 0.5 [8, 56, 75].

Although the duration of infectiousness of mild-to-moderately symptomatic cases may be up to 12 days [69, 73, 74], we assume that the majority of symptomatic individuals self-isolate shortly after developing symptoms. We set the duration of infectiousness of symptomatic cases to be 4 days [3, 14, 38], 2 of which are assumed to represent presymptomatic infectiousness [8, 42, 43]. We further assume that the duration of self-isolation until recovery is 7 days [69, 73, 74].

We calibrate the rate of hospitalization among severe cases (α_{I_S}), assuming a mean of 9 days between the onset of infectiousness and hospitalization: 2 days of presymptomatic infectiousness plus 7 days between the onset of symptoms and hospitalization [78, 79].

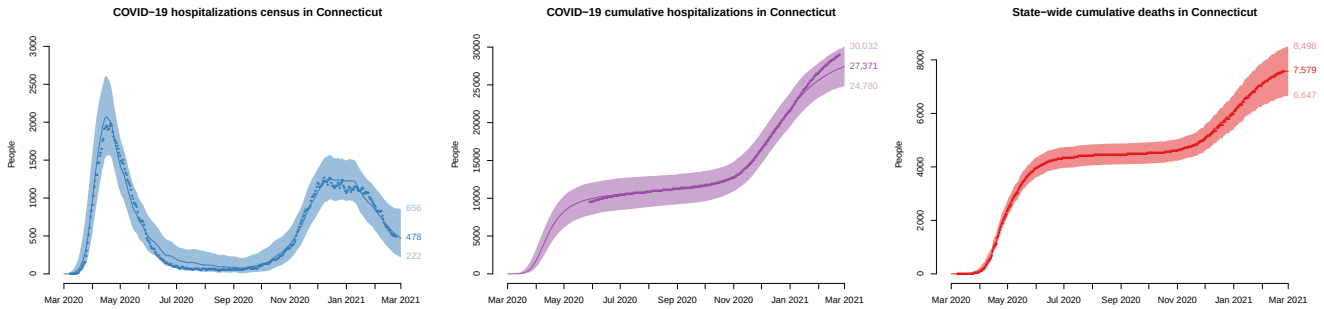

Figure 3: Model calibration results for observed COVID-19 positive hospitalizations census, cumulative hospitalizations and cumulative deaths in Connecticut. Solid lines represent model-projected means and shaded regions represent 95% posterior predictive intervals. Observed time series are shown as points and correspond to total hospitalizations and deaths among all Connecticut residents. The model is calibrated to estimated data series coming from non-congregate settings, and model projections are adjusted by the estimated difference to reflect the totals for congregate and non-congregate settings.

Table 3: Means and 95% credible interval of marginal posterior distributions of model parameters and estimated epidemiologic parameters.

parameter	mean	CI-low	CI-high
β_0	1.15	0.98	1.28
q_A	0.4	0.36	0.44
α_{I_S}	0.1	0.05	0.16
$m_{H,0}$	0.12	0.07	0.17
E_0	150	140	160
crude R_0	4.86	4.06	5.60
cumulative CDR	0.33	0.21	0.59
cumulative IHR	0.036	0.031	0.042
cumulative IFR	0.0096	0.0082	0.0112

CDR, case detection ratio (proportion of all infections detected); IHR, infection hospitalization ratio (proportion of all infections hospitalized); IFR, infection fatality ratio (proportion of all infections leading to death). IHR and IFR include residents of congregate settings.

3 Results

Figure 3 shows the results of model calibration, including the fit to observed dynamics of hospitalizations and deaths. Observed data points track with mean projections and fall within uncertainty intervals. Table 3 shows marginal means and 95% posterior credible intervals of calibrated model parameters and estimated epidemiologic parameters.

Figure 4 shows model projections of several epidemiologic features of SARS-CoV-2 epidemic in Connecticut. We estimate that effective reproduction number (R_{eff}) dropped substantially in mid-March and remained below one through mid-June - early July. For the rest of the summer, mean estimated R_{eff} was slightly above one consistent with low numbers of case counts and hospitalizations in the summer. A major increase of R_{eff} started in late August - early September with the reopening of schools and colleges. It reached a maximum mean value of 1.4 by mid-October, followed by a slow decline through the rest of the year. The dynamics of R_{eff} follows closely the dynamics of close contact. The top right plot of Figure 4 shows that the estimated dynamics of transmission parameter captured by a community contact function (measure of close contact adjusted for estimated random effects) exhibits small deviations from the measure of close contact.

