<u>Original Paper</u>

Adaptive Susceptible-Infectious-Removed Model for Continuous Estimation of the COVID-19 Infection Rate and Reproduction Number in the United States: Modeling Study

Mark B Shapiro, PhD; Fazle Karim, PhD; Guido Muscioni, MSc; Abel Saju Augustine, MSc

Anthem, Inc, Indianapolis, IN, United States

Corresponding Author:
Mark B Shapiro, PhD
Anthem, Inc
220 Virginia Avenue
Indianapolis, IN, 46204
United States
Phone: 1 708 295 8150
Email: mark.shapiro@anthem.com

Abstract

Background: The dynamics of the COVID-19 pandemic vary owing to local population density and policy measures. During decision-making, policymakers consider an estimate of the effective reproduction number R_t, which is the expected number of secondary infections spread by a single infected individual.

Objective: We propose a simple method for estimating the time-varying infection rate and the R_t.

Methods: We used a sliding window approach with a Susceptible-Infectious-Removed (SIR) model. We estimated the infection rate from the reported cases over a 7-day window to obtain a continuous estimation of R_t. A proposed adaptive SIR (aSIR) model was applied to analyze the data at the state and county levels.

Results: The aSIR model showed an excellent fit for the number of reported COVID-19 cases, and the 1-day forecast mean absolute prediction error was <2.6% across all states. However, the 7-day forecast mean absolute prediction error approached 16.2% and strongly overestimated the number of cases when the R_t was rapidly decreasing. The maximal R_t displayed a wide range of 2.0 to 4.5 across all states, with the highest values for New York (4.4) and Michigan (4.5). We found that the aSIR model can rapidly adapt to an increase in the number of tests and an associated increase in the reported cases of infection. Our results also suggest that intensive testing may be an effective method of reducing R_t.

Conclusions: The aSIR model provides a simple and accurate computational tool for continuous R_t estimation and evaluation of the efficacy of mitigation measures.

Introduction

The COVID-19 pandemic is currently underway. As of September 2, 2020, over 6,000,000 individuals in the United States have been reported positive for COVID-19. Modeling studies are key to understanding the factors that drive the spread of the disease and for developing mitigation strategies. Early modeling efforts forecasted very large numbers of infected individuals, which would overwhelm health care systems in many countries [1-3]. These forecasts served as a call to action

for policymakers to introduce policy measures including social distancing, travel restrictions, and eventually lockdowns to avoid the predicted catastrophe [4-6]. The mitigating policy measures have been successful in changing the dynamics of the pandemic and in "flattening the curve," such that fewer people have needed to seek treatment at any given time, and this has prevented the health care system from getting overwhelmed.

One of the most fundamental metrics that describes the pandemic's dynamics is the reproduction number R_t, which is

XSL•FO
RenderX

the expected number of secondary infections spread by a single infectious individual [7]. In 1906, Hamer [8] speculated that the course of an epidemic is determined by the rate of contact between susceptible and infectious individuals. Later, Kermack and McKendrick [9] reported that epidemics end not when there are no susceptible individuals left, but rather when each infectious individual can infect, on average, <1 more individual. The R_t depends on three factors: (1) the likelihood of infection per contact, (2) the period during which infectious individuals freely interact with susceptible individuals and spread the disease, and (3) the rate of contact. The likelihood of infection per contact (factor 1) is determined on the basis of pathogen virulence and protective measures such as social distancing or wearing masks. Free interactions between infectious and susceptible individuals (factor 2) occur until the infectious individual is self-quarantined or hospitalized, either when the individual tests positive or experiences severe symptoms. Finally, the rate of contact (factor 3) is strongly affected by public health measures to mitigate risk [10], such as lockdowns during the COVID-19 pandemic. Thus, R_t is determined on the basis of the biological properties of the pathogen and multiple aspects of social behavior. When $R_t > 1$, the number of cases is expected to increase exponentially. The pandemic is considered to have been contained when R_t decreases and remains at <1. Real-time R_t estimation is critical for determining the effect of implemented mitigation measures and future planning.

We propose a method for continuous estimation of the infection rate and R_t to investigate the effect of mitigation measures and immunity acquired by those who recover from the disease. We estimated R_t with a Susceptible-Infectious-Removed (SIR) model [9] that describes the dynamics of population compartments as follows: individuals are initially "susceptible," contract the viral infection and become "infectious," and are then moved to the "removed" compartment once they are quarantined or hospitalized, recover, or die. The SIR model is one of the simplest epidemiological models that still captures the main properties of an epidemic [11,12], and it has been widely used in epidemic modeling studies. In most SIR modeling studies, the model parameters were constant. An SIR model with constant parameters, however, cannot be applied for the COVID-19 pandemic because various mitigating measures were introduced during pandemic progression. The effect of policy changes on COVID-19 dynamics has been modeled using the combination of an SIR model and Bayesian inference [13,14]. In these modeling studies, the rate of infection spread was assumed to be piece-wise linear among the 3 dates of the implementation of policy changes. In another approach, continuous estimation of R_t and an assessment of the effect of mitigation measures were carried out on the basis of estimates of the distribution of the serial intervals between symptom onset in the primary and secondary cases [15-17]. Bayesian inference and methods based on estimations of the serial interval include multiple parameters whose values are not estimated from the data. In contrast, we propose an adaptive SIR (aSIR) model in which only one parameter—the removal rate—is determined from the literature, while the second parameter—the infection rate—is continuously estimated from the data through a sliding

window approach. A continuous R_t estimate is then obtained using the infection rate estimate. The SIR model is described as a system of differential equations, and the key idea in our proposed method is that the initial conditions for each window are considered as values estimated for the previous window. The only additional hyperparameter is the length of the sliding window. The proposed method retains the conceptual and computational simplicity of SIR-type models and can be easily extended through the introduction of additional compartments supported by data.

