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

Andrea L. Bertozziab,1 , Elisa Francohc,1 , George Mohlerd, Martin B. Shorte, and Daniel Sledgef

aDepartment of Mathematics, University of California, Los Angeles, CA 90095; bDepartment of Mechanical and Aerospace Engineering, University of California, Los Angeles, CA 90095; cDepartment of Bioengineering, University of California, Los Angeles, CA 90095; dDepartment of Computer Science, Indiana University–Purdue University Indianapolis, Indianapolis, IN 46202; eDepartment of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332; and fDepartment of Political Science, University of Texas at Arlington, Arlington, TX 76019

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected in 2018.

Contributed by Andrea L. Bertozzi, April 15, 2020 (sent for review April 7, 2020; reviewed by Tapio Schneider, Chad Topaz, and Thomas P. Witelski)

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

The epidemiological perspective on modeling infectious disease spread involves consideration of a larger number of disease. The branching process can also track changes over time in the dynamic reproductive number.

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

Where this takes place, social distancing efforts that appear to have succeeded in the short term will have little impact on the total number of infections expected over the course of the pandemic.

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

Here, we present three basic models of disease transmission that can be fit to data emerging from local and national governments. While the Imperial College study employed an agent-based method (one that simulates individuals getting sick and recovering through contacts with other individuals in the population), we present three macroscopic models: 1) exponential growth, 2) self-exciting branching process, and 3) the susceptible–infected–resistant (SIR) compartment model. These models have been chosen for their simplicity, minimal number of parameters, and for their ability to describe regional-scale aspects of the pandemic. In presenting these models, we demonstrate how they are connected and note that in different cases one model may fit better than another. Because these models are parsimonious, they are particularly well suited to isolating key features of the pandemic and to developing policy-relevant insights. We order them according to their usefulness at different stages of the pandemic—exponential growth for the initial stage, a self-exciting branching process when one is still analyzing individual count data going into the development of the pandemic, and a macroscopic mean-field model going into the peak of the epidemic modeling at the forefront of worldwide public policy making. Nonetheless, modeling and forecasting the spread of COVID-19 remain a challenge. Here, we present and detail these regional-scale models for forecasting and assessing the course of the pandemic. This work is intended to demonstrate the utility of parsimonious models for understanding the pandemic and to provide an accessible framework for generating policy-relevant insights into its course. We show how these models can be connected to each other and to time series data for a particular region. Capable of measuring and forecasting the impacts of social distancing, these models highlight the dangers of relaxing nonpharmaceutical public health interventions in the absence of a vaccine or antiviral therapies.

The authors declare no competing interest.

This open access article is distributed under Creative Commons Attribution License 4.0 (CC BY).


Reviewers: T.S., California Institute of Technology; C.T., Williams College; and T.P.W., Duke University.

Data deposition: The code and data used to generate Fig. 1, Upper can be downloaded from GitHub at https://github.com/francoelisa/PNAS2020. The code to download data and generate Fig. 1, Lower can be downloaded from GitHub at https://github.com/gomohler/pnas2020/tree/master/dynamic_R. The code to generate Table 1 is available in GitHub at https://github.com/gomohler/pnas2020.

1To whom correspondence may be addressed. Email: bertozzi@ucla.edu.

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

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

Results

A. Exponential Growth. Epidemics naturally exhibit exponential behavior in the early stages of an outbreak, when the number of infections is already substantial but recoveries and deaths are still negligible. If at a given time \( t \) there are \( I(t) \) infected individuals and \( \alpha \) is the rate constant at which they infect others, then at early times (neglecting recovered individuals), \( I(t) = I_0 e^{\alpha t} \). The time it takes to double the number of cumulative infections (doubling time) is a common measure of how fast the contagion spreads: if we start with \( I \) infections, then at time \( T_d = \ln 2 / \alpha \) we achieve \( 2I \) infections. For the COVID-19 outbreak, exponential growth is seen in data from multiple countries (Fig. 1), with remarkably similar doubling times in the early stages of the epidemic. For COVID-19, we expect an exponential growth phase during the first 15 to 20 d of the outbreak in the absence of public health interventions such as social distancing, isolation, or quarantine. This estimate is based on patient data from the Wuhan outbreak, which indicate that the average time from illness onset to death or discharge is between 17 and 21 d for hospitalized patients (20, 21). Because deaths are a fraction of infections, they initially increase at a similar exponential pace,
with some delay relative to the beginning of the outbreak. These observed doubling time estimates (2 to 4 d) are significantly smaller than early estimates (∼7 d) obtained using data collected in Wuhan (22).

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

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

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

Given Eq. 1, the quantity

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

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

\[ \mathcal{R}(t) = \sum_{k=1}^{\beta} r_k \{ t_i \in I_k \}. \]

Here, \( I_k \) are intervals discretizing time, and \( \beta \) is the number of such intervals. The reproduction number, \( r_k \), in each interval \( k \) is determined by

\[ r_k = \sum_{t_i > t_j} p_{ij} \{ t_j \in I_k \} / N_k, \]

where \( N_k \) is the total number of events in interval \( k \).