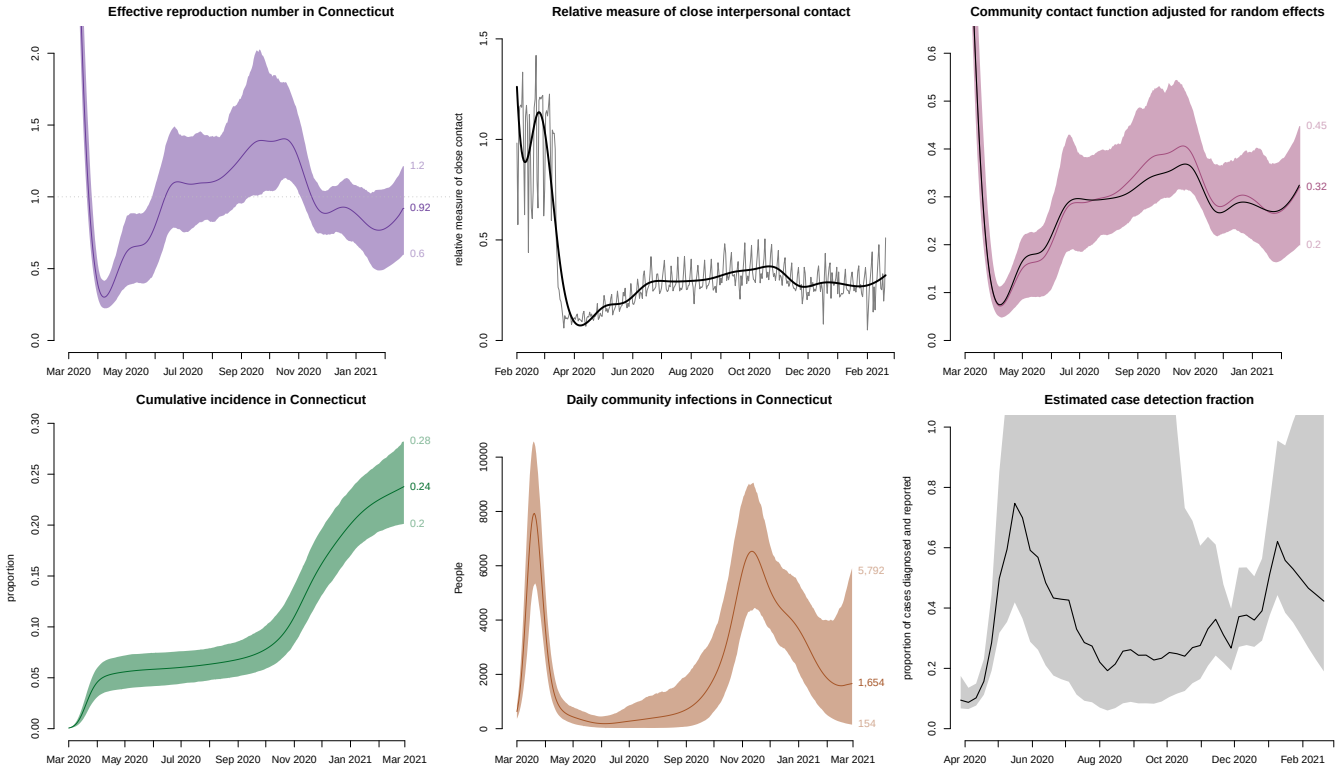

Figure 4: Model projections for selected epidemiologic features of SARS-CoV-2 transmission in Connecticut. Top row from left to right: effective reproduction number, normalized measure of close interpersonal contact relative to the pre-epidemic period (thick solid line shows spline smoothing of the contact measure), and contact function adjusted for estimated random effects, which capture residual variation in transmission that is not explained by the dynamics of close contact and other time-varying parameters. In this plot, black line shows smoothed close contact unadjusted for random effects. Bottom row from left to right: cumulative incidence, daily new infections, and estimated infection detection fraction in non-congregate settings in Connecticut. Solid lines represent model-projected means and shaded regions represent 95% posterior predictive intervals.

We estimate that at the end of February of 2021, cumulative incidence in Connecticut was about 24%, which implies a cumulative case detection ratio of 33% (95%CI: 21 - 59%). Our estimate of cumulative incidence at the beginning of June is 5.8% (95%CI: 4.2 - 7.4%) consistent with community-based seroprevalence surveys conducted in Connecticut between April and June [90, 91]. Our projections suggest that case detection rate varied substantially over time in a way that cannot be explained by the testing volume alone. An increase up to 75% in mid-May is likely due to delayed testing, including postmortem diagnosis of first epidemic wave cases. The second spike in estimated case detection rate in early January is likely a consequence of testing and reporting disruptions related to Christmas and New Year holidays.

4 Discussion

In this report, we have described technical details of a model of SARS-CoV-2 transmission and COVID-19 disease progression developed to support public health decision-making in Connecticut. The model is calibrated to the observed dynamics of hospitalizations and cumulative deaths in Connecticut; its projections reproduce these dynamics accurately. Many COVID-19 models have been developed and analyzed by the CDC in the attempt to perform ensemble forecasting of the epidemic development in the US [2]. Some of these models offer state-level projections [25–35]. The CDC publishes updates of consolidated summary of cumulative death projections in the next four weeks from these models for each state. Local projections from nation-wide models may offer useful insights, however simplifying or uniform assumptions made in most of these models may not hold in all of the locations.

There is substantial uncertainty about epidemiologic parameters that govern aspects of COVID-19 dynamics and have a direct impact on the quality of projections. When local context is not directly taken into account, the effects of parametric uncertainty is exacerbated. State and county-level models are needed to support local decision-making. Our model captures distinct important features of COVID-19 dynamics and the relationship between model features and data reporting in Connecticut. The ability to capture temporal dynamics in key model parameters provides a substantial improvement to the model fit, accuracy and credibility. Indeed, incorporating time trends in such important epidemic features as close contact, risk profile of incident cases, hospital length of stay, and hospital case fatality ratio allowed us to achieve a good model fit to observed data and improve predictive performance of the model over time. However, the calibrated posterior distribution of model parameters is not necessarily generalizable to other settings: model projections are tightly linked to the Connecticut context.

In addition to providing predictions for policymakers, model projections may be useful for prospectively planning epidemiological studies that can inform the state's response. In particular, planning of seroprevalence surveys requires estimates of the proportion of population who have evidence of prior exposure to the virus. Due to limited testing availability and potentially high proportion of asymptomatic individuals, official case counts offer a poor approximation to the true cumulative incidence. Seroprevalence surveys, if properly conducted, can provide an important piece of information that would permit more precise estimates of the fraction of asymptomatic infections.

Prior knowledge and assumptions about plausible ranges of parameter values combined with local data allows us to substantially reduce parametric uncertainty and produce narrow projection intervals. However, several important considerations limit our ability to make reliable long-term projections. First, it is difficult to make predictions about the extent of changes time-varying parameters in the future, in particular changes in human behavior leading to higher or lower rates of interpersonal contact. Second, the effectiveness of widespread testing and contact tracing on timely isolation of infectious individuals, and its subsequent impact on the force of infection is unknown. This effect is a complex function of viral shedding characteristics among symptomatic, presymptomatic and asymptomatic individuals, along with the implementation features of contact tracing, testing, and isolation [92]. Third, even individuals residing in non-congregate settings may not be mixing at random, and there may be unequal depletion of susceptible individuals in various subgroups, leading to lower immunization thresholds necessary to achieve herd immunity [22, 23].

The work presented in this paper and other local models of COVID-19 transmission show that local models are more useful to local policymakers because they are able to incorporate vast data sources and local context information that is only available at the local level. As new data sources become available, the model may be extended to reflect more granular geographic variation, age structure, vaccination, and other important features of the epidemic dynamics.