Methods

Data

Data on daily and cumulative confirmed cases between February 29 and September 2, 2020, were obtained from John Hopkins University (JHU), and the dates of interventions by state (eg, state of emergency and stay-at-home orders) were obtained from Wikipedia. The JHU data were available at 2 levels of aggregation: county and state. JHU considers many sources for reporting these data; county-level information was extracted from the websites of the states' departments of health, and state-level data were extracted directly from the website of the Centers for Disease Control and Prevention.

Model

The SIR model is a system of ordinary differential equations:

$$\frac{dS}{dt} = -\beta \frac{I}{N} S$$
$$\frac{dI}{dt} = \beta \frac{I}{N} S - \gamma I \quad (1)$$
$$\frac{dR}{dt} = \gamma I$$

Here, S is the number of susceptible individuals, I is the number of infectious individuals who freely interact with others and can transmit the infection, R is the number of individuals excluded from the other 2 compartments because they are quarantined or hospitalized, have recovered and acquired immunity, or have died. Several sources of government data on COVID-19 provide the daily number of newly confirmed cases and a cumulative number of confirmed cases. Careful consideration is required to determine whether these numbers should be attributed to the I or R compartment. In the United States, once an individual has been confirmed positive for COVID-19, he/she is expected to be either self-isolated or hospitalized. Therefore, we assigned the data on confirmed cases to the R compartment, and we fit the model to the cumulative number of confirmed cases.

The infection rate is determined as follows:

$$\beta = p \times c \quad (2)$$

where p is the probability of being infected upon contact with an infectious individual, and c is the average number of contacts per day. We have no data that would allow us to estimate p and c separately; hence, we directly estimated β, as is usually performed when using SIR models.

XSL•FO

RenderX

The removal rate γ determines the rate at which infected individuals are moved from the I to the R compartments. In the context of the COVID-19 pandemic, γ is determined from the time taken for the appearance of severe symptoms, such that the individual can be tested and is self-quarantined or hospitalized, as required. Therefore, we assumed the duration of the infectious period as the average time taken for the infected individual to be isolated, not the overall time for recovery. We assumed that an individual is infectious from the day he/she contracts the infection before symptom onset [18-20]. The average time to symptom onset is 5-6 days [21-23]. We assumed that the infectious period before the development of severe symptoms is 6 days; hence, $\gamma=1/6$.

Time-Variant Parameter Estimation

The aSIR model contains two parameters, β and γ, with $\gamma=1/6$ obtained from the literature, and β estimated from the reported data for each region of interest. The time-variant $\beta(t)$ was estimated using a sliding window of $\tau=7$ days and step of $s=1$ day, with the estimated values for S and I obtained from the previous window used as the initial conditions for the next window.

The reproduction number was calculated as follows:

$$R_t(t) = \beta(t)/\gamma \qquad (3)$$

For the first window, we determined the date when the number of confirmed cases began to increase exponentially. This is important because for many states or counties, very few confirmed cases were initially reported for a number of days or even weeks, which suggests that either the epidemic had not started or the true number of infected individuals was unknown. It is not reasonable to apply an SIR model for this initial period. We considered the onset of the pandemic as the first of the 4 consecutive days in which the number of reported confirmed

cases increased in at least 3 days. The initial conditions for system (1) for window 0 were as follows:

$$S_0(0) = N \qquad (4)$$

where N is the population in the region of interest, $I_0(0)=1$, and $R_0(0)=0$. The infection rate β_i and $S(t)$, $I(t)$ for t $[0, \tau-1]$ were estimated from the initial conditions and actual R.

The window was slid by $s=1$ point. For the new $i+1$ window, the initial conditions were considered as the estimated values from the previous window $S_{i+1}(0)=S_i(s)$, $I_{i+1}(0)=I_i(s)$, and actual $R_{i+1}(0)=R(s)$. The actual values of $R(t)$ were used, and the infection rates β_{i+1} and $S_{i+1}(t)$, and $I_{i+1}(t)$ were estimated.

For each window, the $R_{t,i}$ was determined as follows:

$$R_{t,i} = \beta_i/\gamma \qquad (5)$$

The $R_{t,i}$ was assigned to the last time point of the window. To obtain a smooth estimate of R_t, we used a rolling average of 5 points.

Results

We fit the model for each state and county in the United States. Model performance was evaluated by calculating the quality of fit as the root mean squared error between the actual and fitted R data for all windows concatenated (wRMSE). The fit was excellent with wRMSE<6 across all states. Furthermore, we calculated 1-day, 3-day, and 7-day forecasts of R after each window (Figure 1A). The mean absolute prediction error for the forecasts is provided in Table 1. The 1-day forecast error did not exceed 2.6% across all states, while the 7-day forecast error was large and approached 16.2% for New York. In particular, the 7-day forecast strongly overestimated the number of cases when R_t was rapidly decreasing (Figure 1).