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

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

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

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

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

The SIR model equations are

\[ \frac{dS}{dt} = -\beta IS / N, \quad \frac{dI}{dt} = \beta IS / N - \gamma I, \quad \frac{dR}{dt} = \gamma I. \]

Here, \( \beta \) is the transmission rate constant, \( \gamma \) is the recovery rate constant, and \( R_0 = \beta / \gamma \) is the reproduction number. One integrates Eq. 6 forward in time from initial values of \( S, I, \) and \( R \) at time 0. The SEIR model includes an exposed category \( E \):

\[ \frac{dS}{dt} = -\beta IS / N, \quad \frac{dE}{dt} = \beta IS / N - aE, \quad \frac{dI}{dt} = aE - \gamma I, \quad \frac{dR}{dt} = \gamma I. \]

Here, \( a \) is the inverse of the average incubation time. Both models are fit, using maximum likelihood estimation with a Poisson likelihood, to data for three US states (California, New York, and Indiana) (17). Table 1 compares the results with the branching process. We use the relative likelihood based upon the Akaike Information Criterion (AIC) (39) to measure model performance for each dataset; AIC is biased against models with more parameters. The SEIR model performs better on the confirmed data for California and Indiana, possibly due to the larger amount of data, compared with mortality for which SIR is the best for all three states. The branching process performs best for confirmed cases in New York. Our choice of fitting follows the method in ref. 43 for the 1918 pandemic death data. Our focus is on model comparison rather
than measuring uncertainty of parameters in a specific model, as is currently being done for hospital demand forecasts in Los Angeles (16). For interval forecasts with uncertainty quantification, one may consider a negative binomial alternative to Poisson regression that captures overdispersion in case and death counts (44, 45).

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

$$
\frac{ds}{d\tau} = -R_0i, \quad \frac{di}{d\tau} = R_0is - i, \quad \frac{dr}{d\tau} = i,
$$

$$
(s, i, r) |_{\tau = 0} = (1 - \epsilon, \epsilon, 0),
$$

where $0 < \epsilon < 1$ is the initial fraction of the infected population at the start time, and the system retains only one dimensionless parameter $R_0$ that, in conjunction with the initial conditions, completely determines the resulting behavior. For Eq. 7, the shapes of the solution curves $s(\tau), i(\tau), r(\tau)$ do not depend on $\epsilon$, other than exhibiting a time shift that depends logarithmically on $\epsilon$ (Fig. 2). This is a universal solution for the SIR model in the limit of small $\epsilon$ (Fig. 3), depending only on $R_0$. Critically, the height of the peak in $i(\tau)$ and the total number of resistant/susceptible people by the end of the epidemic are determined by $R_0$ alone. However, the sensitivity of the time translation to the parameter $\epsilon$ and the dependence of true time values of the peak on parameter $\gamma$ make SIR challenging to fit to data at the early stages of an epidemic when Poisson statistics and missing information are prevalent. When using early-time death data to fit SIR, the estimate of the percentage of deaths per total number of infections (chosen as 1% here) has a sensitivity that can be understood directly in terms of this shift in the time to peak. This is important information for public health officials, policy makers, and for political leaders interested in decreasing $R_0$ for substantial periods of time. This sensitivity to parameters helps explain why projections of the outbreak can display such large variability and highlights the need for extensive disease testing within the population to more accurately track the epidemic curve. After the surge in infections, the model asymptotes to an end state in which $r$ approaches the end value $r_{\infty}$ and $s$ approaches $1 - r_{\infty}$, where $r_{\infty}$ and the infected population approaches 0. The value $r_{\infty}$ satisfies a well-known transcendental equation (46–48). A phase diagram of the universal solutions for several $R_0$ values is shown in Fig. 2, Upper Right. The dynamics start in the bottom right corner where $s$ is almost 1 and follow the colored line to terminate on the $i = 0$ axis at the value $s_{\infty}$. A rigorous derivation of the limiting state under the assumptions here can be found in refs. 46–48.

Under the SIR (and similar) model(s), if $R_0$ is decreased during the middle of an outbreak, through social distancing and other public health measures, the rate of new infections will decrease. However, unless the number of infected individuals is brought down to zero, the outbreak will likely reemerge, and the total number of infections may still be a large fraction of the population. This is illustrated in Fig. 3, where we present scenarios of no social distancing vs. short-term social distancing with parameters from Table 1 for death data up through the end of March, fit to the SIR model. We caution that the goal of these scenarios is not to produce highly accurate percentages but rather, to present scenarios under different basic assumptions that illustrate the usefulness of social distancing measures and the potential danger in easing them too soon.

For the New York state scenario, with $R_0$ presumed to decrease by a factor of two with distancing measures, the outbreak is not completely controlled by distancing, and the number of infections continues to rise to approximately 10% by mid-April. In contrast, without distancing measures the scenario shows four times that number of infected individuals by mid-April, a level that would have represented a potentially disastrous...
strain on the hospital system. In the California scenario, distancing measures bring the effective $R_0$ closer to one, thus keeping infections at a much smaller portion of the population than the scenario with no social distancing, at least during the period when the distancing is still in effect. We take these scenarios one step further by suddenly stopping distancing on 5 May (this is both extreme and hypothetical, but it serves to illustrate the model). Because of the low initial infected count in California, bringing $R_0$ back to the original predistancing level produces a curve that follows the original peak scenario, just shifted later in time. For New York state, because a nontrivial fraction of the population is initially infected, there are fewer to infect, and the new peak is less steep than the scenario without any distancing.

