Analysis of Current Status and Simulations of Future Spread of COVID-19 in the United States

Xiaoping Liu (xiaoping.liu@hsc.wvu.edu)
West Virginia University Health Science Center https://orcid.org/0000-0002-7516-382X
A DeVries
West Virginia University

Article

Keywords: COVID-19, time-dependent transmission coefficient, social distancing, coronavirus 66 vaccination, simulations and forecast

Posted Date: March 9th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-310822/v1

Analysis of Current Status and Simulations of Future Spread of COVID-19 in the United States

Xiaoping Liu and A. Courtney DeVries

Department of Medicine, Department of Neuroscience, Rockefeller Neuroscience Institute,

West Virginia University Health Science Center, Morgantown, WV 26506 USA

Xiaoping.liu@hsc.wvu.edu

ABSTRACT

COVID-19 has killed more than 500,000 people in the United States, as of mid-February 2021. Forecasting of COVID-19 spread is helpful for key policy discussions. The transmission coefficient k_{n} of COVID-19 spread varies across time. Accurately forecasting COVID-19 spread is difficult because of the time-dependent k_{n} and becomes more complicated when coronavirus vaccination needs to be considered. In this study, the *l-i AIR* model was further developed for analyzing COVID-19 spread accompanied by coronavirus vaccinations in the United States. We determined all values of k_{n} prior to January 13, 2021 and calculated the actual number of cumulative infections (I_{n}) including asymptomatic infected individuals. We observed 4 plateaus of k_{n}, which corresponded to four national social events. This suggests that events that reduce social distancing and/or percentage of mask wearing played an important role in the acceleration of COVID-19 transmission. Our simulations show that if the American people return to their normal life before 100 million of people are vaccinated, there is likely to be at least one large surge of daily COVID-19 cases. However, if the American people partially return ($k_{\mathrm{n}} \leq 0.4$) to normal life after 100 million vaccinations, and completely return ($k_{\mathrm{n}}=1$) to normal life after two thirds of the US population are vaccinated in addition to those who have gained some immunity through coronavirus infections, the US may avoid any additional major surge of COVID-19 cases.

Keywords: COVID-19; time-dependent transmission coefficient; social distancing; coronavirus vaccination; simulations and forecast

INTRODUCTION

The first coronavirus disease 2019 (COVID-19) case was identified in the United States on January 19, 2020[1]. The patient had a history of travel to Wuhan, the epicenter of COVID-19 in China at that time, and returned to US on January 15, 2020. The first non-travel–related U.S. case was confirmed on February 26, 2020[2]. Lines of evidence suggest that limited community transmission in the U.S. likely began in late January or early February 2020, after a single importation from China and then multiple importations from Europe[3]. As of mid-February 2021, the U.S. has more than 28 million confirmed COVID-19 cases with more than 500,000 deaths[4]. In the early period of COVID-19 outbreak in the US, the number of daily new COVID-19 cases increased exponentially. To slow down the outbreak of COVID-19, the State of California became the first to mandate a state-wide order on March 19, 2020 which was followed by the vast majority of states as of early April 2020[5]. These interventions of increasing social distance and/or percentage of mask wearing slowed down the COVID-19 spread in the United States; the reported number (centered 7-day moving average) of the daily new cases of COVID-19 ($\bar{y}_{\mathrm{n}}$) gradually decreased from its peak around 32,800 in early April 2020 down to ~21,500 in early June 2020[4]. Starting from early June, $\bar{y}_{\mathrm{n}}$ surged again to start the second wave of COVID-19 spread, and peaked in late July 2020 (~70,000), and then it gradually decreased until mid-September (~35900 on September 9, 2020). After mid-September 2020, $\bar{y}_{\mathrm{n}}$ surged again to start the third wave of COVID-19 spread, and it was higher than 255,000 in January 8, 2021. The three waves of COVID-19 spread in the U.S. imply that the transmission coefficient of COVID-19 varies with time. This non-constant transmission coefficient makes it difficult to accurately forecast COVID-19 spread. In a recent study, we developed a new individual-based $l\text{-}i\,AIR$ epidemic model[6], which is especially suitable for analyzing epidemic curves with time-dependent transmission coefficient. In this study, the $l\text{-}i\,AIR$ model is further developed for considering the effect of virus vaccination on the spread of infectious diseases. We will use this new epidemical model to determine the time-dependent transmission coefficient since COVID-19 outbreak in the U.S.

from March 2020. Then, we will forecast trends in COVID-19's transmission under different scenarios with or without COVID-19 vaccinations.

METHOD

Brief description of *l-i AIR* model

The *l-i AIR* model was developed and reported in our recent paper[6]. In this model, l and i are two parameters: l represents the latent period and i represents the infectious period. A, I, and R are variables: A stands for the number of active infectious individuals or individuals in the infectious period, I for the number of cumulative infected individuals, and R for the number of recovered individuals. It is assumed that (a) the length of latent period is l days or l time units (1 time unit can be 1 day or less/more than 1 day); (b) the length of infectious period is i days or i time units; and (c) the infectious individual infects one person per day or per time unit in the infectious period. From this model, we obtained the following recursive formulas to calculate the nth term (A_n, I_n, R_n) after its previous terms are determined:

$$A_n = A_{n-1} + (I_{n-l} - I_{n-l-1}) - (I_{n-c} - I_{n-c-1}) \quad (1a)$$

$$I_n = I_{n-1} + k_n[(N - I_{n-1})/N]A_n \quad (1b)$$

$$R_n = R_{n-1} + I_{n-c} - I_{n-c-1} \quad (1c)$$

In the above equations, $c=l+i$, and N is the initial number of total susceptible individuals. To use the above formulas, we assume that the first term (A_1,I_1,R_1)=(0,1,0), all previous terms as $n<1$ are (0,0,0), and k_n=1 until a certain intervention affecting the transmission process of the infectious disease is effective. In order to use the calculated epidemic curve from Eqn. (1) to fit the reported number of daily new COVID-19 ceases, we need to induce a second parameter α, the transient incidence of the disease among the infected people, and assume:

$$y_n = \alpha A_n \quad (1d)$$

where y_n is the reported number of daily new cases of the infectious disease at day n[6].