Figure 1. (A) Estimated Infectious and forecast Removed. (B) Estimated reproduction number R_t. The shaded region indicates the dates of the lockdown. While the 1-day and 3-day forecasts are accurate, the 7-day forecast exhibits marked errors when $R_t > 1$ and is rapidly decreasing.
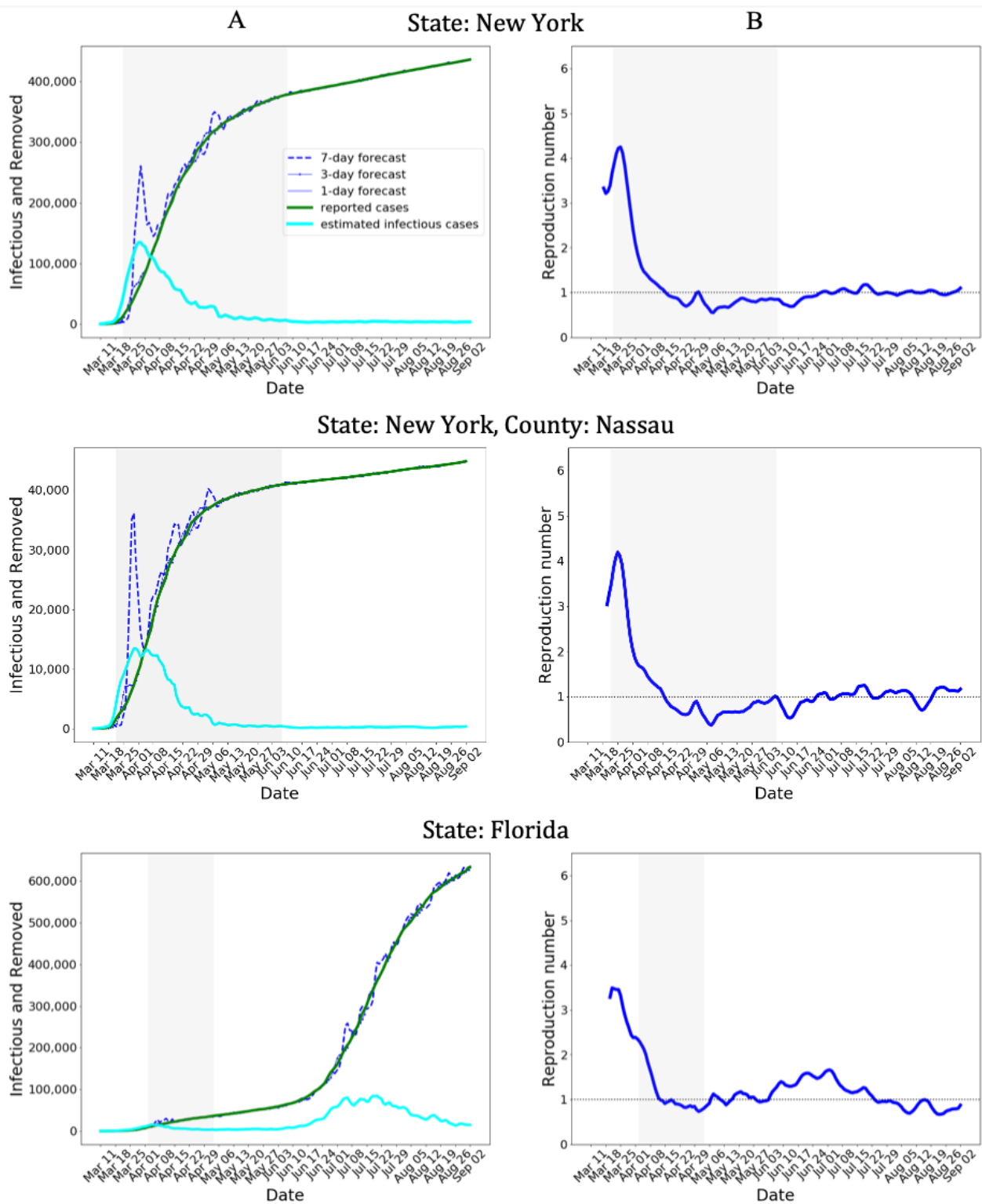
Table 1. Reproduction numbers and forecast accuracy for 50 US states.

State	R_t^a max	MAPE[b] (1-day forecast), %	MAPE (3-day forecast), %	MAPE (7-day forecast), %
Alabama	2.9	1.5	4.2	10.0
Alaska	2.8	1.6	3.7	11.3
Arizona	3.3	1.3	2.9	10.1
Arkansas	2.8	1.5	4.0	12.6
California	2.5	1.7	2.9	6.3
Colorado	2.6	1.1	2.9	7.3
Connecticut	4.1	2.0	3.1	9.3
Delaware	2.4	1.7	2.9	7.9
District of Columbia	2.1	0.8	1.8	4.4
Florida	3.6	2.0	4.4	9.3
Georgia	3.0	1.8	3.7	7.4
Hawaii	2.7	2.0	3.6	9.7
Idaho	3.4	2.4	4.8	13.6
Illinois	4.0	1.3	2.5	8.5
Indiana	3.8	1.4	4.0	10.4
Iowa	2.8	1.8	3.6	8.0
Kansas	3.0	1.6	3.5	8.6
Kentucky	3.0	2.6	4.9	11.2
Louisiana	3.7	1.8	4.0	12.1
Maine	2.0	1.2	2.8	6.7
Maryland	3.3	1.2	2.8	6.2
Massachusetts	3.4	1.3	3.6	9.7
Michigan	4.5	1.6	3.5	12.8
Minnesota	2.7	1.3	2.9	8.0
Mississippi	2.9	1.2	3.0	9.3
Missouri	3.6	1.8	3.3	11.4
Montana	3.3	1.6	3.7	11.9
Nebraska	2.5	2.0	4.1	9.7
Nevada	2.9	2.3	3.8	10.0
New Hampshire	2.3	1.7	3.3	8.5
New Jersey	4.1	1.5	2.3	7.8
New Mexico	2.3	2.2	3.4	7.3
New York	4.4	1.5	4.2	16.2
North Carolina	3.2	1.3	2.4	7.2
North Dakota	2.4	1.8	4.8	12.6
Ohio	3.3	1.2	3.4	9.8
Oklahoma	3.1	1.4	3.5	10.2
Oregon	2.5	1.3	2.8	6.6
Pennsylvania	3.2	1.7	2.9	6.3
Rhode Island	2.4	1.5	3.1	6.8
South Carolina	3.5	2.1	4.3	10.6

State	R_t^a max	MAPE^b (1-day forecast), %	MAPE (3-day forecast), %	MAPE (7-day forecast), %
South Dakota	2.1	1.3	3.2	8.7
Tennessee	3.5	2.2	4.8	12.5
Texas	3.6	2.0	3.9	9.3
Utah	3.2	1.4	3.1	8.3
Vermont	2.9	0.8	2.4	7.7
Virginia	2.5	1.1	2.1	5.1
Washington	3.0	2.0	4.8	8.6
West Virginia	3.5	1.6	3.9	14.0
Wisconsin	3.6	1.5	3.2	10.0
Wyoming	2.9	1.9	4.7	14.2

^{a}R_t: reproduction number.