**Discussion**

Our analysis, employing parsimonious models, illustrates several key points. 1) The reproduction number $R$ is highly variable both over time and by location, and this variability is compounded by distancing measures. These variations can be calculated using a stochastic model, and lower $R$ is critical to decreasing strains on health care systems and to creating time to develop effective vaccines and antiviral therapies. 2) Morbidity data and confirmed case data have statistics that vary by location and by time depending on testing and on accurate accounting of deaths due to the disease. Differences in collection methods and in the accuracy of morbidity and mortality data can lead to different projected outcomes. 3) Nonpharmaceutical public health interventions (NPIs) such as social distancing and shelter in place orders offer an important means of reducing the virus’s reproduction number. Nonetheless, NPIs may not have a substantial impact on the total number of infections unless sustained over time. Policy makers should be cautious about scaling back distancing measures after early signs of effectiveness.

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

![Fig. 2. Solution of the dimensionless SIR model (5) with $R_0 = 2$. Upper Left shows the graphs of $s$ (blue), $i$ (orange), and $r$ (gray) on the vertical axis vs. $\tau$ on the horizontal axis, for different $\epsilon$. The corresponding values of $\epsilon$ from left to right are $10^{-4}$, $10^{-6}$, $10^{-8}$, $10^{-10}$, respectively. Upper Center shows the time until peak infections vs. $\log(\epsilon)$ for the values shown in Upper Left. This asymptotic tail to the left makes it challenging to fit data to SIR in the early stages. Upper Right is a phase diagram for fraction of infected vs. fraction of susceptible with the direction of increasing $\epsilon$ indicated by arrows, for three different values of $R_0$. Lower displays a typical set of SIR solution curves over the course of an epidemic, with important quantities labeled.](image1)

![Fig. 3. Scenarios for the impact of short-term social distancing: fraction of population vs. date. (Left) California SIR model based on mortality data with parameters from Table 1 ($R_0 = 2.7, \gamma = .12, I_0 = .1$) under two scenarios: $R_0$ constant in time (light blue) and $R_0$ cut in half from 27 March (1 wk from the start of the California shutdown) to 5 May but then returned to its original value, to represent a short-term distancing strategy (dark blue). (Right) New York SIR model with parameters from Table 1 ($R_0 = 4.1, \gamma = .1, I_0 = .05$) under the same two scenarios but with short-term distancing occurring over the dates of 30 March (1 wk from the start of the New York shutdown) to 5 May. In both states, the distancing measures suppress the curve and push the peak infected date into the future, but the total number of cases is only slightly reduced.](image2)
Integrating this, we see that

g with a shorter doubling time when

total cumulative infections are

infections, which include recoveries and deaths. Using the SIR model, the

\(\alpha\) with \(\alpha\) 6

in Eq. 38 can vary signifi-
cantly for a given location. Although we have in each case
determined which of these fits appears to have most validity,
in many cases these are not strong indicators. These models
have several sources of uncertainty, including parameter uncer-
tainty, variation based on data or model type used, and most
importantly, uncertainty in the severity and length of social dis-
tancing measures, which can change the peak date by months
or even create multiple peaks. This variability in outcomes high-
lights the challenges of modeling and forecasting the course of
a pandemic during its early stages and with only limited data.
This uncertainty is a major challenge for policy makers, who
must consider the social and economic consequences of disrup-
tive public health interventions while recognizing that relaxing
them may swiftly lead to the reemergence of a devastating
disease.

**Materials and Methods**

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

7. J. Riou, C. L. Althaus, Pattern of early human-to-human transmission of Wuhan 2019-

8. T. A. Perkins et al., Estimating unobserved SARS-Cov-2 infections in the United States. The
https://doi.org/10.1101/2020.03.15.20036582 (26 March 2020).
9. A. J. Kucharski et al., Early dynamics of transmission and control of COVID-19: A

31. J. Wallinga, P. Teunis, Different epidemic curves for severe acute respiratory syndrome
30. S. Cauchemez
18. G. Mohler, F. Schoenberg, M. B. Short, D. Sledge, Analyzing the world-wide
et al.
17. E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in
et al.
16. J. Wallinga, P. Teunis, Different epidemic curves for severe acute respiratory syndrome
et al.
14. N. Alteredi et al., Curating a COVID-19 data repository and forecasting county-level
12. J. T. Wu, K. Leung, P. G. M. Leung, Nowcasting and forecasting the potential domestic
and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A
11. J. Hellewell et al., Nowcasting and forecasting the potential domestic and
international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study
10. L. C. Tindale et al., Transmission interval estimates suggest pre-symptomatic spread of
9. J. K. Muller et al., Estimating the reproductive number of COVID-19 from daily death
counts in New York City and Milwaukee during the first 100 days of the pandemic.