Big Data and Information Analytics

http://www.aimspress.com/journal/BDIA

BDIA, 6: 12–25.
DOI:10.3934/bdia.2021002
Received: 21 October 2020
Accepted: 08 March 2021
Published: 15 March 2021

Research article

Mathematical modeling of transmission dynamics of COVID-19

Shuqi Wang[1], Wen Tang[1], Liyan Xiong[1], Mengyu Fang[1], Bingsong Zhang[1], Chi-Yang Chiu[2] and Ruzong Fan[1,*]

[1] Department of Biostatistics, Bioinformatics, and Biomathematics, Georgetown University Medical Center, 4000 Reservoir Rd., N.W.,Washington, DC 20057, USA

[2] Division of Biostatistics, Department of Preventive Medicine, University of Tennessee Health Science Center, 66 N. Pauline Street,Memphis, TN 38163, USA

* **Correspondence:** Email: rf740@georgetown.edu; Tel: +12026878518.

Abstract: The emergence of coronavirus disease 2019 (COVID-19) demonstrates the importance of research on understanding and accurately modeling the transmission and spread of pandemic. In this paper, we consider a susceptible-exposed-infected-recovered-deceased (SEIRD) system of differential equations to describe relationship among the number of susceptible individuals, the number of exposed individuals who are transmitting the virus, the number of infected individuals among the exposed people, the number of recovered individuals from those infected, and the number of deaths from those infected in a town, state or country. Based on the empirical results of transmission process of COVID-19 in the United States from April 16th to June 30th, 2020, we consider a few cases of contact rate, incidence rate, recovery rate, and mortality rate to model the transmission and dynamics of the virus. Numerical analysis and analytical method are used to explore the dynamics and prediction of the pandemic.

Keywords: COVID-19; incidence rate; mortality rate; susceptible-infected-recovered models; susceptible-exposed-infected-recovered-deceased models

1. Introduction

An outbreak of pneumonia initially took place in Wuhan, China, which turned out to be coronavirus disease 2019 (COVID-19), caused by a novel coronavirus (SARS-CoV-2). The pandemic has brought about devastating public health, economic, and political crises. The outbreak has spread to 222 countries and territories, inflicted more than 116.17 million confirmed cases, and claimed more than 2.58 million lives, as reported on March 7th, 2021 [1]. People residing in the United States (US) have paid a high price; the pandemic has caused more than 29.05 million confirmed cases and claimed

527,726 lives, as reported on March 11th, 2021 [2, 3]. The emergence of COVID-19 pandemic demonstrates the importance of research on accurately understanding and modeling the transmission and spread of the COVID-19 and SARS-CoV-2 [4–6]. It is essential to understand the mathematical mechanism and dynamics of COVID-19 to predict the pandemic better and evaluate the effects of intervention strategies and offer effective policies [7]. This paper is motivated by the need for in-depth investigations.

The COVID-19 is a communicable disease. An exposed person may interact with and transmit the virus to a susceptible individual. An interaction between the exposed person and the susceptible individual may result in the susceptible individual becoming exposed. An exposed individual may experience an incubation period before becoming ill, i.e., infected. Thus, an exposed person is infectious but not yet infected. An infected individual may be lucky enough to recover, at least to some extent, after a period of time. Alternatively, they will die. This process can be described by a dynamical system of susceptible-exposed-infected-recovered-deceased (SEIRD) differential equations, which can connect the numbers of susceptible, exposed, infected, recovered, and deceased individuals. Figure 1 shows the SEIRD model flow diagram of the transmission of COVID-19.

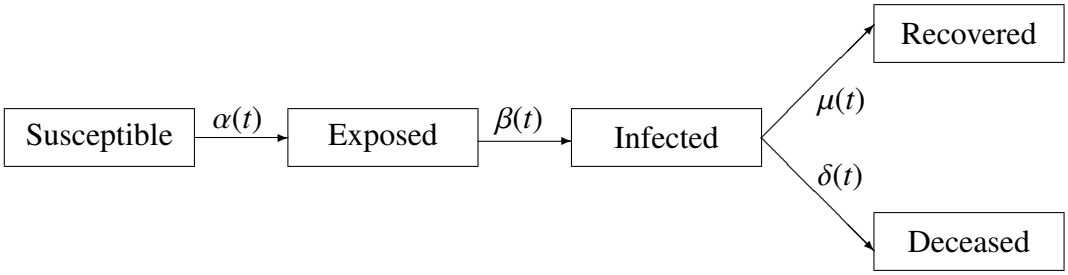

Figure 1. The SEIRD model flow diagram of the transmission of COVID-19.

In the Figure 1, $\alpha(t)$ is the contact rate between susceptible and exposed people, $\beta(t)$ is the incidence rate of the exposed people becoming infected, $\mu(t)$ is the recovery rate of the infected individuals, and $\delta(t)$ is the mortality rate of the infected individuals. In the literature, a system of susceptible-exposed-infected-recovered (SEIR) differential equations was developed to describe diseases like chickenpox and vector-borne diseases such as dengue hemorrhagic fever [8–32]. These diseases have a long incubation period, during which time the exposed individual cannot yet transmit the pathogen to others. The SEIR assumes that only an infected individual is infectious and can transmit the virus to a susceptible person and an exposed person is not infectious, as with susceptible-infected-recovered (SIR) models. For the very dangerous COVID-19, an exposed individual is infectious and can transmit the virus to susceptible people rapidly. An infected person is infectious but may be sent to hospitals quickly and separated from the susceptible people, making them unlikely to transmit the virus to susceptible people. Therefore, the existing SEIR and SIR systems cannot be used to analyze the COVID-19, and we need to build new models to fill the gap.