^{b}MAPE: mean absolute prediction error.

The estimated time course of R_t for New York and Nassau county, one of the most affected counties since the beginning of the COVID-19 pandemic, are shown in Figure 1. The estimated daily number of infectious individuals rapidly increased and then gradually declined after the lockdown was implemented on March 22, 2020 (Figure 1A). The estimated R_t also declined upon implementation of the lockdown (Figure 1B). The time course of R_t exhibits weekly seasonality, which likely reflects the effect of social interactions and possibly the effect of fluctuations in case reporting on weekdays vs weekends. For New York and Nassau county, R_t initially increased, which may reflect the fact that the pandemic in New York was continuously seeded by travelers arriving at John F Kennedy International Airport until a ban on international travel was implemented on March 12, 2020. This may also reflect the fact that not all severe cases were initially recognized and reported as COVID-19 cases. In Florida, R_t decreased to almost 1 by mid-April but then began increasing at the end of May (Figure 1B). In June 2020, Florida authorities introduced more stringent measures to control the pandemic, which is reflected in the reduction in R_t in the second half of July 2020. The opening of multiple states since June 2020 has been accompanied by an increase in R_t beyond 1 (data not shown), and close monitoring of R_t is needed to contain another wave of the pandemic.

Next, we compared aSIR with the model developed by Cori et al [15], implemented as R package EpiEstim, and a model implemented by Systrom, Vladeck, and Krieger in rt.live [24] (Figure 2). In EpiEstim, we assumed an equal probability of infection within the infectious period of 6 days, the R_t estimate was smoothed with a 7-point rolling average window, same as that in aSIR. While all 3 models show similar estimates when R_t approaches 1, their estimates differ considerably in the beginning of the pandemic. In particular, the rt.live model [24] returned a lower maximum R_t than the other 2 models and estimated that R_t already decreased to 1 by the time the lockdown was announced in New York on March 22, 2020 (Figure 2, shaded region). The EpiEstim and aSIR models estimated similar peak values of R_t, and both models estimated that R_t decreased and approached 1 in the first week of April 2020. Although both models show a rapid reduction in R_t in March, the aSIR model shows a lagged change. However, we are not aware of the ground truth data to determine which model yields a more accurate estimate.

Figure 2. Comparison of models that generate continuous R_t estimates. The three R_t estimates differ widely in the beginning of the COVID-19 pandemic. In particular, the R_t estimated using the rt.live model of Systrom, Vladeck, and Krieger [24] decreased to 1 at the onset at the lockdown on March 22, 2020 (shaded region). aSIR: adaptive Susceptible-Infectious-Removed.
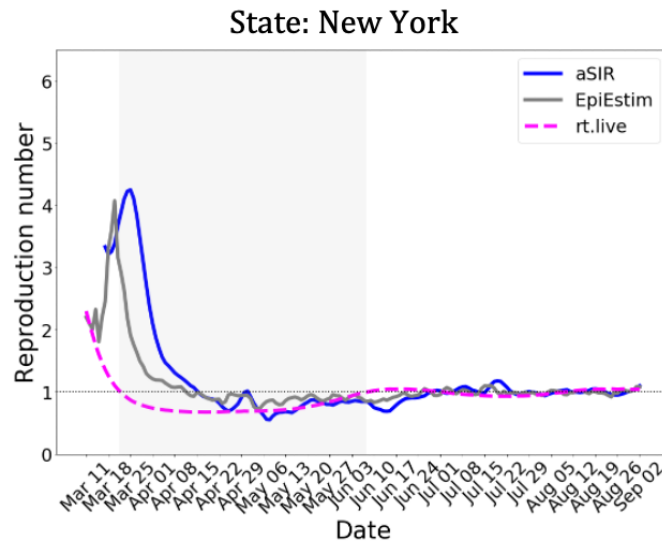
Finally, we investigated the effect of an abrupt increase in testing on the estimated R_t (Figure 3). We assumed a step-wise 50% increase in testing, which persisted after April 12, 2020 (Figure 3, left panel). Both aSIR and EpiEstim models exhibited a spike in R_t. However, an increase in testing would help identify and quarantine infectious individuals sooner, resulting in a shorter infectious period and larger removal rate γ, in turn decreasing R_t. We did not model a potential increase in γ. Instead, we assumed that the underlying dynamics of the pandemic did not change, and within 2 weeks both models returned to the R_t time course estimated without an increase in testing.