Determinations of parameters (l and i) and coefficients (α and k_n) in the l-i *AIR* model. The values of l, i and α were determined from the reported numbers (y_n) of daily new COVID-19 cases in the early period of the COVID-19 outbreak in the United States before any major social interventions were applied for reducing COVID-19 spread. Plot of logarithm of y_n, log(y_n), in the early period vs date or n formed a straight line. The slope (S_0) of this plotted straight line was obtained from the best fitting line of the plotted data. Then, we calculated A_n and I_n from Eqn. (1) for different pairs of parameters l and i assuming k_n=1 under the condition $I_n \ll N$. Plot of logarithm of A_n, log(A_n), vs n for a given pair of l and i would also form a straight line, and thus we could obtain the slope $S(l,i)$ of the straight line for any given pair of l and i. When a pair of l and i makes the slope $S(l,i)$ to be closest to S_0, this pair of l and i was chosen to be used in this study for simulating an epidemic curve of COVID-19 in the United States. Using the chosen pair of l and i to calculate A_n from Eqn. (1), the slope $S(l,i)$ that is obtained from the linear plot of log(A_n) vs n is closest to the slope S_0 that is obtained from the linear plot of log($\bar{y}_n$) vs n; however, the intercepts of the two straight lines may have large difference. By regulating the value of α, we could change the intercept of the plot of log(αA_n) vs n. In this way, we could find a value of α, which minimizes the difference between the two straight lines (log(αA_n)~n and log($\bar{y}_n$)~n), by the method of least squares. The values of k_n will decrease after major social interventions are started. The varied k_n are determined by fitting the calculated αA_n to $\bar{y}_n$ based on Eqn. (1d). All COVID-19 data in this studied were obtained from Worldometer website (https://www.worldometers.info/coronavirus/country/us/).

Simulations of COVID-19 transmission in the United States with/without coronavirus vaccinations. Methods for simulations of US COVID-19 transmission without coronavirus vaccinations are similar to what we have reported in our recent paper[6]. In the simulations with coronavirus vaccinations, we set the starting day of coronavirus vaccinations to mid-December 2020[7]. Because 16.525 million doses of vaccines were administered between December 14, 2020 and January 20, 2021, and 29.865 million doses of vaccines were administered between January 20, 2021 and February 12, 2021[8], we assumed that 434,200 doses were injected per day on average from 12/14/2020 to 1/20/2021, and 1.3 million doses per day from 1/20/2021 to 2/11/2021. Starting from February 12, 2021, the per-day doses were set at 1.7 million[9] until the whole US population is vaccinated. On March 3, 2021, about 52.9 million people have received at least one dose of a COVID-19 vaccine, including about 27 million people who have been fully vaccinated (received both doses)[10]. Thus, a total of 80 million doses, 52.9+27≈80, have been injected by March 3, 2021. Because the number of people who have received at least one dose is 52.9 million on March 3, 2021, this number is nearly 2/3 of the total injected doses (80 million) by March 3, 2021. Using the ratio 2/3, we can convert the number of total doses injected to the number of people who have received at least one dose of vaccine.

To count in the effect of vaccination on the epidemic curve, Eqn. (1b) needs to be revised as below:

$$I_n = I_{n-1} + k_n[(N - I_{n-1} - W_{n-14})/N]A_n \qquad (1b')$$

where W_n represents the number of cumulative individuals who have not been infected by coronavirus before the vaccination on day n. The subscript n-14 in W_{n-14} means that the vaccination day is 14 days prior to day n. The 14-day waiting period is required, because the vaccine is highly efficacious from 2 weeks after the first dose[11]. It needs to be noted that W_n is not the number of the cumulative individuals who have been vaccinated. To calculate I_n from Eqn. (1b'), we need to calculate W_n first. For this purpose, we introduce two more variables U_n and V_n. Here, U_n stands for the number of cumulative individuals who have been vaccinated on day n, and V_n stands for the number of cumulative individuals who have been infected by coronavirus before the vaccination on day n. Based on these definitions, we have the relation $U_n = V_n + W_n$. The values of U_n, V_n and W_n are determined by the following 3 formulas:

$$U_n = \sum_{i=s}^{n} Z_i \qquad (2a)$$

$$V_n = V_{n-1} + Z_n(I_{n-1} - V_{n-1})/(N - U_{n-1}) \qquad (2b)$$

$$W_n = W_{n-1} + Z_n(N - U_{n-1} - (I_{n-1} - V_{n-1}))/(N - U_{n-1}) \qquad (2c)$$

In Equation (2a), Z_i is the number of individuals who are given the vaccine on day i; and s stands for the starting date of vaccination. N in Eqn. (2b) is the US population or the initial number of total susceptible individuals in the United States. No vaccinations are given to any people before day s, so all values of Z_{s-1}, U_{s-1}, V_{s-1}, W_{s-1} are 0. Because V_n represent those who have been infected by coronavirus before vaccinations, they (V_n) should be subtracted from I_n in Eqns. (2b) and (2c).

RESULTS

Parameters l, i, and coefficient α for simulating epidemic curve of COVID-19 transmission in the U.S. In the early period of COVID-19 outbreak, both the number of total COVID-19 cases and the number of daily new COVID-19 cases increase exponentially. A portion of data, between February 24, 2020 and March 22, 2020, about COVID-19 cases in the United States is listed in Table 1. Column 1 in Table 1 is date. The data of the total COVID-19 cases (t_n), which are listed in column 2, were read from the curve of total COVID-19 cases in the United States on the Worldometer website[4]. The US daily new cases (y_n), which are listed in column 3, were obtained from the differences of t_n in column 2 between two adjacent days. The centered 7-day moving averages of US daily new cases ($\bar{y}_n$), which are listed in column 4, were calculated from the US daily new cases in column 3. The logarithms of 7-day averages of the US daily new cases, log($\bar{y}_n$), are listed in column 5. In Figure 1, we demonstrated the linear plot of log($\bar{y}_n$) vs date (red open rings) and the best fitting line (black solid line) of the linear plot. The slope of the best fitting line is 0.1368. Based on Eqn. (1), we calculated A_n for different pairs of parameters l and i, assuming that k_n=1 under conditions $I_n \ll N$. Using the calculated A_n, we plotted log(A_n) vs n, which formed a straight line for each pair of parameters l and i. Slopes of these straight lines for different pairs of l and i are listed in Table 2. It can be seen that when l=4 and i=10, the slope is 0.1372, which is closest to the slope 0.1368 of the plot of log($\bar{y}_n$) vs date. Therefore, this pair of parameters l=4 and i=10 is chosen to use in this study to further determine α. It was found that when α=0.01453 and the first non-travel–related U.S.

case was assumed to begin on February 6, 2020, the difference between $\log(\alpha A_n)$ and $\log(\bar{y}_n)$ has the smallest sum of squared residuals. Using these determined parameters l=4, i=10 and α=0.01453, we calculated αA_n and $\log(\alpha A_n)$ and compared them with the reported $\bar{y}_n$ and $\log(\bar{y}_n)$ as shown in Figure 2.