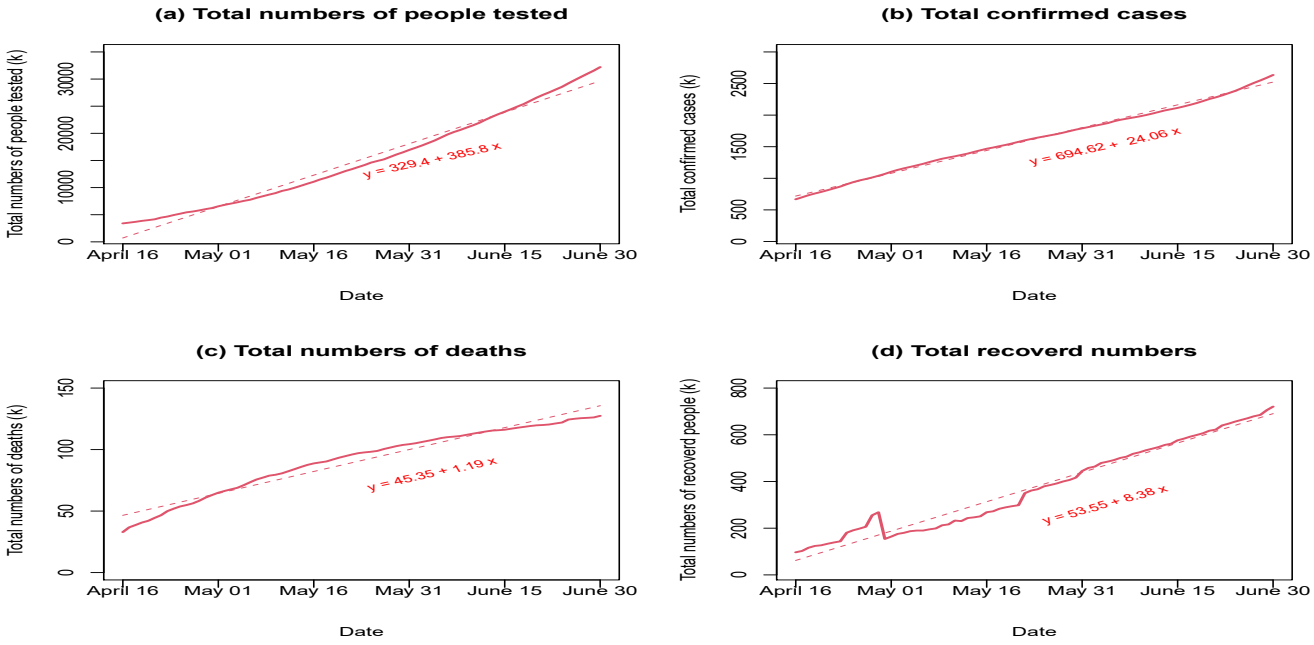

Figure 2. The trajectories of total numbers of people tested, confirmed cases, deaths, and recovered as well as related linear regression lines for COVID-19 in the United States.

To get an intuitive idea of these rates for the COVID-19 transmission process, we obtain population data from US Census Bureau and COVID-19 daily reports from the Johns Hopkins University website through a GitHub repository [4–6]. The data are then combined for a unified analysis. We then calculate the empirical incidence, mortality, and recovered rates. Figure 2 provides trajectories of total numbers of people tested, confirmed cases, deaths, and recovered as well as related linear regression lines for COVID-19 in the US from April 16th to June 30th, 2020.

In Figures 3 and S1, we show trajectories of empirical testing rate, incidence rate, mortality rate, and recovery rate as well as related quadratic and linear regression lines for COVID-19 based on the results of Figure 2, respectively. It can be seen from the Figures 3 and S1 that both the testing rate and recovery rate increase, the incidence rate decreases, and the mortality rate increases at the beginning and then remains relatively stable before decreasing. The features of these rate may be due to people's being inexperienced at the beginning of the pandemic but gradually becoming more experienced at dealing with it and getting better prepared.

The empirical results shown in Figures 1, 3 and S1 provide us with a rationale for building the models. One noticeable feature is that these rates are not constants but time-dependent. Figures 1, 3 and S1 are based on COVID-19 data in the US from April 16th to June 30th, 2020, in which the sample size is large, and the deterministic SEIRD system is appropriate. The results of Figures 1, 3 and S1 provide a good estimate of the incidence, mortality, and recovery rates of COVID-19. Based on the empirical results, we can consider a few cases of contact, incidence, recovery, and mortality rates to model dynamics of COVID-19.

In this article, we build a system of SEIRD ordinary differential equations to describe relationship among the numbers of susceptible, exposed, infected, recovered, and deceased individuals in a town, state or country based on the empirical rates. The organization of the paper is as follows. In section 2,

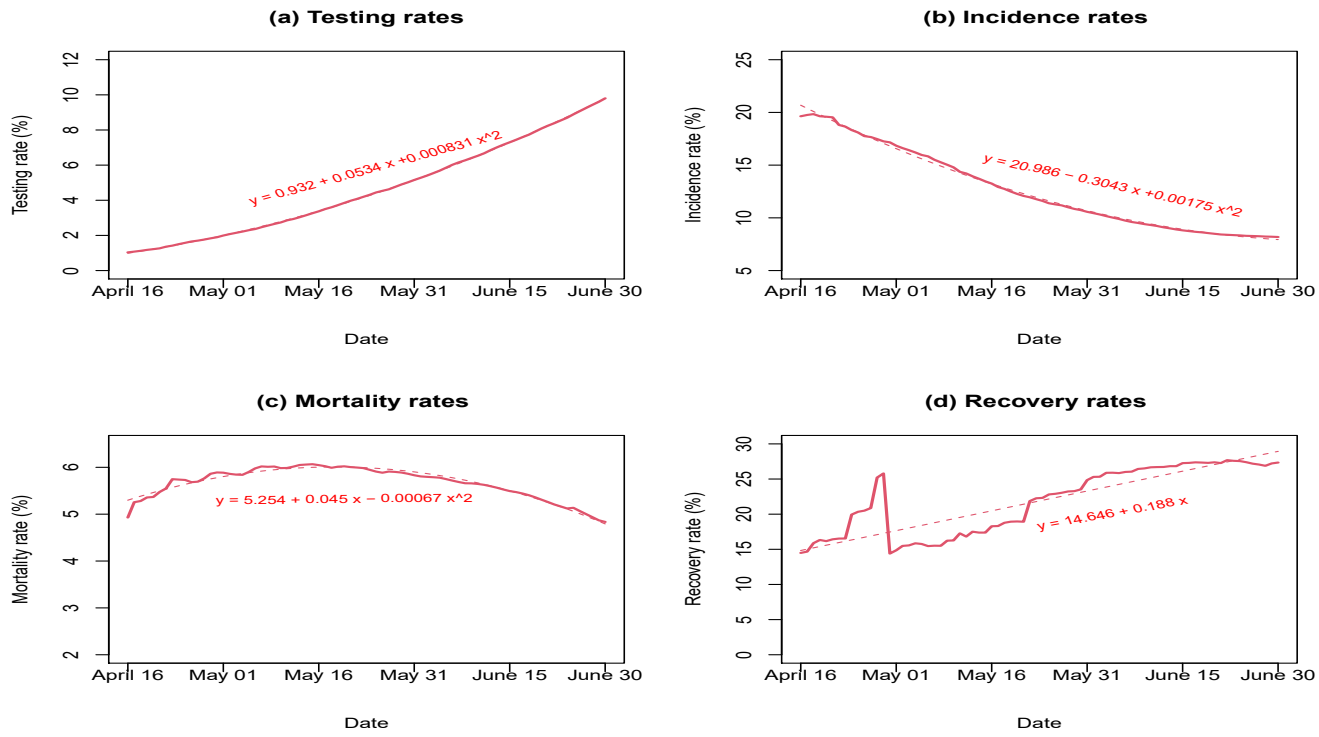

Figure 3. The trajectories of testing rates, incidence rates, mortality rates, and recovery rates as well as related quadratic and linear regression lines for COVID-19 in the United States.