Figure 3. Effect of a step-wise 50% increase in testing (left panel, dashed line). The 1-day forecast by the aSIR model adapts within a week. For the Rt estimate, both EpiEstim and our aSIR models produced a spike, followed by a reduction (right panel, dashed lines) before returning to the unperturbed Rt time course (solid lines). aSIR: adaptive Susceptible-Infectious-Removed.
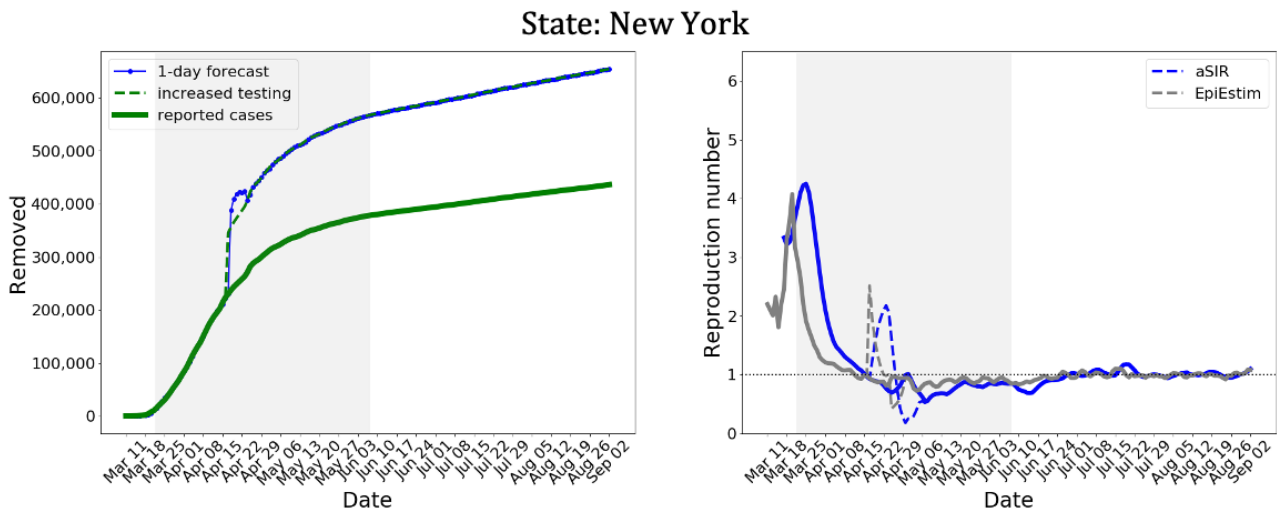
Discussion

Principal Findings

We developed a simple approach to adaptively estimate the time-varying parameters of the SIR model, using reported data on the number of confirmed COVID-19 cases. This approach adds to the already large literature on COVID-19 modeling in 2 ways. First, we estimate the parameters of the SIR model with a sliding window of a limited duration (7 days) to account for rapid changes in transmissibility and contact patterns in response to changes in social behavior and government mitigation measures. The window duration is a hyperparameter that can be changed as needed, the trade-off being the accuracy of the parameter estimates versus the rapid reaction to changes in the underlying pandemic. Because the proposed model is so simple, a number of scenarios can be explored as needed.

Second, we attribute the data on reported cases to the Removed compartment rather than the Infectious compartment. This modeling decision is based on the realities of the COVID-19 pandemic in the United States, where individuals with confirmed COVID-19 are supposed to self-isolate or be hospitalized. Although these individuals remain infectious and can infect

other family members or caregivers even when self-isolated or hospitalized, they would not freely interact with the susceptible population, as would be required to attribute them to the *I* compartment. The addition of a new X compartment in the SIR model has been proposed to model symptomatic quarantined infectious individuals [25]. However, we have no data to independently estimate this additional parameter of quarantine rate. For the same reason, we did not use the Susceptible-Exposed-Infected-Removed (SEIR) model because we are not aware of reliable data on the duration of the exposure period during which an infected person is not yet infectious. Moreover, it has been reported that the SIR model performed better than the SEIR model in representing the information contained in the confirmed-case data on COVID-19 [26].

The reported number of positive COVID-19 cases represents a fraction of infected individuals because of the limited testing capacity in March and April 2020; consequently, only those who developed severe symptoms were tested. Up to 80% of infected individuals may have been asymptomatic or may have experienced mild symptoms [27] and were not tested; hence, for that period, our model applies only to the small subpopulation with severe symptoms. However, this subpopulation is of particular interest because it represents those who are at the greatest risk, and R_t estimated from these limited data can be used to guide policy decisions aimed at protecting the most vulnerable population [28]. As the number of the tested individuals increases, the short sliding window approach makes our model adaptable to an increasing proportion of the population (Figure 3).

Across all US states, the maximal R_t values were estimated for New York (4.4) and Michigan (4.5) (Table 1), which is similar to the mean value of 4.34 estimated for Italy [29] but higher than that obtained with a stochastic transmission model [30,31]. The wide range of maximal values of R_t of 2.0-4.5 (Table 1) likely reflects the differences in contact rates owing to the population density [32,33]. Increased social distancing is required to contain the spread of the pandemic [34,35], with more stringent mitigation measures, including lockdown, considered necessary to decrease the contact rate in high-density states and counties. Another measure to decrease R_t is to increase the removal rate γ through intensive testing and quarantining of individuals who test positive. This targeted intervention would strongly decrease the interaction between infectious and susceptible individuals and maintain an R_t of <1 until a vaccine is available and while vaccination efforts are ramping up. Intensive testing combined with social distancing and mask wearing, followed by the isolation of individuals confirmed with COVID-19, are key features of reopening strategies for schools and universities [36-38]. Our model allows researchers and policymakers to monitor R_t in different geographic regions of the United States, better understand the effect of government policies on the dynamics of the pandemic,

and develop further mitigation strategies as we continue to battle COVID-19 [39,40].