Acknowledgements: We are grateful to Edward H. Kaplan for helpful comments. We thank the Connecticut Hospital Association for providing data on COVID-19 hospitalizations and deaths. We thank Stony Brook Research Computing and Cyberinfrastructure, and the Institute for Advanced Computational Science at Stony Brook University for access to the high-performance SeaWulf computing system, which was made possible by a USD $1.4M National Science Foundation grant (# 1531492).

Funding: We acknowledge NIH/NICHD grant 1DP2HD091799-01.

Author contributions: Conceptualization: FWC, OM; Methodology Development: FWC, OM, ZRL; Software Development: FWC, OM, ZRL; Validation: OM; Formal Analysis: OM, ZRL; Resources: FWC; Data Curation: OM, FWC; Writing – Original Draft: OM; Writing – Review & Editing: FWC, ZRL, OM; Visualization Preparation: FWC, OM, ZRL; Supervision Oversight: FWC; Project Administration: FWC, OM; Funding Acquisition: FWC.

Competing interests: FWC is a paid consultant to Whitespace Solutions.

References

[1] Inga Holmdahl and Caroline Buckee. Wrong but useful – what Covid-19 epidemiologic models can and cannot tell us. *New England Journal of Medicine*, In press, 2020.

[2] Centers for Disease Control and Prevention. COVID-19 Forecasts. https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html, 2020. Accessed: 2020-05-17.

[3] Stephen M Kissler, Christine Tedijanto, Edward Goldstein, Yonatan H Grad, and Marc Lipsitch. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science*, 2020.

[4] Neil M Ferguson, Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Zulma Cucunubá, Gina Cuomo-Dannenburg, Amy Dighe, Ilaria Dorigatti, Han Fu, Katy Gaythorpe, Will Green, Arran Hamlet, Wes Hinsley, Lucy C Okell, Sabine van Elsland, Hayley Thompson, Robert Verity, Erik Volz, Haowei Wang, Yuanrong Wang, Patrick GT Walker, Caroline Walters, Peter Winskill, Charles Whittaker, Christl A Donnelly, Steven Riley, and Azra C Ghani. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand (Imperial College COVID-19 Response Team), 2020.

[5] H Juliette T Unwin, Swapnil Mishra, Valerie C Bradley, Axel Gandy, Michaela Vollmer, Thomas Mellan, Helen Coupland, Kylie Ainslie, Charles Whittaker, Jonathan Ish-Horowicz, Sarah Lucie Filippi, Xiaoyue Xi, Melodie Monod, Oliver Ratmann, Michael Hutchinson, Fabian Valka, Harrison Zhu, Iwona Hawryluk, Philip Milton, Marc Baguelin, Adhiratha Boonyasiri, Nick Brazeau, Lorenzo Cattarino, Giovanni Charles, Laura V Cooper, Zulma Cucunuba, Gina Cuomo-Dannenburg, Bimandra Djaafara, Ilaria Dorigatti, Oliver J Eales, Jeff Eaton, Sabine van Elsland, Richard FitzJohn, Katy Gaythorpe, William Green, Timothy Hallett, Wes Hinsley, Natsuko Imai, Ben Jeffrey, Edward Knock, Daniel Laydon, John Lees, Gemma Nedjati-Gilani, Pierre Nouvellet, Lucy Okell, Alison Ower, Kris V Parag, Hayley A Thompson, Robert Verity, Patrick Walker, Caroline Walters, Yuanrong Wang, Oliver J Watson, Lilith Whittles, Azra Ghani, Neil M Ferguson, Steven Riley, Christl Donnelly, Samir Bhatt, and Seth Flaxman. Report 23: State-level tracking of COVID-19 in the United States (Imperial College COVID-19 Response Team), 2020.

[6] Tobias Brett and Pejman Rohani. COVID-19 herd immunity strategies: walking an elusive and dangerous tightrope. *medRxiv*, 2020.

[7] Andrew C Miller, Nicholas J Foti, Joseph A Lewnard, Nicholas P Jewell, Carlos Guestrin, and Emily B Fox. Mobility trends provide a leading indicator of changes in SARS-CoV-2 transmission. *medRxiv*, 2020.

[8] Alberto Aleta, David Martin-Corral, Ana Pastore y Piontti, Marco Ajelli, Maria Litvinova, Matteo Chinazzi, Natalie E Dean, M. Elizabeth Halloran, Ira M Longini, Stefano Merler, Alex Pentland, Alessandro Vespignani, Esteban Moro, and Yamir Moreno. Modeling the impact of social distancing, testing, contact tracing and household quarantine on second-wave scenarios of the COVID-19 epidemic. *medRxiv*, 2020.

[9] Marissa L Childs, Morgan P Kain, Devin Kirk, Mallory Harris, Lisa Couper, Nicole Nova, Isabel Delwel, Jacob Ritchie, and Erin A Mordecai. The impact of long-term non-pharmaceutical interventions on COVID-19 epidemic dynamics and control. *medRxiv*, 2020.

[10] Teresa Yamana, Sen Pei, Sasikiran Kandula, and Jeffrey Shaman. Projection of COVID-19 cases and deaths in the US as individual states re-open May 4, 2020. *medRxiv*, 2020.

[11] Difan Zou, Lingxiao Wang, Pan Xu, Jinghui Chen, Weitong Zhang, and Quanquan Gu. Epidemic model guided machine learning for COVID-19 forecasts in the United States. *medRxiv*, 2020.

[12] Ting Tian, Jianbin Tan, Yukang Jiang, Xueqin Wang, and Heping Zhang. Evaluate the timing of resumption of business for the states of New York, New Jersey, and California via a pre-symptomatic and asymptomatic transmission model of COVID-19. *MedRxiv*, 2020.

[13] Francisco Perez-Reche and Norval Strachan. Importance of untested infectious individuals for the suppression of COVID-19 epidemics. *MedRxiv*, 2020.