we build SEIRD models in details. We perform empirical analysis for the SEIRD models in section 3. In section 4, we provide some discussion about the property of the SEIRD models and some thoughts about future research direction.

2. The SEIRD models

Consider an independent time variable t which is measured in days. At time t, let $s(t)$ be the number of susceptible individuals in a town, state or country; $e(t)$, the number of exposed individuals that have transmitted the virus; $i(t)$, the number of exposed individuals that are infected; $r(t)$, the number of infected individuals that have recovered; and $d(t)$, the number of infected individuals that have died. To describe the transmission process of COVID-19, consider a SEIRD system of ordinary differential equations

$$\begin{aligned}
\frac{ds(t)}{dt} &= -\alpha(t)s(t)e(t), \\
\frac{de(t)}{dt} &= \alpha(t)s(t)e(t) - \beta(t)e(t), \\
\frac{di(t)}{dt} &= \beta(t)e(t) - \mu(t)i(t) - \delta(t)i(t), \\
\frac{dr(t)}{dt} &= \mu(t)i(t),
\end{aligned} \tag{2.1}$$

$$\frac{dd(t)}{dt} = \delta(t)i(t),$$

with initial conditions $s(0) = S_0, e(0) = E_0, i(0) = I_0, r(0) = 0$, and $d(0) = 0$. In the system (2.1), $\alpha(t)$ is contact rate, $\beta(t)$ is incidence rate, $\mu(t)$ is recovery rate, and $\delta(t)$ is mortality rate displayed in the Figure 1. Note all these rates are non-negative. The first equation in the system (2.1) is called susceptible equation, the second equation in the system (2.1) is called exposed equation, the third is called infected equation, the fourth is called recovery equation, and the fifth is called deceased equation, respectively. The system (2.1) is different from those in the literature of SEIR system as the susceptible and exposed equations are different. In the SEIR models, it is assumed that only an infected individual can transmit the virus to a susceptible person like SIR models and so the first two SEIR equations are

$$\begin{aligned}\frac{ds(t)}{dt} &= -\alpha(t)s(t)i(t),\\ \frac{de(t)}{dt} &= \alpha(t)s(t)i(t) - \beta(t)e(t).\end{aligned}$$

For COVID-19, however, an exposed individual is infectious and may interact with a susceptible individual, thus transmitting the virus. This transmission is described by the susceptible equation in system (2.1). The exposed equation in system (2.1) implies that an exposed individual may contract coronavirus and experience an incubation period before becoming ill, i.e., infected. Once contracting the virus, an exposed person can transmit the virus to a susceptible individual. The infected can either recover with a recovery rate $\mu(t)$ or die with a mortality rate $\delta(t)$.

Let us denote $N = S_0 + E_0 + I_0$. Note

$$\frac{ds(t)}{dt} + \frac{de(t)}{dt} + \frac{di(t)}{dt} + \frac{dr(t)}{dt} + \frac{dd(t)}{dt} = 0.$$

It follows that $s(t) + e(t) + i(t) + r(t) + d(t) = N$ is a constant. Since $\frac{ds(t)}{dt} = -\alpha(t)s(t)e(t) \leq 0$, the susceptible function $s(t)$ is a decreasing function. Similarly, the recovery function $r(t)$ and deceased function $d(t)$ are increasing functions since $\frac{dr(t)}{dt} = \mu(t)i(t) \geq 0$ and $\frac{dd(t)}{dt} = \delta(t)i(t) \geq 0$, respectively. In the following, we impose conditions on incidence rate $\beta(t)$, recovery rate $\mu(t)$, and mortality rate $\delta(t)$ based on the empirical results of the Figures 3 and S1 as follows

$$\begin{aligned}&\text{Quadratic decline of incidence rate: } \beta(t) = \beta_0 - \beta_1 t + \beta_2 t^2 \geq 0, \beta_0 > 0, \beta_1 > 0, \beta_2 > 0,\\ &\text{Linear decline of incidence rate: } \beta(t) = \beta_0 - \beta_1 t \geq 0, \beta_0 > 0, \beta_1 > 0,\\ &\text{Linear increase of recovery rate: } \mu(t) = \mu_0 + \mu_1 t, \mu_0 \geq 0, \mu_1 \geq 0,\\ &\text{Quadratic mortality rate: } \delta(t) = \delta_0 + \delta_1 t + \delta_2 t^2 \geq 0, \delta_0 \geq 0, \delta_1 \geq 0, \delta_2 < 0.\end{aligned} \tag{2.2}$$

As shown in Figure 3, a declining incidence rate implies that the proportion of infected people is high at the beginning but decreases as time goes by. The empirical mortality rate is assumed to be a quadratic curve. It implies that the mortality rate increases initially, after which it remains stable and eventually decreases. One possible reason for this is that the virus is strongly virulent at the beginning but gradually weakens. Also, healthcare for infected people improves, and this contributes to an eventual decline in mortality. On the other hand, an increasing recovery rate is most likely due to improving healthcare and the fact that people are more experienced at handling the disease as

time progresses. We assume that the recovery rate is linearly increasing from the empirical result of COVID-19 in the US, as shown in Figure 3.

For the contact rate $\alpha(t)$, we consider an exponential declining rate: $\alpha(t) = \alpha_0 \exp(-\alpha_1 t)$, $\alpha_0 > 0$, $\alpha_1 > 0$. The rationale for a declining contact rate is that people may contact each other as normal at the beginning stage of the pandemic but reduce contact quickly as time moves on. The exponential decline means that the contact reduces to a minimum fast.