Calculation of the total number (I_n) of individuals infected by coronavirus in the United States. It has been known that there are many asymptomatic and mildly symptomatic infected individuals during coronavirus transmission[12]. Using the *l-i AIR* model, we calculated the total number (I_n) of individuals infected by coronavirus, including asymptomatic and mildly symptomatic individuals that were not reported, in New York City (NYC)[6] and New York State (NYS)[13]. The calculated total numbers of infected individuals (I_n) are very close to the reported results measured by antibody tests [14-16]. To calculate I_n from Eqn. (1), we need to know k_n. Using the determined parameters l=4, i=10 and α=0.01453, we can calculate I_n, A_n and αA_n from Eqn. (1) assuming k_n=1 prior to March 19, 2020. Starting from March 19, 2020, k_n was reduced due to social interventions. In this situation, the value of k_n on each day as of January 13, 2021 was determined by fitting αA_n to the centered 7-day moving average of US daily new cases ($\bar{y}_n$). On January 13, 2021, k_n was determined as 0.18. By assuming that k_n remains unchanged at 0.18 after January 13, 2021, we simulated US daily cases αA_n, which fit well with the reported US daily cases $\bar{y}_n$ from January 13, 2021 to late February 2021 (Figure 3A). Both I_n and A_n varied with k_n. In Figure 3B, we demonstrated the calculated total number (I_n) of individuals infected by coronavirus in the U.S. and also showed the reported number of US total COVID-19 cases (t_n). The ratio of I_n to t_n varied with time. In the end of March 2020, the ratio was around 20, but it decreased to 7.8 on May 12, 2020. The calculated I_n was 52.5 million on September 30, 2020. In comparison, the reported number of total US COVID-19 cases was only about 7.5 million. The calculated I_n reached 203 million on February 18, 2021, while the reported number of total US COVID-19 cases on the same day was 28.6 million. The plot of k_n against date is shown in Figure 3C.

Simulations of future trend of COVID-19 transmission in the United States with/without coronavirus vaccinations. Accurate forecasting of COVID-19 spread is difficult because of the transmission coefficient k_n varies with time. Multiple factors, such as social distancing interventions, total number of vaccinations, percentage of people wearing face mask, and infectivity of new variants of coronavirus[17] may change k_n to affect the future course of COVID-19 transmission. The predicted future trend (solid green line) of COVID-19 transmission in the United States shown in Fig. 3A was simulated by assuming that k_n is a constant of 0.18 on and after January 13, 2021. However, people hope to go back to normal life, which means k_n = 1. Therefore, we simulated the possible future trends of COVID-19 transmission in the U.S. when people return to normal life with or without vaccinations (Figs. 4A & 4C). The red open squares in the figures represent the 7-day average of reported US daily new cases ($\bar{y}_n$) before late February 2021; the green line represents the calculated US daily cases without vaccinations; and the purple line represents the calculated US daily cases with vaccinations. Before and on January 13, 2021, k_n was determined by fitting the simulated αA_n to the reported data $\bar{y}_n$. The value of k_n (0.18) on January 13, 2021 was assumed to remain unchanged until April 12, 2021 in the simulations. We forecasted US daily new cases with/without vaccinations under two different scenarios. The first scenario is that k_n is a constant at 0.18 until April 12, 2021, and increases to 1 on April 12, 2021 (Figs. 4A & 4B). In the simulation with vaccinations (the purple line in Fig. 4A), 100 million of people in the US will have been administered coronavirus vaccines as of April 12, 2021. The second scenario is the same as the first one until April 12, 2021; however, k_n only increases to 0.4 on April 12, 2021, and to 1 on September 1, 2021 (Figs. 4C & 4D). In the simulation with vaccination (purple line), by September 1, 2021, more than 260 million (two thirds of the US population) people in the United States will be vaccinated.

DISCUSSION

In this study, we first demonstrated a linear relationship between the logarithm of US daily new cases $\log(\bar{y}_n)$ and the date in the early stage of COVID-19 outbreak (Fig. 1 and Tab. 1). Using this linear relationship, we determined parameters l and i, and further determined coefficient α (Fig. 2 and Tab. 2) The two parameters l and i determined for COVID-19 transmission in the U.S. are the same as what we determined for those in NYC[6] and NYS[13], but the determined value of α (0.0145) in the U.S. is slightly greater than the one (0.0118) in NYC and NYS. Therefore, the characteristics of COVID-19 transmission in the U.S. are very similar to those in NYC and NYS.

The determined parameters l and i imply that the average latent period and the average infectious period are 4 days and 10 days, respectively. Because the transient incidence rate α of COVID-19 for infected people in the infectious period is 0.0145, this implies that if the number of the daily new cases with COVID-19 symptoms in a large community is 1 per day, the community may have had around 1/0.0145 or 69 people in the infectious period already. Because the infectious period is 10 days, more infected people among the 69 people will show the COVID-19 symptoms each day during the infectious period. However, many infected individuals will not show symptoms or will only have mild symptoms, recovering without being tested. Our simulations show that the calculated total number of people infected by coronavirus (I_n) in the US is ~20 to ~8 times (decreasing with time) the total COVID-19 cases (t_n) reported in the US in the period from the end of March 2020 to May 12, 2020 (Fig. 3B). In comparison, tests of antibody to coronavirus in 10 US sites between March 23, 2020 to May 12, 2020 show that 6 to 24 times more infections than the total COVID-19 cases reported[18]. This indicates that these changes in the ratio simulated from our *l-i AIR* model are close to the range of the reported data from antibody tests in the same period. The calculated I_n on September 30, 2020 is 52.5 million, which is very close to 52.9 million that was estimated in a recent paper[19].