[14] Henrik Salje, Cecile Tran Kiem, Noemie Lefrancq, Noemie Courtejoie, Paolo Bosetti, Juliette Paireau, Alessio Andronico, Nathanael Hoze, Jehanne Richet, Claire-Lise Dubost, Yann Le Strat, Justin Lessler, Daniel Levy Bruhl, Arnaud Fontanet, Lulla Opatowski, Pierre-Yves Boelle, and Simon Cauchemez. Estimating the burden of SARS-CoV-2 in France. *medRxiv*, 2020.

[15] Laura Di Domenico, Giulia Pullano, Chiara E Sabbatini, Pierre-Yves Boëlle, and Vittoria Colizza. Expected impact of reopening schools after lockdown on COVID-19 epidemic in Île-de-France. *medRxiv*, 2020.

[16] Jonas Dehning, Johannes Zierenberg, F Paul Spitzner, Michael Wibral, Joao Pinheiro Neto, Michael Wilczek, and Viola Priesemann. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science*, 2020.

[17] Luca Ferretti, Chris Wymant, Michelle Kendall, Lele Zhao, Anel Nurtay, Lucie Abeler-Dörner, Michael Parker, David Bonsall, and Christophe Fraser. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*, March 2020.

[18] Kiesha Prem, Yang Liu, Timothy W Russell, Adam J Kucharski, Rosalind M Eggo, Nicholas Davies, Stefan Flasche, Samuel Clifford, Carl A B Pearson, James D Munday, Sam Abbott, Hamish Gibbs, Alicia Rosello, Billy J Quilty, Thibaut Jombart, Fiona Sun, Charlie Diamond, Amy Gimma, Kevin van Zandvoort, Sebastian Funk, Christopher I Jarvis, W John Edmunds, Nikos I Bosse, Joel Hellewell, Mark Jit, and Petra Klepac. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *The Lancet Public Health*, March 2020. ISSN 2468-2667.

[19] Matteo Chinazzi, Jessica T. Davis, Marco Ajelli, Corrado Gioannini, Maria Litvinova, Stefano Merler, Ana Pastore y Piontti, Kunpeng Mu, Luca Rossi, Kaiyuan Sun, Cécile Viboud, Xinyue Xiong, Hongjie Yu, M. Elizabeth Halloran, Ira M. Longini, and Alessandro Vespignani. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*, March 2020.

[20] Adam J Kucharski, Timothy W Russell, Charlie Diamond, Yang Liu, John Edmunds, Sebastian Funk, Rosalind M Eggo, and Centre for Mathematical Modelling of Infectious Diseases COVID-19 working group. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The Lancet Infectious Diseases*, 2020.

[21] Kathy Leung, Joseph T Wu, Di Liu, and Gabriel M Leung. First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. *The Lancet*, 2020.

[22] Tom Britton, Frank Ball, and Pieter Trapman. The disease-induced herd immunity level for Covid-19 is substantially lower than the classical herd immunity level. *arXiv preprint arXiv:2005.03085*, 2020.

[23] M Gabriela M Gomes, Ricardo Aguas, Rodrigo M Corder, Jessica G King, Kate E Langwig, Caetano Souto-Maior, Jorge Carneiro, Marcelo U Ferreira, and Carlos Penha-Goncalves. Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold. *medRxiv*, 2020.

[24] Jonathan Fintzi, Damon Bayer, Isaac Goldstein, Keith Lumbard, Emily Ricotta, Sarah Warner, Lindsay M Busch, Jeffrey R Strich, Daniel S Chertow, Daniel M Parker, Bernadette Boden-Albala, Alissa Dratch, Richard Chhuon, Nichole Quick, Matthew Zahn, and Vladimir N Minin. Using multiple data streams to estimate and forecast sars-cov-2 transmission dynamics, with application to the virus spread in orange county, california. *arXiv preprint arXiv:2009.02654*, 2020.

[25] Imperial College COVID-19 Response Team. COVID-19 epidemic in the United States: state-level projections. https://mrc-ide.github.io/covid19usa/#/, 2020. Accessed: 2020-05-27.

[26] Max Henderson, Rep Jonathan Kreiss-Tomkins, Igor Kofman, and Zack Rosen. Covid Act Now. https://covidactnow.org, 2020. Accessed: 2020-05-17.

[27] Northeastern University Network Science Institute. COVID-19 modeling: United States. https://covid19.gleamproject.org, 2020. Accessed: 2020-05-27.

[28] Columbia University Mailman School of Public Health Epidemiology. COVID-19 risk mapping in the US. https://columbia.maps.arcgis.com/apps/webappviewer/index.html?id=ade6ba85450c4325a12a5b9c09ba796c, 2020. Accessed: 2020-05-27.

[29] IHME. IHME | COVID-19 Projections. https://covid19.healthdata.org/projections, 2020. Accessed: 2020-05-27.

[30] Los Alamos National Laboratory. COVID-19 U.S. Forecasts. https://covid-19.bsvgateway.org, 2020. Accessed: 2020-05-27.

[31] MIT COVID Analytics. Predictions of infections and deaths under a variety of policies. https://www.covidanalytics.io/policies, 2020. Accessed: 2020-05-27.

[32] UCLA. Combating COVID-19. https://covid19.uclaml.org, 2020. Accessed: 2020-05-27.

[33] UMass-Amherst Influenza Forecasting Center of Excellence. Reich lab COVID-19 forecast hub. https://reichlab.io/covid19-forecast-hub/, 2020. Accessed: 2020-05-27.

[34] The University of Texas COVID-19 Modeling Consortium. COVID-19 mortality projections for US states and metropolitan areas. https://covid-19.tacc.utexas.edu/projections, 2020. Accessed: 2020-05-27.

[35] Youyang Gu. COVID-19 projections using machine learning. https://covid19-projections.com, 2020. Accessed: 2020-05-27.

[36] Forrest W. Crawford, Zehang Richard Li, and Olga Morozova. COVID-19 projections for reopening Connecticut. *medRxiv*, 2020. doi: 10.1101/2020.06.16.20126425. URL https://www.medrxiv.org/content/early/2020/06/19/2020.06.16.20126425.

[37] Matt J Keeling and Pejman Rohani. *Modeling infectious diseases in humans and animals*. Princeton University Press, 2011.