From the susceptible equation and the exposed equation in the system (2.1), we have

$$\frac{de(t)}{ds(t)} = -1 + \frac{\beta(t)}{\alpha(t)s(t)}.$$

Integrating the above equation gives

$$e(t) = E_0 + S_0 - s(t) + \int_0^t \frac{\beta(u)}{\alpha(u)s(u)} ds(u).$$

If $\alpha(t) = \alpha_0$ and $\beta(t) = \beta_0$ are constants, the above equation provides an exact solution

$$e(t) = E_0 + S_0 - s(t) + \frac{1}{q}\ln(s(t)/S_0), q = \frac{\alpha_0}{\beta_0}. \tag{2.3}$$

In addition, $\frac{de(t)}{ds(t)} = -1 + \frac{\beta_0}{\alpha_0 s(t)} = 0$ implies that $e(t)$ reaches its maximum when $s(t) = 1/q$. Hence, the maximal number of exposed individual is given by

$$e_{\max} = E_0 + S_0 - \frac{1}{q}\left(1 + \ln(qS_0)\right), \text{ if } \alpha(t) = \alpha_0 \text{ and } \beta(t) = \beta_0. \tag{2.4}$$

3. Numerical results

Under an assumption of constant contacting rate $\alpha(t) = \alpha_0$ and constant incidence rate $\beta(t) = \beta_0$, we obtain the exposed function $e(t)$ and the maximal number of exposed individual $e_{\max}$ in relations (2.3) and (2.4). In general, the contact rates, incidence rates, recovery rates, and mortality rates are functions of time variable t. An exact solutions of SEIRD system (2.1) is not possible. Instead, a numerical approximation procedure can provide us a solution via Euler's method [33–35]. Taking a small time interval Δt and increment $\Delta s(t) = s(t + \Delta t) - s(t)$, we may make an approximation $\frac{ds(t)}{dt} = \frac{\Delta s(t)}{\Delta t}$. Similarly, we may approximate $\frac{di(t)}{dt} = \frac{\Delta i(t)}{\Delta t}$, where $\Delta i(t) = i(t + \Delta t) - i(t)$, etc. Then, we may obtain a system of recurrence equations as follows to approximate the numbers of susceptible, exposed, infected, recovered, and deceased

$$\begin{aligned}
s(t + \Delta t) &= s(t) - \alpha(t)s(t)e(t)\Delta t, \\
e(t + \Delta t) &= e(t) + \left[\alpha(t)s(t)e(t) - \beta(t)e(t)\right]\Delta t, \\
i(t + \Delta t) &= i(t) + \left[\beta(t)e(t) - \mu(t)i(t) - \delta(t)i(t)\right]\Delta t, \\
r(t + \Delta t) &= r(t) + \mu(t)i(t)\Delta t, \\
d(t + \Delta t) &= d(t) + \delta(t)i(t)\Delta t.
\end{aligned} \tag{3.1}$$

Using the recurrence system (3.1) and initial conditions $s(0) = S_0, e(0) = E_0, i(0) = I_0, r(0) = 0$, and $d(0) = 0$, we perform numerical approximations for a few cases of the rates defined in relations (2.2).

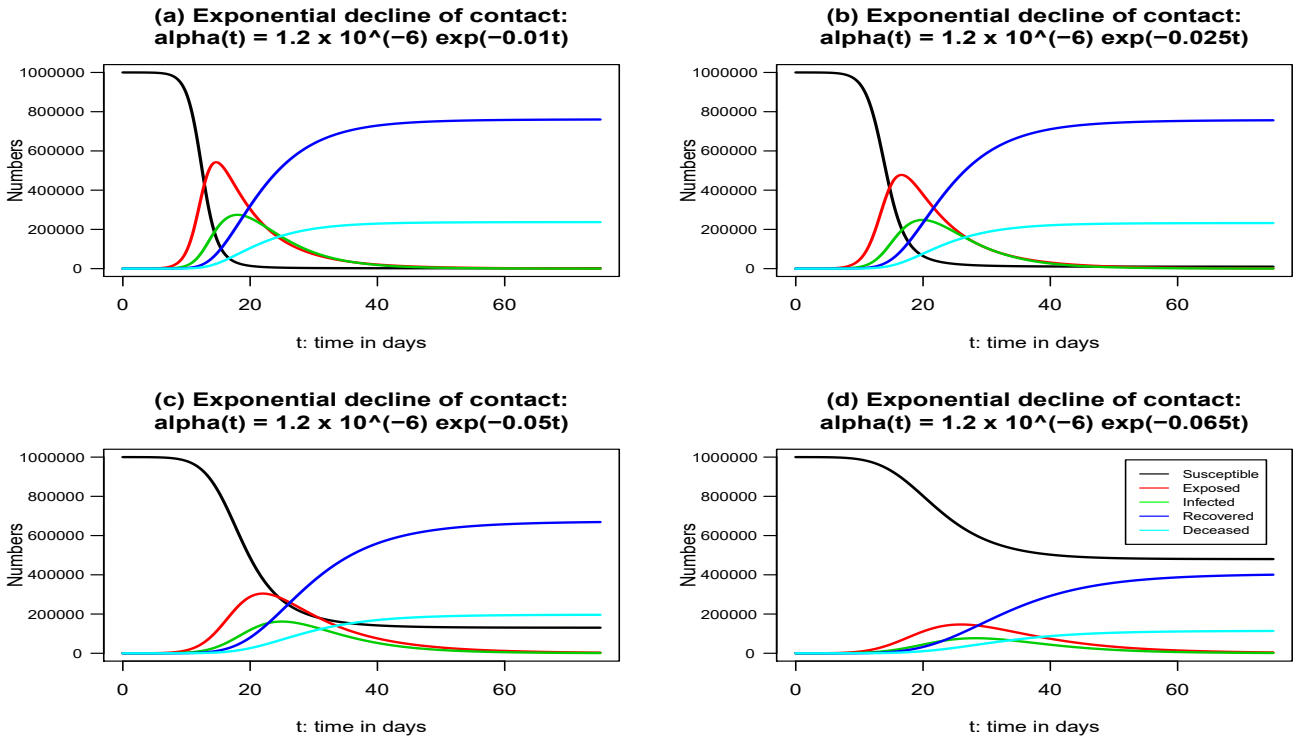

Figure 4. Approximate numbers of susceptible, exposed, infected, and deceased using exponential decline contact function $\alpha(t) = 1.2 \times 10^{-6} \exp(\alpha_1 t)$, (a) $\alpha_1 = -0.01$, (b) $\alpha_1 = -0.025$, (c) $\alpha_1 = -0.05$, and (d) $\alpha_1 = -0.065$, when other rates are fixed: $\beta(t) = 0.01(21.0 - 0.30t + 0.0018t^2)$, $\mu(t) = 0.01(14.65 + 0.188t)$, and $\delta(t) = 0.01(5.25 + 0.045t - 0.00067t^2)$. The initial conditions are given by $S_0 = 10^6$, $E_0 = 10$, and $I_0 = 2$.