As of mid-February 2021, the US has experienced three surges of COVID-19 transmissions. The first surge of daily new COVID-19 cases ($\bar{y}_n$, the centered 7-day moving average) in the U.S. peaked around

April 7, 2020, and then $\bar{y}_n$ slowly decreased until early June 2020 due to the “stay at home” order[20]. During this period, k_n decreased from 1 on March 18, 2020 down to 0.095 gradually until early June 2020. The second surge started from early June 2020 and peaked around July 22, 2020. The third surge started from the end of August 2020, consisting of three peaks around November 22 and December 15, 2020 and January 8, 2021. Corresponding to the four peaks of the last two surges, k_n reached four plateaus or local maximums: the first plateau of k_n is 0.155 (in the 2nd surge) from June 8 to 24, 2020; the second plateau is 0.18 (in the 3rd surge) from October 17 to Nov. 4, 2020; the third plateau is 0.17 (in the 3rd surge) from November 23 to 29 of 2020; and the fourth plateau is 0.19 (in the 3rd surge) from December 20, 2020 to January 5, 2021. The time ranges of the four plateaus of k_n overlap with the following four large social events, which reduce social distance and percentage of mask wearing, respectively: 1) reopening of business after “Stay at Home” orders[20,21], 2) election related activities, 3) Thanksgiving week, and 4) Christmas and New Year holidays. This overlap in time ranges between plateaus of k_n and social events indicates that events (reducing physical distance and/or percentage of mask wearing) play an important role in accelerating COVID-19 spread in the US by increasing k_n, the transmission coefficient of COVID-19. Notably, k_n=0.18 from October 17 to November 4, 2020 caused large increases in US daily new cases. However, nearly 3 months later or starting from January 13, 2021, when k_n remains at 0.18, the number of daily new cases in the United States is rapidly decreasing. Based on our simulations, the reason for this situation is that a large portion of people have been infected by coronavirus during the 3 months (I_n=174 million on January 13, 2021). This large increase in infections is slowing down COVID-19 spread by reducing the percentage of uninfected susceptible individuals who appear around an infectious individual. Furthermore, more than 40 million people have received at least one dose of a COVID-19 vaccine from mid-December to February 21, 2021[10]. The vaccination will have a bigger and bigger effect on slowing down COVID-19 spread in the United States with the increase in the number of vaccinated people. Although the number of US daily new cases steadily decreases currently in February, 2021 and coronavirus vaccines are being given to more and more people, can we avoid another surge of COVID-19 cases in the US if people return to their normal life? In the view of our *l-i AIR* model, normal life means k_n=1. In Figs. 4A, we demonstrated that if we increase k_n to 1 on April 12, 2021 and let k_n remain at 1 after this day (Fig. 4B), we will see a very large surge of daily COVID-19 cases, and the maximal number of US daily cases will pass 800,000 if no coronavirus vaccines are given to people (green line in Fig. 4A). In comparison, under the same conditions but assuming that 100 million of people will be vaccinated by April 12, 2021, we will see a much smaller, but still a rather large surge of daily COVID-19 cases, and the maximal number of US daily cases will pass 220,000 (purple line in Fig. 4A). This result indicates that regardless of whether coronavirus vaccines (less than 100 million) are given to people, we will still see a big surge of daily COVID-19 cases when returning to normal life too early. In contrast, if we set k_n = 0.4 on April 12, 2021 and k_n = 1 on September 1, 2021 as shown in Fig. 4D, we will see 2 big surges of daily COVID-19 cases if no people are vaccinated (green line in Fig. 4C). However, under the same conditions but assuming that 100 million of people will be vaccinated by April 12, 2021 and 260 million of people will be vaccinated by September 1, 2021, we will not see any big surges of COVID-19 in the United States (purple line in Fig. 4C). This result indicates that it is possible to avoid a big surge of COVID-19 if k_n is limited to less than 0.4 after 100 million people have received coronavirus vaccines, and k_n is increased to 1 after 260 million people have received coronavirus vaccines in the United States.

In summary, in this study, we determined parameters l, i, and coefficient α of *l-i AIR* epidemic model from early data of COVID-19 transmission in the United States by using the least-squares method. Based on the determined parameters, we further determined the time dependent k_n from the beginning of the COVID-19 outbreak to January 13, 2021 by fitting the calculated US daily new cases to the reported US daily new cases (centered 7-day moving average). It was observed that, during the second and third wave of COVID-19 spreads from early June 2020 to early January 2021, the determined k_n reached four plateaus. In terms of their time ranges, the four k_n plateaus correspond to the four social events respectively: 1) reopening of business after "Stay at Home" orders, 2) election related activities, 3) Thanksgiving week, and 4) Christmas and New Year holidays. If assuming k_n=0.18 after January 13, 2021, then the simulated number of US daily new cases concurs with the reported number of US daily new cases so far (February 21, 2021). Simulations show that we will still see at least one large surge of daily COVID-19 cases in the future if people return to their normal life too early. However, if people partially return to their normal life after 100 million people have been vaccinated, and completely return to their normal life after more than 260 million people have received coronavirus vaccines in addition to those who have gained some immunity through coronavirus infections, the US population may avoid another big surge of daily COVID-19 cases in the future, assuming that the coronavirus vaccines are highly effective to current coronaviruses and their new variants, and those who were infected by coronavirus will not be re-infected.

ACKNOWLEDGEMENTS

We thank Dr. Randy Nelson for reading and commenting on drafts of this manuscript