[38] Ruiyun Li, Sen Pei, Bin Chen, Yimeng Song, Tao Zhang, Wan Yang, and Jeffrey Shaman. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science*, 2020.

[39] COVID-19 Statistics, Policy modeling, and Epidemiology Collective and Joshua A Salomon. Defining high-value information for COVID-19 decision-making. *medRxiv*, 2020.

[40] Nathan W Furukawa, John T Brooks, and Jeremy Sobel. Evidence supporting transmission of severe acute respiratory syndrome coronavirus 2 while presymptomatic or asymptomatic. *Emerging Infectious Diseases*, 26(7), 2020.

[41] Yang Liu, Sebastian Funk, and Stefan Flasche. The contribution of pre-symptomatic transmission to the COVID-19 outbreak, 2020.

[42] Andrew William Byrne, David McEvoy, Aine B Collins, Kevin Hunt, Miriam Casey, Ann Barber, Francis Butler, John Griffin, Elizabeth A Lane, Conor McAloon, Kirsty O'Brien, Patrick Wall, Kieran A Walsh, and Simon J More. Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases. *BMJ Open*, 10(8), 2020. ISSN 2044-6055. doi: 10.1136/bmjopen-2020-039856. URL https://bmjopen.bmj.com/content/10/8/e039856.

[43] Wycliffe E Wei, Zongbin Li, Calvin J Chiew, Sarah E Yong, Matthias P Toh, and Vernon J Lee. Presymptomatic transmission of SARS-CoV-2 – Singapore, January 23–March 16, 2020. *Morbidity and Mortality Weekly Report*, 69(14):411, 2020.

[44] Xi He, Eric H. Y. Lau, Peng Wu, Xilong Deng, Jian Wang, Xinxin Hao, Yiu Chung Lau, Jessica Y. Wong, Yujuan Guan, Xinghua Tan, Xiaoneng Mo, Yanqing Chen, Baolin Liao, Weilie Chen, Fengyu Hu, Qing Zhang, Mingqiu Zhong, Yanrong Wu, Lingzhai Zhao, Fuchun Zhang, Benjamin J. Cowling, Fang Li, and Gabriel M. Leung. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine*, 26(5): 672–675, 2020.

[45] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. URL https://www.R-project.org/.

[46] Karline Soetaert, Thomas Petzoldt, and R. Woodrow Setzer. Solving differential equations in R: Package deSolve. *Journal of Statistical Software*, 33(9):1–25, 2010.

[47] United States Census Bureau. 2010 cartographic boundary file, current block group for Connecticut. Data retrieved from http://magic.lib.uconn.edu/connecticut_data.html, 2010. Accessed: 2020-04-14.

[48] Forrest W. Crawford, Sydney Jones, Matthew Cartter, Samantha G. Dean, Joshua L. Warren, Zehang Richard Li, Jacqueline Barbieri, Jared Campbell, Patrick Kenney, Thomas Valleau, Robert Zalot, and Olga Morozova. Impact of close interpersonal contact on COVID-19 incidence: evidence from one year of mobile device data. *Working Paper*, 2021. URL https://forrestcrawford.shinyapps.io/ct_social_distancing/.

[49] Tegan K Boehmer, Jourdan DeVies, Elise Caruso, Katharina L van Santen, Shichao Tang, Carla L Black, Kathleen P Hartnett, Aaron Kite-Powell, Stephanie Dietz, Matthew Lozier, et al. Changing age distribution of the covid-19 pandemic—united states, may–august 2020. *Morbidity and Mortality Weekly Report*, 69(39): 1404, 2020.

[50] Connecticut Hospital Association, 2020. URL https://cthosp.org/.

[51] Connecticut State Department of Public Health. COVID-19 Data Resources, 2021. URL https://data.ct.gov/stories/s/COVID-19-data/wa3g-tfvc/.

[52] United States Census Bureau. *American Community Survey (ACS)*, 2020. URL https://www.census.gov/programs-surveys/acs.

[53] CHIMEData, 2020. URL https://chimedata.org/.

[54] Iain Murray, Ryan Adams, and David MacKay. Elliptical slice sampling. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 541–548, 2010.

[55] Daniel P Oran and Eric J Topol. Prevalence of asymptomatic sars-cov-2 infection: A narrative review. *Annals of Internal Medicine*, 2020.

[56] Diana Buitrago-Garcia, Dianne Egli-Gany, Michel J Counotte, Stefanie Hossmann, Hira Imeri, Aziz Mert Ipekci, Georgia Salanti, and Nicola Low. Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis. *PLoS medicine*, 17(9):e1003346, 2020.

[57] Centers for Disease Control and Prevention. COVID-19 Pandemic Planning Scenarios. https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html, 2021. Accessed: 2021-01-18.

[58] Hiroshi Nishiura, Tetsuro Kobayashi, Takeshi Miyama, Ayako Suzuki, Sung-Mok Jung, Katsuma Hayashi, Ryo Kinoshita, Yichi Yang, Baoyin Yuan, Andrei R Akhmetzhanov, and Natalie M Linton. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *International Journal of Infectious Diseases*, 94:154 – 155, 2020.

[59] Kenji Mizumoto, Katsushi Kagaya, Alexander Zarebski, and Gerardo Chowell. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Eurosurveillance*, 25(10):2000180, 2020.

[60] Timothy W Russell, Joel Hellewell, Christopher I Jarvis, Kevin Van Zandvoort, Sam Abbott, Ruwan Ratnayake, Stefan Flasche, Rosalind M Eggo, W John Edmunds, and Adam J Kucharski. Estimating the infection and case fatality ratio for coronavirus disease (COVID-19) using age-adjusted data from the outbreak on the Diamond Princess cruise ship, February 2020. *Euro surveillance*, 25(12):2000256, 2020.