By using the rate functions provided in the Figure 3, we perform numerical calculation for approximate numbers of susceptible, exposed, infected, and deceased. The incidence, recovery, and mortality rates are from Figure 3: $\beta(t) = 0.01(21.0 - 0.30t + 0.0018t^2)$, $\mu(t) = 0.01(14.65 + 0.188t)$, and $\delta(t) = 0.01(5.25 + 0.045t - 0.00067t^2)$. The initial conditions are given by $S_0 = 10^6$, $E_0 = 10$, and $I_0 = 2$. We consider four exponential decline contact functions $\alpha(t) = 1.2 \times 10^{-6} \exp(\alpha_1 t)$: (a) $\alpha_1 = -0.01$, (b) $\alpha_1 = -0.025$, (c) $\alpha_1 = -0.05$, and (d) $\alpha_1 = -0.065$, respectively. Thus, the exponential declining rates of the four contact functions decrease from -0.01 to -0.065. The results are plotted in Figure 4.

It can be seen from Figure 4(a),(b) that the number of susceptible individuals rapidly decreases to zero when $\alpha_1 = -0.01$ and $\alpha_1 = -0.025$. If $\alpha_1 = -0.05$, the number of susceptible individuals decreases and stabilizes around 100,000 when the time t is about 30 days (Figure 4 (c)); if $\alpha_1 = -0.065$, the number of susceptible individuals decreases and stabilizes around 500,000 when the time t is about 30 days (Figure 4 (d)). For each plot, the numbers of exposed and infected start low and then increase. Once they reach their peak, they decrease to low numbers again. When $\alpha_1 = -0.01$ and $\alpha_1 = -0.025$, both the numbers of exposed and infected reach high levels when the time t is about 15 days, while fewer people are exposed and infected when $\alpha_1 = -0.05$ and $\alpha_1 = -0.065$, respectively. The numbers of recovered and deceased individuals increase. The numbers of recovered and deceased decrease as the exponential declining rate decreases exponentially from –0.01 to –0.065, e.g., the curve of

recovered in Figure 4(a) is higher than that in Figure 4(d).

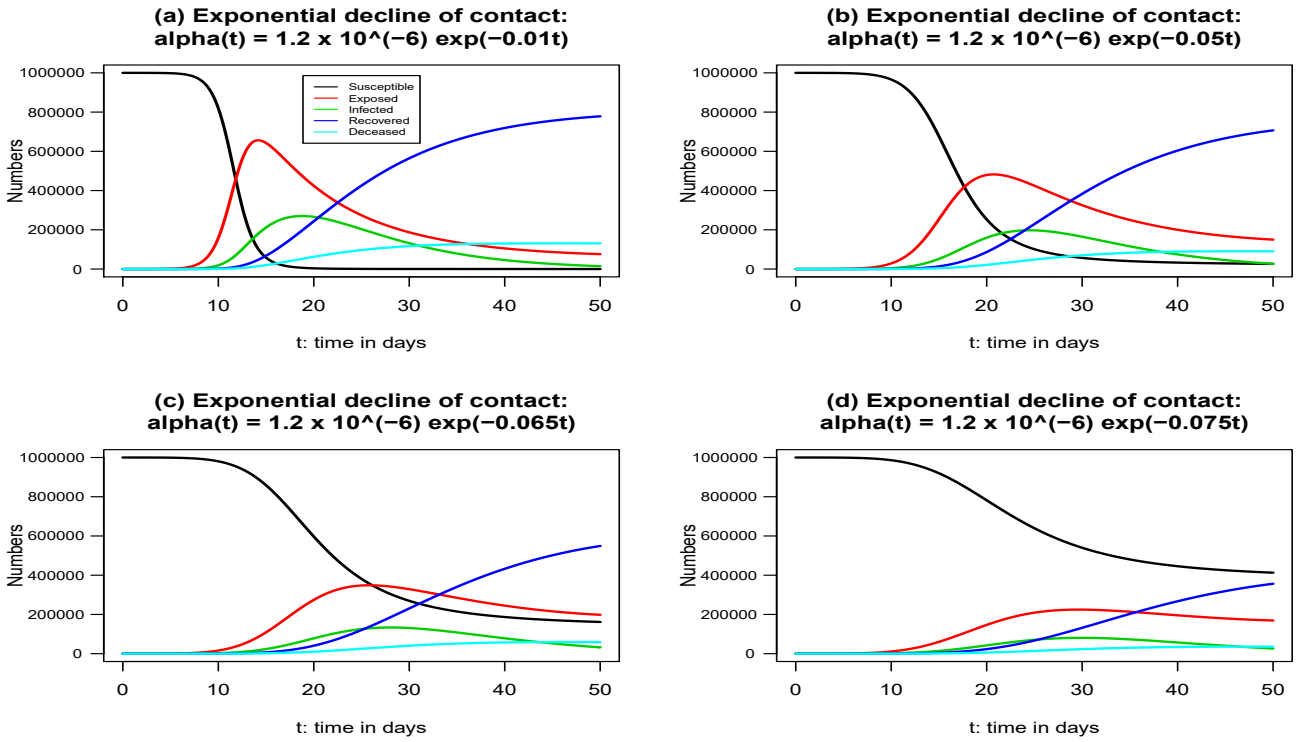

Figure 5. Approximate numbers of susceptible, exposed, infected, and deceased using exponential decline contact function $\alpha(t) = 1.2 \times 10^{-6} \exp(\alpha_1 t)$, (a) $\alpha_1 = -0.01$, (b) $\alpha_1 = -0.05$, (c) $\alpha_1 = -0.065$, and (d) $\alpha_1 = -0.075$, when other rates are fixed: $\beta(t) = 0.01(14.5 - 0.25t), \mu(t) = 0.01(8.5 + 0.30t)$, and $\delta(t) = 0.01(4.5 - 0.05t - 0.001t^2)$. The initial conditions are given by $S_0 = 10^6, E_0 = 10$, and $I_0 = 2$.