REFERENCES

1. Holshue, M.L., *et al.* First Case of 2019 Novel Coronavirus in the United States. *The New England journal of medicine* **382**, 929-936 (2020).
2. Heinzerling, A., *et al.* Transmission of COVID-19 to Health Care Personnel During Exposures to a Hospitalized Patient - Solano County, California, February 2020. *MMWR. Morbidity and mortality weekly report* **69**, 472-476 (2020).
3. Team, C.C.-R., *et al.* Evidence for Limited Early Spread of COVID-19 Within the United States, January-February 2020. *MMWR. Morbidity and mortality weekly report* **69**, 680-684 (2020).
4. Worldometer. United States- Coronavirus Cases and Deaths. https://www.worldometers.info/coronavirus/country/us/ (2020).
5. Kates, J., Michaud, J. & Tolbert, J. Stay-At-Home Orders to Fight COVID-19 in the United States: The Risks of a Scattershot Approach. *Kaiser Family Foundation*, https://www.kff.org/coronavirus-policy-watch/stay-at-home-orders-to-fight-covid19/ (2020).
6. Liu, X. A simple, SIR-like but individual-based epidemic model: Application in comparison of COVID-19 in New York City and Wuhan. *Results Phys* **20**, 103712 (2021).
7. Guarino, B., Cha, A.E., Wood, J. & Witte, G. 'The weapon that will end the war': First coronavirus vaccine shots given outside trials in U.S. *The Washington Post*, https://www.washingtonpost.com/nation/2020/2012/2014/first-covid-vaccines-new-york/ (2020).
8. Wilson, C. The U.S. COVID-19 Vaccine Rollout Is Getting Faster. But Is It Enough? *Time. https://time.com/5938128/covid-19-vaccine-rollout-biden/* (2021).
9. Rummler, O. U.S. administering average of 1.7 million vaccine doses per day. *AXIOS. https://www.axios.com/coronavirus-vaccine-doses-average-biden-d17dc13e-ec8c-491a-89d4-c1ac651310ba.html* (2021).
10. Ivory, D., Smith, M., Lee, J.C., Walker, A.S. & Gamio, L. See How the Vaccine Rollout Is Going in Your State. *The New York Times, https://www.nytimes.com/interactive/2020/us/covid-19-vaccine-doses.html* (2021).
11. Skowronski, D.M. & De Serres, G. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N Engl J Med* **384**(2021).
12. Camero, K. COVID-19 infections in US likely 8 times higher than reported, CDC says. Here's why. *https://www.miamiherald.com/news/coronavirus/article247457275.html* **November 27**(2020).
13. Liu, X. Excel generated New York State's COVID-19 epidemic curve based on the I-i AIR model. *Mendeley Data* **doi: 10.17632/w6zjz63bmf.1**, https://data.mendeley.com/datasets/w6zjz63bmf/61 (2020).
14. Worldometer. Coronavirus (COVID-19) Mortality Rate. *https://www.worldometers.info/coronavirus/coronavirus-death-rate/?fbclid=IwAR26hCMKGKHO8_OVqgF7ySvczGxzmVlEcxLVaxwLPAdjwy6hXtRZIvjrdTw* **May 14**(2020).
15. Governor Andrew M. Cuomo announced the results of the state's completed antibody testing study on May 2, 2020. *https://www.governor.ny.gov/news/amid-ongoing-covid-19-pandemic-governor-cuomo-announces-results-completed-antibody-testing*.
16. LaVito, A., Brown, K.V. & Clukey, K. New York Finds Virus Marker in 13.9%, Suggesting Wide Spread. Bloomberg. *https://www.bloomberg.com/news/articles/2020-04-23/new-york-finds-virus-marker-in-13-9-suggesting-wide-spread* (2020).
17. Lauring, A.S. & Hodcroft, E.B. Genetic Variants of SARS-CoV-2-What Do They Mean? *JAMA* **325**, 529-531 (2021).

18. Havers, F.P., *et al.* Seroprevalence of Antibodies to SARS-CoV-2 in 10 Sites in the United States, March 23-May 12, 2020. *JAMA Intern Med* (2020).
19. Reese, H., *et al.* Estimated incidence of COVID-19 illness and hospitalization - United States, February-September, 2020. *Clin Infect Dis* (2020).
20. Moreland, A., *et al.* Timing of State and Territorial COVID-19 Stay-at-Home Orders and Changes in Population Movement — United States, March 1–May 31, 2020. *Morbidity and Mortality Weekly Report (MMWR)* **69**, 1198–1203, https://www.cdc.gov/mmwr/volumes/1169/wr/mm6935a1192.htm#suggestedcitation (2020).
21. Washington Post Staff. Where states reopened and cases spiked after the U.S. shutdown. *The Washington Post*, https://www.washingtonpost.com/graphics/2020/national/states-reopening-coronavirus-map/ (2020).

TABLES

Table 1. Reported US total COVID-19 cases (t_n), daily new cases (y_n), centered 7-day moving average of daily new cases ($\bar{y}_n$), and log($\bar{y}_n$)

Date	t_n	y_n	$\bar{y}_n$	log($\bar{y}_n$)
2/24/2020	53			
2/25/2020	57	4		
2/26/2020	60	3		
2/27/2020	60	0		
2/28/2020	63	3	7	0.827
2/29/2020	68	5	10	0.981
3/1/2020	75	7	14	1.146
3/2/2020	100	25	23	1.362
3/3/2020	124	24	37	1.563
3/4/2020	158	34	52	1.720
3/5/2020	221	63	67	1.823
3/6/2020	319	98	86	1.936
3/7/2020	435	116	124	2.094
3/8/2020	541	106	163	2.213
3/9/2020	704	163	210	2.322
3/10/2020	994	290	282	2.450
3/11/2020	1301	307	358	2.554
3/12/2020	1690	389	470	2.672
3/13/2020	2290	600	606	2.783
3/14/2020	2942	652	848	2.928
3/15/2020	3830	888	1262	3.101
3/16/2020	4948	1118	1898	3.278
3/17/2020	6928	1980	2672	3.427
3/18/2020	10134	3206	3315	3.520
3/19/2020	14978	4844	4566	3.660
3/20/2020	20997	6019		
3/21/2020	26145	5148		
3/22/2020	35790	9645		

Table 2. Slope of linear plot of log(A_n) vs n for different pairs of parameters l and i

i=	2	3	4	5	6	7	8	9	10	11	12
l=1	0.209	0.2646									
l=2	0.1221	0.166	0.1859								
l=3		0.1222	0.1400	0.1499	0.1557						
l=4			0.1128	0.1221	0.128	0.1317	0.1342	0.1359	0.1372	0.138	0.1385
l=5				0.1031	0.1089	0.1129	0.1155	0.1173	0.1185	0.1194	0.1201