[61] Robert Verity, Lucy C Okell, Ilaria Dorigatti, Peter Winskill, Charles Whittaker, Natsuko Imai, Gina Cuomo-Dannenburg, Hayley Thompson, Patrick GT Walker, Han Fu, Amy Dighe, Jamie T Griffin, Marc Baguelin, Sangeeta Bhatia, Anne Boonyasiri, Adhiratha andd Cori, Zulma Cucunubá, Rich FitzJohn, Katy Gaythorpe, Will Green, Arran Hamlet, Wes Hinsley, Daniel Laydon, Gemma Nedjati-Gilani, Steven Riley, Sabine van

Elsland, Erik Volz, Haowei Wang, Yuanrong Wang, Xiaoyue Xi, Christl A Donnelly, Azra C Ghani, and Neil M Ferguson. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet Infectious Diseases*, 2020.

[62] Matthew Biggerstaff, Benjamin J Cowling, Zulma M Cucunubá, Linh Dinh, Neil M Ferguson, Huizhi Gao, Verity Hill, Natsuko Imai, Michael A Johansson, Sarah Kada, et al. Early insights from statistical and mathematical modeling of key epidemiologic parameters of COVID-19. *Emerging infectious diseases*, 26 (11), 2020.

[63] Conor McAloon, Áine Collins, Kevin Hunt, Ann Barber, Andrew W Byrne, Francis Butler, Miriam Casey, John Griffin, Elizabeth Lane, David McEvoy, et al. Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research. *BMJ open*, 10(8):e039652, 2020.

[64] Stephen A Lauer, Kyra H Grantz, Qifang Bi, Forrest K Jones, Qulu Zheng, Hannah R Meredith, Andrew S Azman, Nicholas G Reich, and Justin Lessler. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Annals of Internal Medicine*, 2020.

[65] Jantien A Backer, Don Klinkenberg, and Jacco Wallinga. Incubation period of 2019 novel coronavirus (2019-ncov) infections among travellers from wuhan, china, 20–28 january 2020. *Eurosurveillance*, 25(5), 2020.

[66] Qun Li, Xuhua Guan, Peng Wu, Xiaoye Wang, Lei Zhou, Yeqing Tong, Ruiqi Ren, Kathy S.M. Leung, Eric H.Y. Lau, Jessica Y. Wong, Xuesen Xing, Nijuan Xiang, Yang Wu, Chao Li, Qi Chen, Dan Li, Tian Liu, Jing Zhao, Man Liu, Wenxiao Tu, Chuding Chen, Lianmei Jin, Rui Yang, Qi Wang, Suhua Zhou, Rui Wang, Hui Liu, Yinbo Luo, Yuan Liu, Ge Shao, Huan Li, Zhongfa Tao, Yang Yang, Zhiqiang Deng, Boxi Liu, Zhitao Ma, Yanping Zhang, Guoqing Shi, Tommy T.Y. Lam, Joseph T. Wu, George F. Gao, Benjamin J. Cowling, Bo Yang, Gabriel M. Leung, and Zijian Feng. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New England Journal of Medicine*, 382(13):1199–1207, 2020.

[67] Natalie M Linton, Tetsuro Kobayashi, Yichi Yang, Katsuma Hayashi, Andrei R Akhmetzhanov, Sung-mok Jung, Baoyin Yuan, Ryo Kinoshita, and Hiroshi Nishiura. Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data. *Journal of Clinical Medicine*, 9(2):538, 2020.

[68] Jing Qin, Chong You, Qiushi Lin, Taojun Hu, Shicheng Yu, and Xiao-Hua Zhou. Estimation of incubation period distribution of COVID-19 using disease onset forward time: A novel cross-sectional and forward follow-up study. *Science Advances*, 6(33), 2020. doi: 10.1126/sciadv.abc1202. URL https://advances.sciencemag.org/content/6/33/eabc1202.

[69] Muge Cevik, Matthew Tate, Ollie Lloyd, Alberto Enrico Maraolo, Jenna Schafers, and Antonia Ho. SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. *The Lancet Microbe*, 2020.

[70] Kieran A Walsh, Karen Jordan, Barbara Clyne, Daniela Rohde, Linda Drummond, Paula Byrne, Susan Ahern, Paul G Carty, Kirsty K O'Brien, Eamon O'Murchu, et al. SARS-CoV-2 detection, viral load and infectivity over the course of an infection. *Journal of Infection*, 2020.

[71] Xueting Qiu, Ali Ihsan Nergiz, Alberto Enrico Maraolo, Isaac I. Bogoch, Nicola Low, and Muge Cevik. Differences in secondary attack rates based on symptom status of index case(s)– a living systematic review. *medRxiv*, 2021. doi: 10.1101/2020.09.01.20135194. URL https://www.medrxiv.org/content/early/2021/01/16/2020.09.01.20135194.

[72] Rongrong Yang, Xien Gui, and Yong Xiong. Comparison of clinical characteristics of patients with asymptomatic vs symptomatic coronavirus disease 2019 in Wuhan, China. *JAMA Network Open*, 3(5):e2010182–e2010182, 2020.

[73] Kieran A Walsh, Susan Spillane, Laura Comber, Karen Cardwell, Patricia Harrington, Jeff Connell, Conor Teljeur, Natasha Broderick, Cillian F de Gascun, Susan M Smith, et al. The duration of infectiousness of individuals infected with SARS-CoV-2. *Journal of Infection*, 2020.

[74] Roman Wölfel, Victor M. Corman, Wolfgang Guggemos, Michael Seilmaier, Sabine Zange, Marcel A. Müller, Daniela Niemeyer, Terry C. Jones, Patrick Vollmar, Camilla Rothe, Michael Hoelscher, Tobias Bleicker, Sebastian Brünink, Julia Schneider, Rosina Ehmann, Katrin Zwirglmaier, Christian Drosten, and Clemens

Wendtner. Virological assessment of hospitalized patients with COVID-2019. *Nature*, pages 1–5, 2020.

[75] David Mc Evoy, Conor G McAloon, Aine B Collins, Kevin Hunt, Francis Butler, Andrew W Byrne, Miriam Casey, Ann Barber, John M Griffin, Elizabeth A Lane, et al. The relative infectiousness of asymptomatic SARS-CoV-2 infected persons compared with symptomatic individuals: A rapid scoping review. *medRxiv*, 2020.

[76] Daihai He, Shi Zhao, Qianying Lin, Zian Zhuang, Peihua Cao, Maggie H Wang, and Lin Yang. The relative transmissibility of asymptomatic COVID-19 infections among close contacts. *International Journal of Infectious Diseases*, 2020.