In the supplementary materials, we consider another set of incidence, recovery, and mortality rates from Figure S1: $\beta(t) = 0.01(19.25 - 0.17t)$, $\mu(t) = 0.01(14.65 + 0.20t)$, and $\delta(t) = 0.01(5.25 + 0.045t - 0.001t^2)$. In this set of rates, the incidence rate is fitted by linear regression, instead of a quadratic regression in Figure 3. Again, the initial conditions are given by $S_0 = 10^6, E_0 = 10$, and $I_0 = 2$. Using the rates, approximate numbers of susceptible, exposed, infected, and deceased are given in Figure S2, which provides similar results as Figure 4.

In Figure 5, we use another set of rates to calculate approximate numbers of susceptible, exposed, infected, and deceased: $\beta(t) = 0.01(14.5 - 0.25t), \mu(t) = 0.01(8.5 + 0.30t)$, and $\delta(t) = 0.01(4.5 - 0.05t - 0.001t^2)$. The features of the numbers of susceptible, recovered, and deceased are similar to that of Figure 4. However, the number of exposed is still high and does no reach a low level when time t is around 50 days, and the number of infected does not stabilize when time t is around 50 days. Hence, this set of rate functions leads to a good fit for the transmissions and progressions of US COVID-19 pandemic.

4. Discussion

In this article, we build SEIRD differential equations to describe relationships among the numbers of susceptible, exposed, infected, recovered, and deceased individuals. Based on the empirical rates, empirical analyses for the SEIRD models are carried out to explore the applications and implications of the models. For COVID-19, a susceptible individual may interact with an exposed person and contract the virus. Hence, an exposed individual is infectious and can transmit the virus to susceptible people rapidly. The SEIRD system differs from SEIR and SIR differential equations in the literature. In these literature models, only an infected individual is infectious and can transmit the virus to a susceptible person. The proposed SEIRD system can directly describe the transmission of COVID-19 in a susceptible population.

Using the US empirical incidence, mortality, and recovery rate functions of COVID-19, we calculate approximate numbers of susceptible, exposed, infected, recovered, and deceased. The incidence, recovery, and mortality rate functions are functions of time in days and are not constant. We consider four exponential decline contact functions to investigate the impact of contact rate on the transmission of COVID-19. An exponential declining contact rate implies that people have contact with each other as usual at the beginning, with a reduction in contact quickly later on. Our empirical results show that a very rapid decline in contact significantly reduces the numbers of exposed, infected, and deceased. Therefore, it makes sense to wear masks and reduce social activities during the pandemic.

In the literature, the SEIR models were extended to be a susceptible-exposed-infected-recovered-susceptible (SEIRS) models, where recovered people may become susceptible again, i.e., recovery does not confer lifelong immunity [25]. For example, rotavirus and malaria are diseases with long incubation periods, and recovery only confers temporary immunity. In the SEIRD system (2.1), we assume immunity for COVID-19 and that recovery confers lifelong immunity. The models can be extended to remove this assumption. As time progresses, we may have empirical data supporting the fact that recovery does not confer lifelong immunity for COVID-19; SEIRD system (2.1) can then be extended to accommodate the finding. Another extension of SEIRD system (2.1) allows for immunity by vaccination. Because vaccine rollout began in December 2020, its effectiveness and influence can be evaluated and provide us with data to help us develop an extension of the SEIRD system that will accommodate immunity by vaccination.

In this article, we only consider a deterministic system of SEIRD differential equations which can not capture random fluctuations of the numbers of susceptible, exposed, infected, recovered, and deceased individuals. The deterministic models can be extended to include stochastic components to describe random fluctuations of the virus transmissions of COVID-19 and viral phylodynamics of SARS-CoV-2 [36–45]. More research is needed for further investigations.

Acknowledgments

This study was supported by U.S. National Science Foundation grant DMS-1915904 (Ruzong Fan). One anonymous reviewer and Dr. Jianhong Wu, provided very good and insightful comments for us to improve the manuscript. Mr. Marion L. Hartley helped us to edit the paper. Ruzong Fan conceives the project; Shuqi Wang, Wen Tang, Liyan Xiong, Menyu Fang, and Chi-Yang Chiu provide technical

support for the project.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. World Health Organization (WHO), Coronavirus disease 2019 (COVID-19) situation reports. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/.
2. Center for Disease Control and Prevention (CDC), Coronavirus disease 2019 (COVID-19). Available from: https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html.
3. Center for Disease Control and Prevention (CDC), COVID-19 forecasts: cumulative deaths. Available from: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html.
4. Dong E, Du H and Gardner L, (2020) An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 20: 533–543.
5. USA daily state reports (csse_covid_19_daily_reports_us), the Johns Hopkins University. Available from:https://github.com/TWtangtang/COVID-19/tree/master/csse_covid_19_data.
6. The United States Census Bureau. Available from: https://www.census.gov/data/datasets/time-series/demo/popest/2010s-counties-total.html.
7. Kissler SM, Tedijanto C, Goldstein E, et al. (2020) Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* 368: 860–868.
8. Anderson RM, Anderson B and May RM, (1992) *Infectious Diseases of Humans: Dynamics and Control*, Oxford: Oxford University Press.
9. Antia R, Regoes R, Koella JC, et al. (2003) The role of evolution in the emergence of infectious diseases. *Nature* 426: 658–661.
10. Bailey NTJ, (1975) *The Mathematical Theory of Infectious Diseases and its Applications*, London: Griffin.
11. Bertozzi AL, Francob E, Mohlerd G, et al. (2020) The challenges of modeling and forecasting the spread of COVID-19. *Proc Natl Acad Sci* 117: 16732–16738.
12. Bjornstad ON, (2018) *Epidemics: Models and Data using R*, Springer.
13. Brauer F, (2008) Compartmental models in epidemiology. *Lect Notes Math Epidemiol* 1945: 19–79.
14. Chatterjee K, Chatterjee K, Kumar A, et al. (2020) Healthcare impact of COVID-19 epidemic in India: A stochastic mathematical model. *Med J Armed Force* 76: 147–155.
15. Earn DJD, (2008) A light introduction to modelling recurrent epidemics. *Lect Notes Math Epidemiol* 1945: 3–18.
16. Earn DJD, Rohani P, Bolker BM, et al. (2000) A simple model for complex dynamical transitions in epidemics. *Science* 287: 667–670.

17. Heesterbeek H, Anderson RM, Andreasen V, et al. (2015) Modeling infectious disease dynamics in the complex landscape of global health. *Science* 347: aaa4339.