FIGURES

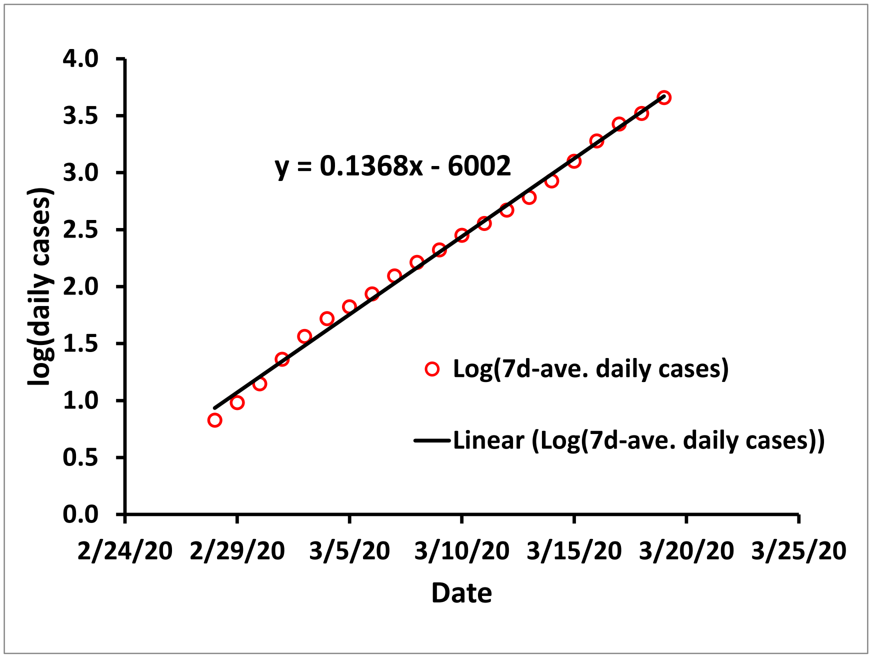
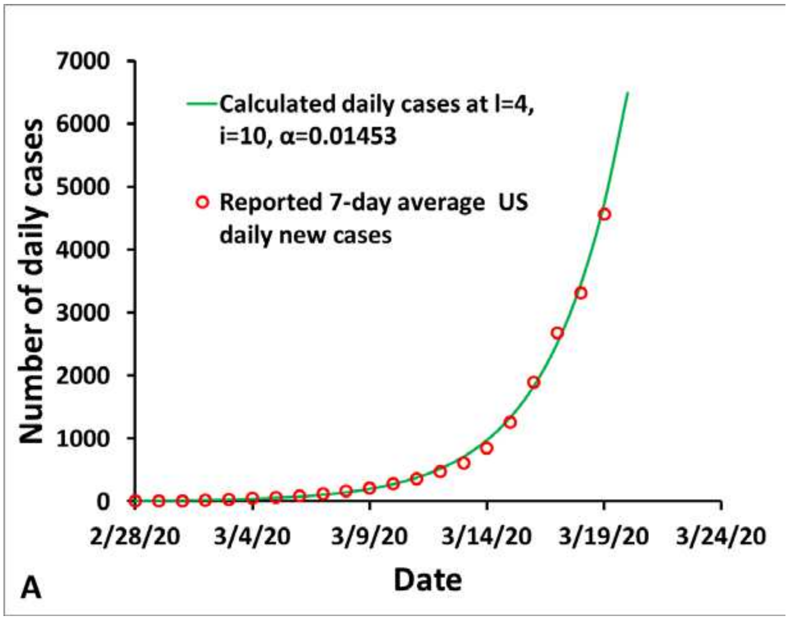


Figure 1. Logarithm of daily new cases in the U.S. in the early period between Feb 28, 2020 and Mar 19, 2020 is linear with date. The slope of the best fitting line is 0.1368.



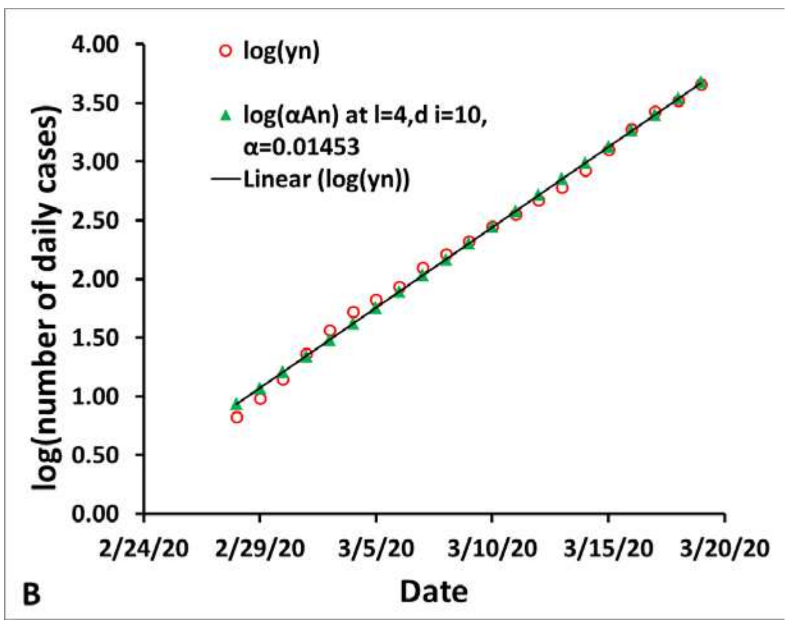


Figure 2. Exponential growth of daily cases in early period of COVID-19 outbreak in the United States between Feb 28 and Mar 19, 2020. (A) The red open circles represent centered 7-day moving average of the numbers of reported daily coronavirus cases ($\bar{y}_n$), and the green solid line represents the numbers of calculated daily coronavirus cases (αA_n) assuming l=4, i=10 and α=0.01453. (B) The red open circles represents log($\bar{y}_n$); the black solid line represents the best fitting line of log(y_n); and the green triangles represent logarithm of the calculated daily coronavirus cases (log(αA_n)).