[77] Centers for Disease Control and Prevention. Discontinuation of isolation for persons with COVID-19 not in healthcare settings. interim guidance. https://www.cdc.gov/coronavirus/2019-ncov/hcp/disposition-in-home-patients.html, 2020. Accessed: 2021-01-18.

[78] Shikha Garg, Lindsay Kim, Michael Whitaker, Alissa O'Halloran, Charisse Cummings, Rachel Holstein, Mila Prill, Shua J Chai, Pam D Kirley, Nisha B Alden, Breanna Kawasaki, Kimberly Yousey-Hindes, Linda Niccolai, Evan J Anderson, Kyle P Openo, Andrew Weigel, Maya L Monroe, Patricia Ryan, Justin Henderson, Sue Kim, Kathy Como-Sabetti, Ruth Lynfield, Daniel Sosin, Salina Torres, Alison Muse, Nancy M Bennett, Laurie Billing, Melissa Sutton, Nicole West, William Schaffner, H. Keipp Talbot, Aquino Clarissa, Andrea George, Alicia Budd, Lynnette Brammer, Gayle Langley, Aron J Hall, and Alicia Fry. Hospitalization rates and characteristics of patients hospitalized with laboratory-confirmed coronavirus disease 2019 – COVID-NET, 14 states, March 1–30, 2020. *MMWR. Morbidity and Mortality Weekly Report*, 69, 2020.

[79] Pablo N Perez-Guzman, Anna Daunt, Sujit Mukherjee, Peter Crook, Roberta Forlano, Mara D Kont, Alessandra Løchen, Michaela Vollmer, Paul Middleton, Rebekah Judge, , Chris Harlow, Anet Soubieres, Graham Cooke, Peter J White, Timothy B Hallett, Paul Aylin, Neil Ferguson, Katharina Hauck, Mark Thursz, and Shevanthi Nayagam. Clinical characteristics and predictors of outcomes of hospitalized patients with coronavirus disease 2019 in a multiethnic London national health service trust: a retrospective cohort study. 2020.

[80] Jon C Emery, Timothy W Russel, Yang Liu, Joel Hellewell, Carl AB Pearson, CMMID 2019-nCoV working group, Gwen M Knight, Rosalind M Eggo, Adam J Kucharski, Sebastian Funk, Stefan Flasche, and Rein M G J Houben. The contribution of asymptomatic SARS-CoV-2 infections to transmission-a model-based analysis of the Diamond Princess outbreak. *medRxiv*, 2020.

[81] Mercedes Yanes-Lane, Nicholas Winters, Federica Fregonese, Mayara Bastos, Sara Perlman-Arrow, Jonathon R Campbell, and Dick Menzies. Proportion of asymptomatic infection among COVID-19 positive persons and their transmission potential: A systematic review and meta-analysis. *PloS one*, 15(11): e0241536, 2020.

[82] Marina Pollán, Beatriz Pérez-Gómez, Roberto Pastor-Barriuso, Jesús Oteo, Miguel A Hernán, Mayte Pérez-Olmeda, Jose L Sanmartín, Aurora Fernández-García, Israel Cruz, Nerea Fernández de Larrea, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. *The Lancet*, 396(10250):535–544, 2020.

[83] Helen Ward, Christina J Atchison, Matthew Whitaker, Kylie E. C. Ainslie, Joshua Elliott, Lucy C Okell, Rozlyn Redd, Deborah Ashby, Christl A. Donnelly, Wendy Barclay, Ara Darzi, Graham Cooke, Steven Riley, and Paul Elliott. Antibody prevalence for SARS-CoV-2 in England following first peak of the pandemic: REACT2 study in 100,000 adults. *medRxiv*, 2020. doi: 10.1101/2020.08.12.20173690. URL https://www.medrxiv.org/content/early/2020/08/21/2020.08.12.20173690.

[84] Barnaby Edward Young, Sean Wei Xiang Ong, Shirin Kalimuddin, Jenny G. Low, Seow Yen Tan, Jiashen Loh, Oon-Tek Ng, Kalisvar Marimuthu, Li Wei Ang, Tze Minn Mak, Sok Kiang Lau, Danielle E. Anderson, Kian Sing Chan, Thean Yen Tan, Tong Yong Ng, Lin Cui, Zubaidah Said, Lalitha Kurupatham, Mark I-Cheng Chen, Monica Chan, Shawn Vasoo, Lin-Fa Wang, Boon Huan Tan, Raymond Tzer Pin Lin, Vernon Jian Ming Lee, Yee-Sin Leo, David Chien Lye, and for the Singapore 2019 Novel Coronavirus Outbreak Research Team. Epidemiologic features and clinical course of patients infected with SARS-CoV-2 in Singapore. *JAMA*, 323 (15):1488–1494, 2020.

[85] The COVID-19 Investigation Team. Clinical and virologic characteristics of the first 12 patients with coronavirus disease 2019 (COVID-19) in the United States. *Nature Medicine*, 2020.

[86] Enrico Lavezzo, Elisa Franchin, Constanze Ciavarella, Gina Cuomo-Dannenburg, Luisa Barzon, Claudia Del Vecchio, Lucia Rossi, Riccardo Manganelli, Arianna Loregian, Nicolò Navarin, et al. Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. *Nature*, 584(7821):425–429, 2020.

[87] Seungjae Lee, Tark Kim, Eunjung Lee, Cheolgu Lee, Hojung Kim, Heejeong Rhee, Se Yoon Park, Hyo-Ju Son, Shinae Yu, Jung Wan Park, et al. Clinical course and molecular viral shedding among asymptomatic and symptomatic patients with SARS-CoV-2 infection in a community treatment center in the Republic of Korea. *JAMA internal medicine*, 180(11):1447–1452, 2020.

[88] Muge Cevik, Krutika Kuppalli, Jason Kindrachuk, and Malik Peiris. Virology, transmission, and pathogenesis of SARS-CoV-2. *bmj*, 371, 2020.