18. Hethcote HW, (1976) Qualitative analyses of communicable disease models. *Math Biosci* 28: 335–356.

19. Hethcote HW, (2000) The mathematics of infectious diseases. *SIAM Rev* 42: 599–653.

20. Hethcote HW and van den Driessche P, (1991) Some epidemiological models with nonlinear incidence. *J Math Biol* 29: 271–287.

21. Huppert A and Katriel G, (2013) Mathematical modelling and prediction in infectious disease epidemiology. *Clin Microbiol Infect* 19: 999–1005.

22. Keeling MJ and Danon L, (2009) Mathematical modelling of infectious diseases. *Br Med Bull* 92: 33–42.

23. Keeling MJ and Rohani P, (2008) *Modeling Infectious Diseases in Humans and Animals*, Princeton University Press.

24. Kermack WO and McKendrick AG, (1927) A contribution to the mathematical theory of epidemics. *Proc R Soc A* 115: 700–721.

25. Li MY, Muldowney JS and van den Driessche P, (1991) Global stability of SEIRS models in epidemiology. *Can Appl Math Quarterly* 7: 409–425.

26. Liu X and Stechlinski P, (2012) Infectious disease models with time-varying parameters and general nonlinear incidence rate. *Appl Math Model* 36: 1974–1994.

27. Miller JC, (2012) A note on the derivation of epidemic final sizes. *Bull Math Biol* 74: 2125–2141.

28. Miller JC, (2017) Mathematical models of SIR disease spread with combined non-sexual and sexual transmission routes. *Infec Dis Model* 2: 35–55.

29. Osemwinyen AC and Diakhaby A, (2015) Mathematical modelling of the transmission dynamics of Ebola virus. *Appl Comput Math* 4: 313–320.

30. Rodrigues HS, (2016) Application of SIR epidemiological model: new trends. *Int J Appl Math Inf* 10: 92–97.

31. Siettos CI and Russo L, (2013) Mathematical modeling of infectious disease dynamics. *Virulence* 4: 295–306.

32. Tang L, Zhou Y, Wang L, et al. (2020) A review of multi-compartment infectious disease models. *Int Stat Rev* 88: 462–513.

33. Butcher JC, (2016) *Numerical Methods for Ordinary Differential Equations*, Chichester, United Kingdom: John Wiley & Sons.

34. Schittkowsky K, (2002) *NNumerical Data Fitting in Dynamical Systems - A Practical Introduction with Applications and Software*, Kluwer Academic Publishers.

35. Stoer J and Bulirsch R, (2013) *Introduction to Numerical Analysis*, New York, United States: Springer.

36. Allen LJS, (2003) *An Introduction to Stochastic Processes with Applications to Biology*, Prentice Hall.

37. Allen LJS, (2008) An introduction to stochastic epidemic models. *Lect Notes Math Epidemiol* : 81–130.

38. Allen LJS, (2017) A primer on stochastic epidemic models: formulation, numerical simulation, and analysis. *Infec Dis Model* 2: 128–142.

39. Aing RX, Liu JM, Cheung WKW, et al. (2016) Stochastic modelling of infectious diseases for heterogeneous populations. *Infec Dis Poverty* 5: 107.

40. Plank M, Binny RN, Hendy SC, et al. (2020) A stochastic model for COVID-19 spread and the effects of Alert Level 4 in Aotearoa New Zealand. *medRxiv*.

41. AHadfield J, Colin Megill C, Bell1 SM, et al. (2018) Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 34: 4121–4123.

42. Quick J, Loman NJ, Duraffour S, et al. (2016) Real-time, portable genome sequencing for Ebola surveillance. *Nature* 530: 228–232.

43. Sagulenko P, Puller V and Neher RA, (2018) Treetime: maximum-likelihood phylodynamic analysis. *Virus Evol* 4: vex042.

44. Volz EM, Pond SLK, Ward MJ, et al. (2009) Phylodynamics of infectious disease epidemics. *Genetics* 183: 1421–1430.

45. Volz EM, Koelle K and Bedford T, (2013) Viral phylodynamics. *PLOS Comput Biol* 9: e1002947.

Supplementary

In the supplementary materials, we consider another set of incidence, recovery, and mortality rates from Figure S1: $\beta(t) = 0.01(19.25 - 0.17t), \mu(t) = 0.01(14.65 + 0.20t)$, and $\delta(t) = 0.01(5.25 + 0.045t - 0.001t^2)$. In this set of rates, the incidence rate is fitted by linear regression, instead of a quadratic regression in Figure 3. Again, the initial conditions are given by $S_0 = 10^6, E_0 = 10$, and $I_0 = 2$. Using the rates, approximate numbers of susceptible, exposed, infected, and deceased are given in Figure S2, which provides similar results as Figure 4.