Limitations

The SIR model is perhaps the simplest model that captures the dynamics of a pandemic. It is based on several assumptions that are valid only to some degree as we consider real-life scenarios. The 2 main limitations of the original SIR model are that it has constant parameters and it is deterministic. Our proposed aSIR model allows us to estimate time-varying parameters and thus removes the first limitation. The other limitation remains, however. It is assumed that infectious individuals freely interact with the susceptible population. The infection rate β encompasses both the probability of transmission and the average number of contacts per day. The SIR model does not reflect interaction dynamics that are stochastic in nature and are described by stochastic epidemiologic models [15-17,41]. In its simplest form, the SIR model does not reflect the heterogeneity of viral transmission reflected in overdispersion or superspreading where few numbers of infected individuals infect a large number of susceptible individuals [42-44]. The removal rate γ is an average number of days until an infectious individual is excluded and does not reflect the variability of this interval, nor does it allow one to model a possibility of a subsequent, albeit reduced transmission to caregivers or other susceptible individuals as may happen in a real-life scenario. Consequently, R_t, which is calculated using constants β and γ rather than their distributions, does not reflect the stochastic nature of the dynamics of the pandemic. Parameter distributions can be obtained by applying Bayesian methods to SIR modeling [45]. Moreover, a single value of R_t estimated for a large population does not reflect differences in subpopulations, such as age groups, which is especially relevant for COVID-19 [46-48]. The generation of a bank of aSIR models for each subpopulation or region can provide a more realistic insight into the dynamics of the pandemic across a larger population [49,50]. Another critical assumption is that once an individual is infected and recovers, he/she is no longer susceptible to repeated infection; however, this assumption does not appear strongly violated. Although some cases of repeated infection with SARS-CoV-2 have been reported [51-53], the risk of re-infection is considered low [54,55]. Overall, the SIR model is a trade-off between the computational simplicity and veracity of describing the real-life complexities of a pandemic.

Conclusions

SIR type models, particularly the proposed time-variant aSIR model, have an advantage over more complex models in the initial stages of a pandemic when critical public policy decisions need to be made while the empirical data on interaction dynamics, transmission rates, and the disease progression and contagiousness from the moment of infection are not yet readily available. Our model provides a simple and efficient method to assess the efficacy of interventions as the pandemic progresses.

Conflicts of Interest

None declared.

References

1. Arenas A, Cota W, Gomez-Gardenes J, Gómez S, Granell C, Matamalas J, et al. A mathematical model for the spatiotemporal epidemic spreading of COVID19. medRxiv. 2020 Mar 23 Preprint posted online March 23, 2020. [doi: 10.1101/2020.03.21.20040022]

2. Gandhi KRR, Murthy KVR, Rao SSP, Casella F. Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality. SSRN. 2020 Mar 25 Preprint posted online March 25, 2020. [doi: 10.2139/ssrn.3560688]

3. Mitjà O, Arenas À, Rodó X, Tobias A, Brew J, Benlloch JM, 62 signatories. Experts' request to the Spanish Government: move Spain towards complete lockdown. Lancet 2020 Apr 11;395(10231):1193-1194 [FREE Full text] [doi: 10.1016/S0140-6736(20)30753-4] [Medline: 32224297]

4. Adam D. Special report: The simulations driving the world's response to COVID-19. Nature 2020 Apr;580(7803):316-318. [doi: 10.1038/d41586-020-01003-6] [Medline: 32242115]

5. Anderson RM, Heesterbeek H, Klinkenberg D, Hollingsworth TD. How will country-based mitigation measures influence the course of the COVID-19 epidemic? Lancet 2020 Mar 21;395(10228):931-934 [FREE Full text] [doi: 10.1016/S0140-6736(20)30567-5] [Medline: 32164834]

6. Enserink M, Kupferschmidt K. With COVID-19, modeling takes on life and death importance. Science 2020 Mar 27;367(6485):1414-1415. [doi: 10.1126/science.367.6485.1414-b] [Medline: 32217707]

7. Delamater PL, Street EJ, Leslie TF, Yang YT, Jacobsen KH. Complexity of the Basic Reproduction Number (R). Emerg Infect Dis 2019 Jan;25(1):1-4 [FREE Full text] [doi: 10.3201/eid2501.171901] [Medline: 30560777]

8. Hamer WH. The Milroy Lectures ON EPIDEMIC DISEASE IN ENGLAND—THE EVIDENCE OF VARIABILITY AND OF PERSISTENCY OF TYPE. The Lancet 1906 Mar;167(4305):569-574. [doi: 10.1016/S0140-6736(01)80187-2]

9. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proc. R. Soc. Lond. A 1997 Jan;115(772):700-721. [doi: 10.1098/rspa.1927.0118]

10. Non-pharmaceutical public health measures for mitigating the risk and impact of epidemic and pandemic influenza. World Health Organization. 2019 Oct. URL: http://www.who.int/influenza/publications/public_health_measures/publication/en/ [accessed 2020-07-07]

11. Anderson RM. Discussion: the Kermack-McKendrick epidemic threshold theorem. Bull Math Biol 1991;53(1-2):3-32. [doi: 10.1016/s0092-8240(05)80039-4] [Medline: 2059740]

12. Anderson R, May R. Infectious diseases of humans: dynamics and control. Oxford, UK: Oxford University Press; 1992.

13. Dehning J, Zierenberg J, Spitzner F, Wibral M, Neto J, Wilczek M, et al. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. medRxiv. 2020 May 06 Preprint posted online May 6, 2020. [doi: 10.1101/2020.04.02.20050922]

14. Karnakov P, Arampatzis G, Kičić I, Wermelinger F, Wälchli D, Papadimitriou C, et al. Data driven inference of the reproduction number (R0) for COVID-19 before and after interventions for 51 European countries. medRxiv. 2020 May 23 Preprint posted online May 23, 2020. [doi: 10.1101/2020.05.21.20109314]

15. Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. Am J Epidemiol 2013 Nov 01;178(9):1505-1512 [FREE Full text] [doi: 10.1093/aje/kwt133] [Medline: 24043437]

16. Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. Proc Biol Sci 2007 Mar 22;274(1609):599-604 [FREE Full text] [doi: 10.1098/rspb.2006.3754] [Medline: 17476782]

17. Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. Am J Epidemiol 2004 Sep 15;160(6):509-516 [FREE Full text] [doi: 10.1093/aje/kwh255] [Medline: 15353409]

18. Arons MM, Hatfield KM, Reddy SC, Kimball A, James A, Jacobs JR, Public Health–SeattleKing CountyCDC COVID-19 Investigation Team. Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility. N Engl J Med 2020 May 28;382(22):2081-2090 [FREE Full text] [doi: 10.1056/NEJMoa2008457] [Medline: 32329971]

19. Hu Z, Song C, Xu C, Jin G, Chen Y, Xu X, et al. Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China. Sci China Life Sci 2020 May;63(5):706-711 [FREE Full text] [doi: 10.1007/s11427-020-1661-4] [Medline: 32146694]

20. Tindale LC, Stockdale JE, Coombe M, Garlock ES, Lau WYV, Saraswat M, et al. Evidence for transmission of COVID-19 prior to symptom onset. Elife 2020 Jun 22;9:e57149 [FREE Full text] [doi: 10.7554/eLife.57149] [Medline: 32568070]

21. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N Engl J Med 2020 Mar 26;382(13):1199-1207 [FREE Full text] [doi: 10.1056/NEJMoa2001316] [Medline: 31995857]

22. Linton N, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov A, Jung S, et al. Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data. medRxiv. 2020 Feb 18 Preprint posted online February 18, 2020. [doi: 10.1101/2020.01.26.20018754]

23. Coronavirus disease (COVID-19). World Health Organization. 2020 Oct 12. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/q-a-coronaviruses [accessed 2021-03-29]

24. Systrom K, Vladeck T, Krieger M. Rt.live. 2020. URL: https://github.com/rtcovidlive/covid-model

25. Maier BF, Brockmann D. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. Science 2020 May 15;368(6492):742-746 [FREE Full text] [doi: 10.1126/science.abb4557] [Medline: 32269067]

26. Roda WC, Varughese MB, Han D, Li MY. Why is it difficult to accurately predict the COVID-19 epidemic? Infect Dis Model 2020;5:271-281 [FREE Full text] [doi: 10.1016/j.idm.2020.03.001] [Medline: 32289100]

27. Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. JAMA 2020 Feb 24:1239-1242. [doi: 10.1001/jama.2020.2648] [Medline: 32091533]

28. Lourenço J, Paton R, Ghafari M, Kraemer M, Thompson C, Simmonds P, et al. Fundamental principles of epidemic spread highlight the immediate need for large-scale serological surveys to assess the stage of the SARS-CoV-2 epidemic. medRxiv. 2020 Dec 22 Preprint posted online December 22, 2020. [doi: 10.1101/2020.03.24.20042291]

29. Wangping J, Ke H, Yang S, Wenzhe C, Shengshu W, Shanshan Y, et al. Extended SIR Prediction of the Epidemics Trend of COVID-19 in Italy and Compared With Hunan, China. Front Med (Lausanne) 2020;7:169 [FREE Full text] [doi: 10.3389/fmed.2020.00169] [Medline: 32435645]

30. Abbott S, Hellewell J, Thompson RN, Sherratt K, Gibbs HP, Bosse NI, et al. Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. Wellcome Open Res 2020 Jun 1;5:112. [doi: 10.12688/wellcomeopenres.16006.1]

31. Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, Centre for Mathematical Modelling of Infectious Diseases COVID-19 working group. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. Lancet Infect Dis 2020 May;20(5):553-558 [FREE Full text] [doi: 10.1016/S1473-3099(20)30144-4] [Medline: 32171059]

32. Hu H, Nigmatulina K, Eckhoff P. The scaling of contact rates with population density for the infectious disease models. Math Biosci 2013 Aug;244(2):125-134 [FREE Full text] [doi: 10.1016/j.mbs.2013.04.013] [Medline: 23665296]

33. Sy K, White L, Nichols B. Population density and basic reproductive number of COVID-19 across United States counties. medRxiv. 2020 Jun 13 Preprint posted online June 13, 2020. [doi: 10.1101/2020.06.12.20130021]

34. Prem K, Liu Y, Russell TW, Kucharski AJ, Eggo RM, Davies N, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, et al. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. Lancet Public Health 2020 May;5(5):e261-e270 [FREE Full text] [doi: 10.1016/S2468-2667(20)30073-6] [Medline: 32220655]

35. Delen D, Eryarsoy E, Davazdahemami B. No Place Like Home: Cross-National Data Analysis of the Efficacy of Social Distancing During the COVID-19 Pandemic. JMIR Public Health Surveill 2020 May 28;6(2):e19862 [FREE Full text] [doi: 10.2196/19862] [Medline: 32434145]

36. Mukherjee UK, Bose S, Ivanov A, Souyris S, Seshadri S, Sridhar P, et al. Evaluation of reopening strategies for educational institutions during COVID-19 through agent based simulation. Sci Rep 2021 Mar 17;11(1):6264 [FREE Full text] [doi: 10.1038/s41598-021-84192-y] [Medline: 33731722]

37. Paltiel AD, Zheng A, Walensky RP. Assessment of SARS-CoV-2 Screening Strategies to Permit the Safe Reopening of College Campuses in the United States. JAMA Netw Open 2020 Jul 01;3(7):e2016818 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.16818] [Medline: 32735339]

38. Panovska-Griffiths J, Kerr CC, Stuart RM, Mistry D, Klein DJ, Viner RM, et al. Determining the optimal strategy for reopening schools, the impact of test and trace interventions, and the risk of occurrence of a second COVID-19 epidemic wave in the UK: a modelling study. Lancet Child Adolesc Health 2020 Nov;4(11):817-827 [FREE Full text] [doi: 10.1016/S2352-4642(20)30250-9] [Medline: 32758453]