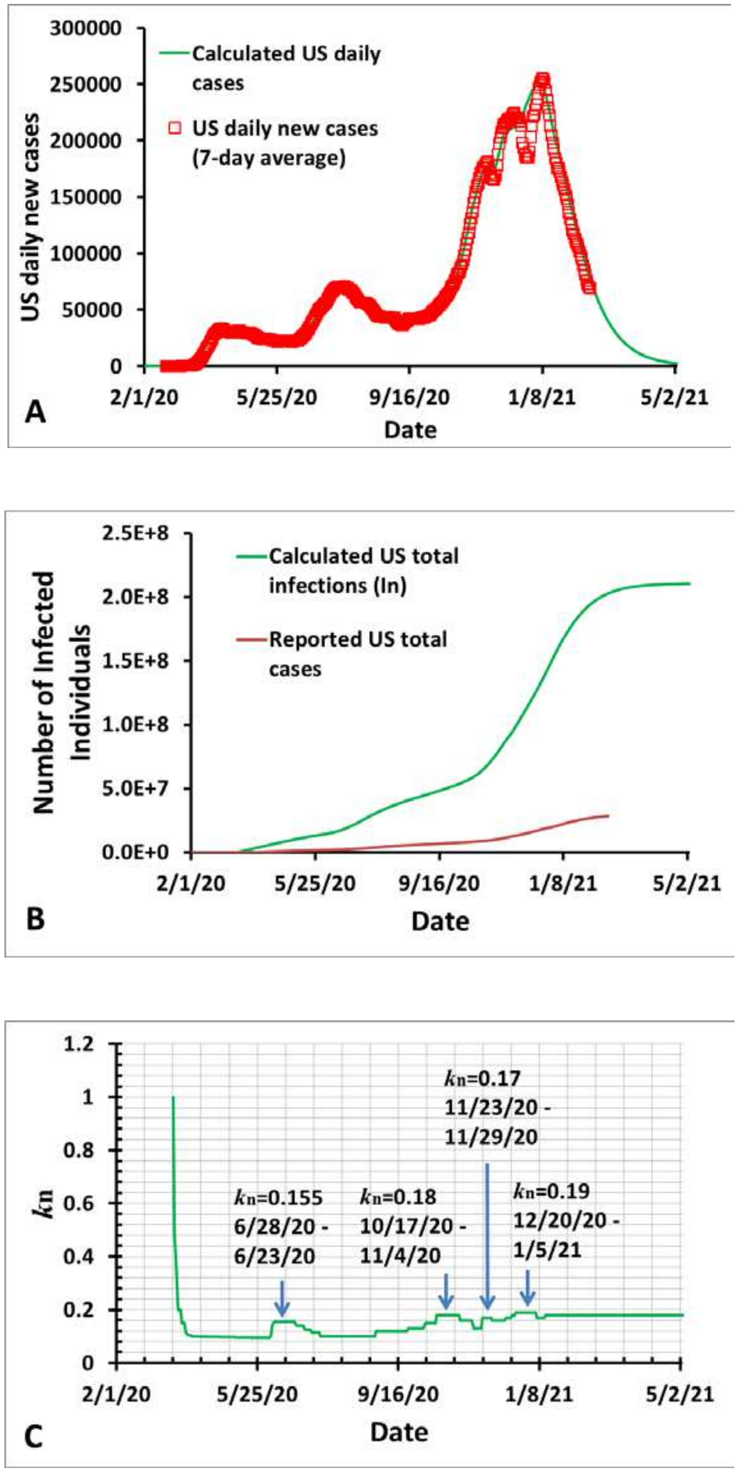


Figure 3. Determination of time-dependent transmission coefficient (k_n) and calculation of the number of US cumulative infected individuals (I_n) by fitting the calculated number of US daily cases (αA_n) to the centered 7-day moving average of reported number of US daily COVID-19 ceases ($\bar{y}_n$).

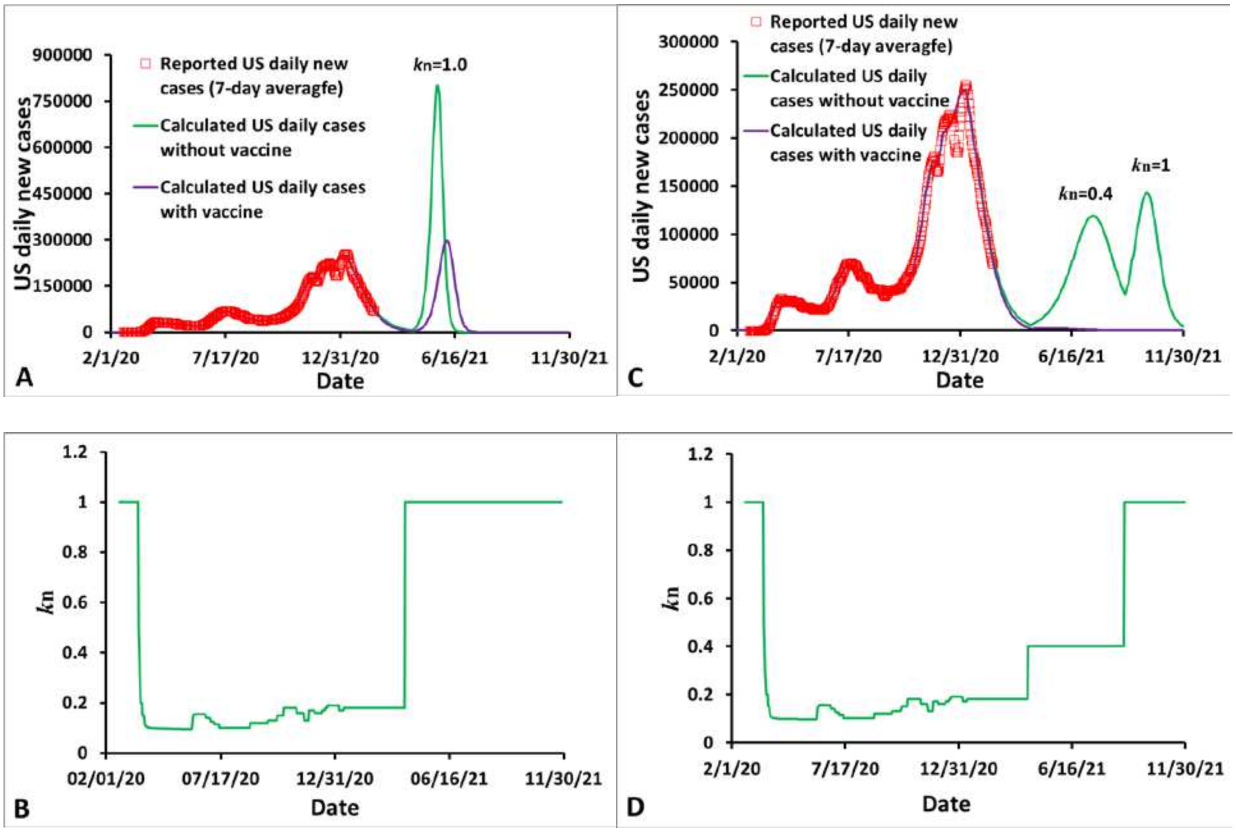


Figure 4. Simulations of future spread of COVID-19 in the United States under two different scenarios.