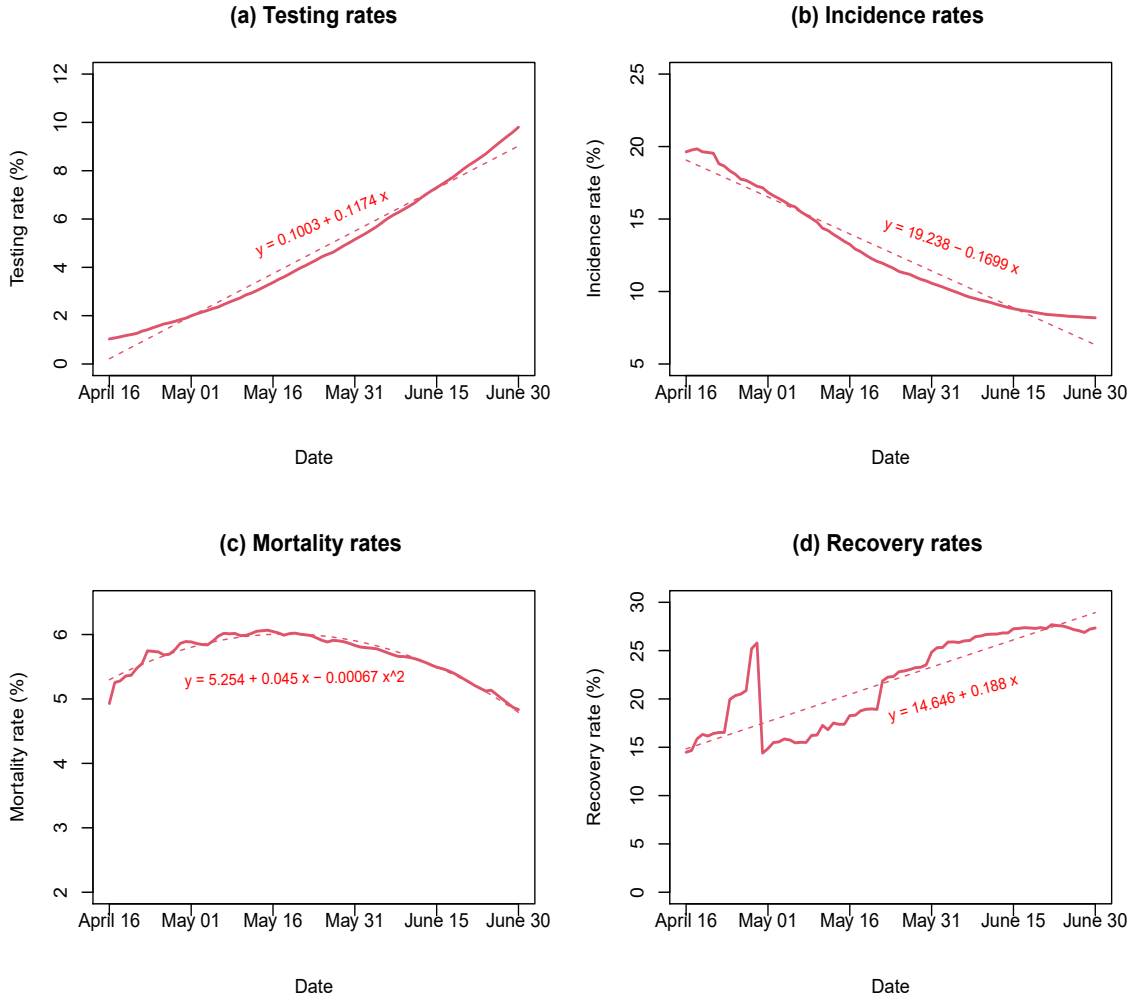

Figure S1. The testing rates, incidence rates, mortality rates, and recovery rates as well as related quadratic and linear regression lines for COVID-19 in the United States.

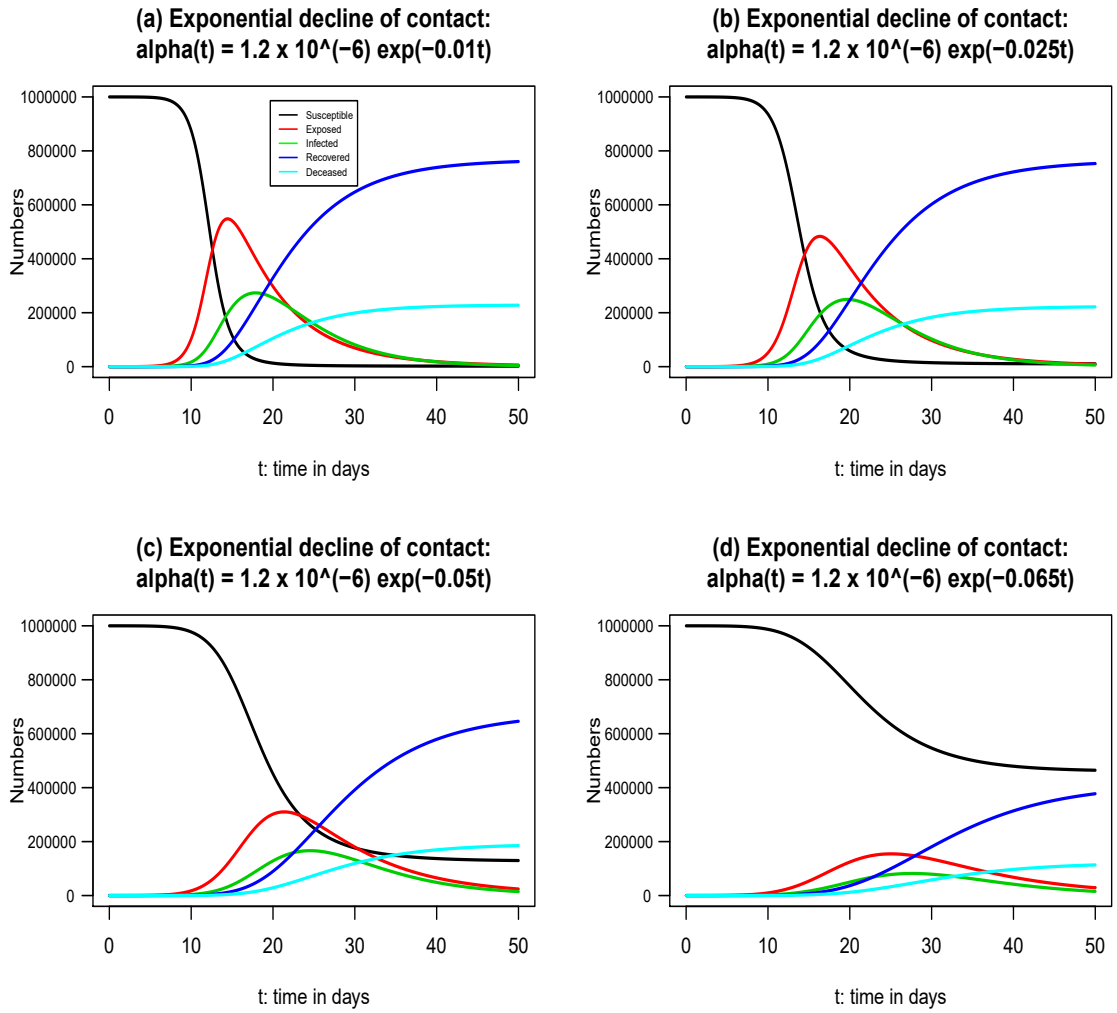

Figure S2. Approximate numbers of susceptible, exposed, infected, and deceased using exponential decline contact function $\alpha(t) = 1.2 \times 10^{-6} \exp(\alpha_1 t)$, (a) $\alpha_1 = -0.01$, (b) $\alpha_1 = -0.025$, (c) $\alpha_1 = -0.05$, and (d) $\alpha_1 = -0.065$, when other rates are fixed: $\beta(t) = 0.01(19.25 - 0.17t), \mu(t) = 0.01(14.65 + 0.20t)$, and $\delta(t) = 0.01(5.25 + 0.045t - 0.001t^2)$. The initial conditions are given by $S_0 = 10^6, E_0 = 10$, and $I_0 = 2$.