[89] Liling Chaw, Wee Chian Koh, Sirajul Adli Jamaludin, Lin Naing, Mohammad Fathi Alikhan, and Justin Wong. Analysis of SARS-CoV-2 transmission in different settings, Brunei. *Emerging infectious diseases*, 26(11), 2020.

[90] Fiona P Havers, Carrie Reed, Travis Lim, Joel M Montgomery, John D Klena, Aron J Hall, Alicia M Fry, Deborah L Cannon, Cheng-Feng Chiang, Aridth Gibbons, et al. Seroprevalence of antibodies to SARS-CoV-2 in 10 sites in the United States, March 23-May 12, 2020. *JAMA internal medicine*, 180(12):1576–1586, 2020.

[91] Shiwani Mahajan, Rajesh Srinivasan, Carrie A Redlich, Sara K Huston, Kelly M Anastasio, Lisa Cashman, Dorothy S Massey, Andrew Dugan, Dan Witters, Jenny Marlar, et al. Seroprevalence of SARS-CoV-2-specific IgG antibodies among adults living in Connecticut: Post-infection prevalence (PIP) study. *The American Journal of Medicine*, 2020.

[92] Edward H Kaplan and Howard P Forman. Logistics of aggressive community screening for Coronavirus 2019. *JAMA Health Forum*, 1(5):e200565–e200565, 2020.

[93] Connecticut State Department of Public Health. Nursing Homes with Residents Positive for COVID-19, 2020. URL https://data.ct.gov/Health-and-Human-Services/Nursing-Homes-with-Residents-Positive-for-COVID-19/wyn3-qphu.

[94] Joseph T Wu, Kathy Leung, Mary Bushman, Nishant Kishore, Rene Niehus, Pablo M de Salazar, Benjamin J Cowling, Marc Lipsitch, and Gabriel M Leung. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nature Medicine*, 26(4):506–510, 2020.

[95] Qifang Bi, Yongsheng Wu, Shujiang Mei, Chenfei Ye, Xuan Zou, Zhen Zhang, Xiaojian Liu, Lan Wei, Shaun A Truelove, Tong Zhang, Wei Gao, Cong Cheng, Xiujuan Tang, Xiaoliang Wu, Yu Wu, Binbin Sun, Suli Huang, Yu Sun, Juncen Zhang, Ting Ma, Justin Lessler, and Tiejian Feng. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *The Lancet Infectious Diseases*, 2020.

[96] Elizabeth A Lane, Damien J Barrett, Miriam Casey, Conor G McAloon, Aine B Collins, Kevin Hunt, Andrew W Byrne, David McEvoy, Ann Barber, John M Griffin, Patrick Wall, and Simon J More. Country differences in hospitalisation, length of stay and admission to Intensive Care Units due to SARS-CoV-2 infection: a rapid review of available literature. *medRxiv*, 2020.

[97] Graziano Onder, Giovanni Rezza, and Silvio Brusaferro. Case-fatality rate and characteristics of patients dying in relation to COVID-19 in Italy. *JAMA*, 2020.

[98] Joseph A Lewnard, Vincent X Liu, Michael L Jackson, Mark A Schmidt, Britta L Jewell, Jean P Flores, Chris Jentz, Graham R Northrup, Ayesha Mahmud, Arthur L Reingold, Maya Petersen, Nicholas P Jewell, Scott Young, and Jim Bellows. Incidence, clinical outcomes, and transmission dynamics of hospitalized 2019 coronavirus disease among 9,596,321 individuals residing in California and Washington, United States: a prospective cohort study. *medRxiv*, 2020.

[99] Ishan Paranjpe, Adam Russak, Jessica K De Freitas, Anuradha Lala, Riccardo Miotto, Akhil Vaid, Kipp W Johnson, Matteo Danieletto, Eddye Golden, Dara Meyer, Manbir Singh, Sulaiman Somani, Sayan Manna, Udit Nangia, Arjun Kapoor, Ross O'Hagan, Paul F O'Reilly, Laura M Huckins, Patricia Glowe, Arash Kia, Prem Timsina, Robert M Freeman, Matthew A Levin, Jeffrey Jhang, Adolfo Firpo, Patricia Kovatch, Joseph Finkelstein, Judith A Aberg, Emilia Bagiella, Carol R Horowitz, Barbara Murphy, Zahi A Fayad, Jagat Narula, Eric J Nestler, Valentin Fuster, Carlos Cordon-Cardo, Dennis S Charney, David L Reich, Allan C Just, Erwin P Bottinger, Alexander W Charney, Benjamin S Glicksberg, Girish Nadkarni, and on behalf of the Mount Sinai

Covid Informatics Center (MSCIC). Clinical characteristics of hospitalized COVID-19 patients in New York City. *medRxiv*, 2020.

[100] Annemarie B Docherty, Ewen M Harrison, Christopher A Green, Hayley E Hardwick, Riinu Pius, Lisa Norman, Karl A Holden, Jonathan M Read, Frank Dondelinger, Gail Carson, Laura Merson, James Lee, Daniel Plotkin, Louise Sigfrid, Sophie Halpin, Clare Jackson, Carrol Gamble, Peter W Horby, Jonathan S Nguyen-Van-Tam, Antonia Ho, Clark D Russell, Jake Dunning, Peter JM Openshaw, J Kenneth Baillie, and Malcolm G Semple. Features of 20 133 UK patients in hospital with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ*, 369, 2020.

[101] Daniel Weinberger, Ted Cohen, Forrest Crawford, Farzad Mostashari, Don Olson, Virginia E Pitzer, Nicholas G Reich, Marcus Russi, Lone Simonsen, Anne Watkins, and Cecile Viboud. Estimating the early death toll of COVID-19 in the United States. *Medrxiv*, 2020.

[102] Gideon Meyerowitz-Katz and Lea Merone. A systematic review and meta-analysis of published research data on COVID-19 infection-fatality rates. *medRxiv*, 2020.

[103] Richard Grewelle and Giulio De Leo. Estimating the global infection fatality rate of COVID-19. *medRxiv*, 2020.

[104] Gianluca Rinaldi and Matteo Paradisi. An empirical estimate of the infection fatality rate of COVID-19 from the first Italian outbreak. *medRxiv*, 2020.