39. Friedman E, Friedman J, Johnson S, Landsberg A. Transitioning Out of the Coronavirus Lockdown: A Framework for Evaluating Zone-Based Social Distancing. Front Public Health 2020;8:266 [FREE Full text] [doi: 10.3389/fpubh.2020.00266] [Medline: 32587845]

40. Inglesby TV. Public Health Measures and the Reproduction Number of SARS-CoV-2. JAMA 2020 Jun 02;323(21):2186-2187. [doi: 10.1001/jama.2020.7878] [Medline: 32356869]

41. Allen LJS. A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis. Infect Dis Model 2017 May;2(2):128-142 [FREE Full text] [doi: 10.1016/j.idm.2017.03.001] [Medline: 29928733]

42. Zhang Y, Li Y, Wang L, Li M, Zhou X. Evaluating Transmission Heterogeneity and Super-Spreading Event of COVID-19 in a Metropolis of China. Int J Environ Res Public Health 2020 May 24;17(10):3705 [FREE Full text] [doi: 10.3390/ijerph17103705] [Medline: 32456346]

43. Lau MSY, Grenfell B, Thomas M, Bryan M, Nelson K, Lopman B. Characterizing superspreading events and age-specific infectiousness of SARS-CoV-2 transmission in Georgia, USA. Proc Natl Acad Sci U S A 2020 Sep 08;117(36):22430-22435 [FREE Full text] [doi: 10.1073/pnas.2011802117] [Medline: 32820074]

44. Adam DC, Wu P, Wong JY, Lau EHY, Tsang TK, Cauchemez S, et al. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. Nat Med 2020 Nov;26(11):1714-1719. [doi: 10.1038/s41591-020-1092-0] [Medline: 32943787]

45. Lawson AB, Kim J. Space-time covid-19 Bayesian SIR modeling in South Carolina. PLoS One 2021;16(3):e0242777 [FREE Full text] [doi: 10.1371/journal.pone.0242777] [Medline: 33730035]

46. Goldstein E, Lipsitch M, Cevik M. On the effect of age on the transmission of SARS-CoV-2 in households, schools and the community. medRxiv. 2020 Jul 28 Preprint posted online July 28, 2020. [doi: 10.1101/2020.07.19.20157362]

47.   Davies NG, Klepac P, Liu Y, Prem K, Jit M, CMMID COVID-19 working group, et al. Age-dependent effects in the transmission and control of COVID-19 epidemics. Nat Med 2020 Aug;26(8):1205-1211. [doi: 10.1038/s41591-020-0962-9] [Medline: 32546824]

48.   Liu Y, Gu Z, Xia S, Shi B, Zhou X, Shi Y, et al. What are the underlying transmission patterns of COVID-19 outbreak? An age-specific social contact characterization. EClinicalMedicine 2020 May;22:100354 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100354] [Medline: 32313879]

49.   Turk PJ, Chou S, Kowalkowski MA, Palmer PP, Priem JS, Spencer MD, et al. Modeling COVID-19 Latent Prevalence to Assess a Public Health Intervention at a State and Regional Scale: Retrospective Cohort Study. JMIR Public Health Surveill 2020 Jun 19;6(2):e19353 [FREE Full text] [doi: 10.2196/19353] [Medline: 32427104]

50.   Mehta M, Julaiti J, Griffin P, Kumara S. Early Stage Machine Learning-Based Prediction of US County Vulnerability to the COVID-19 Pandemic: Machine Learning Approach. JMIR Public Health Surveill 2020 Sep 11;6(3):e19446 [FREE Full text] [doi: 10.2196/19446] [Medline: 32784193]

51.   Murillo-Zamora E, Mendoza-Cano O, Delgado-Enciso I, Hernandez-Suarez C. Predictors of severe symptomatic laboratory-confirmed SARS-COV-2 reinfection. medRxiv. 2020 Oct 18 Preprint posted online October 18, 2020. [doi: 10.1101/2020.10.14.20212720]

52.   Gousseff M, Penot P, Gallay L, Batisse D, Benech N, Bouiller K, in behalf of the COCOREC study group. Clinical recurrences of COVID-19 symptoms after recovery: Viral relapse, reinfection or inflammatory rebound? J Infect 2020 Nov;81(5):816-846 [FREE Full text] [doi: 10.1016/j.jinf.2020.06.073] [Medline: 32619697]

53.   To K, Hung I, Ip J, Chu A, Chan W, Tam A, et al. COVID-19 re-infection by a phylogenetically distinct SARS-coronavirus-2 strain confirmed by whole genome sequencing. Clin Infect Dis 2020 Aug 25:ciaa1275 [FREE Full text] [doi: 10.1093/cid/ciaa1275] [Medline: 32840608]

54.   Raddad L, Chemaitelly H, Malek J, Ahmed A, Mohamoud Y, Younuskunju S, et al. Assessment of the risk of SARS-CoV-2 reinfection in an intense re-exposure setting. medRxiv. 2020 Sep 28 Preprint posted online September 28, 2020. [doi: 10.1101/2020.08.24.20179457]

55.   McMahon A, Robb NC. Reinfection with SARS-CoV-2: Discrete SIR (Susceptible, Infected, Recovered) Modeling Using Empirical Infection Data. JMIR Public Health Surveill 2020 Nov 16;6(4):e21168 [FREE Full text] [doi: 10.2196/21168] [Medline: 33052872]

Abbreviations

SIR: Susceptible-Infectious-Removed
aSIR: adaptive Susceptible-Infectious-Removed
JHU: Johns Hopkins University