Figures

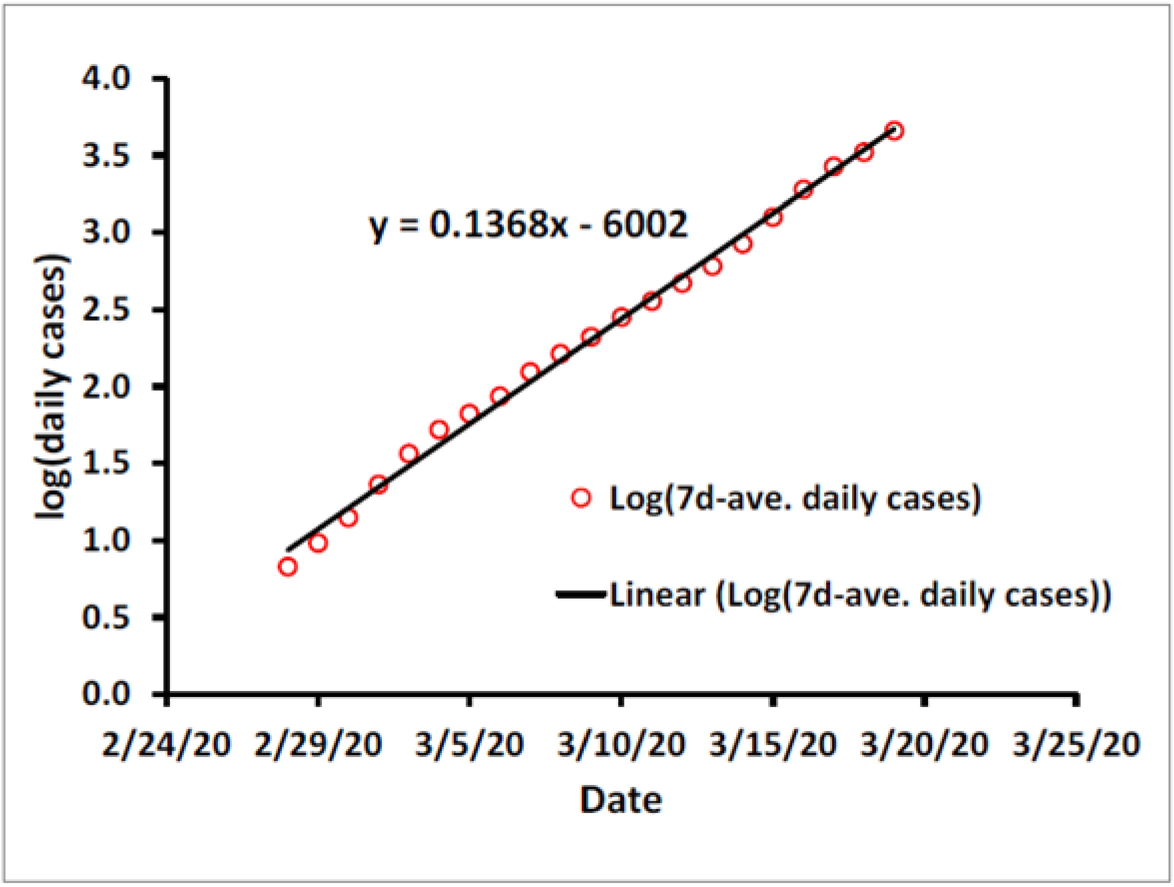


Figure 1

See manuscript for full figure caption.

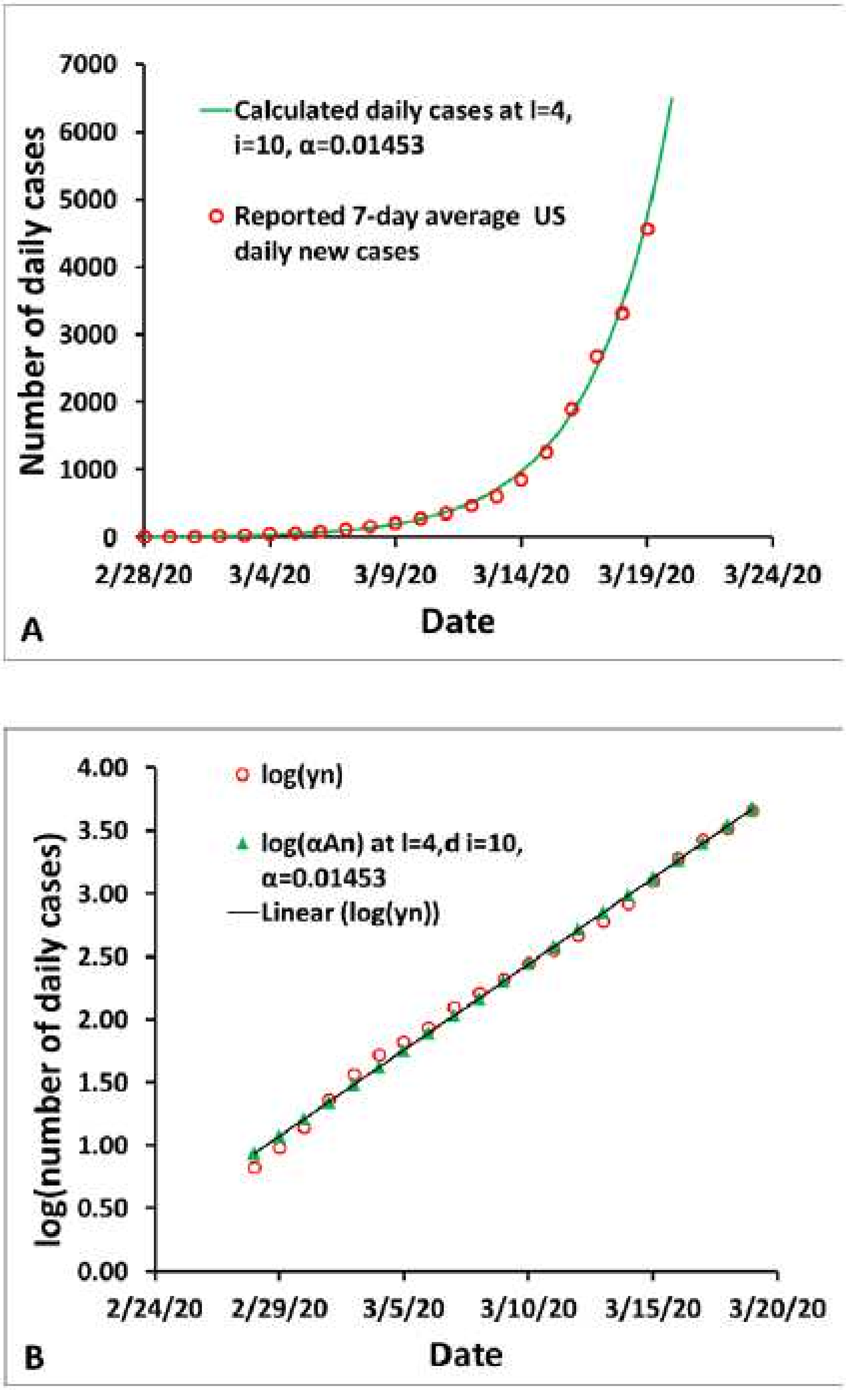


Figure 2

See manuscript for full figure caption.

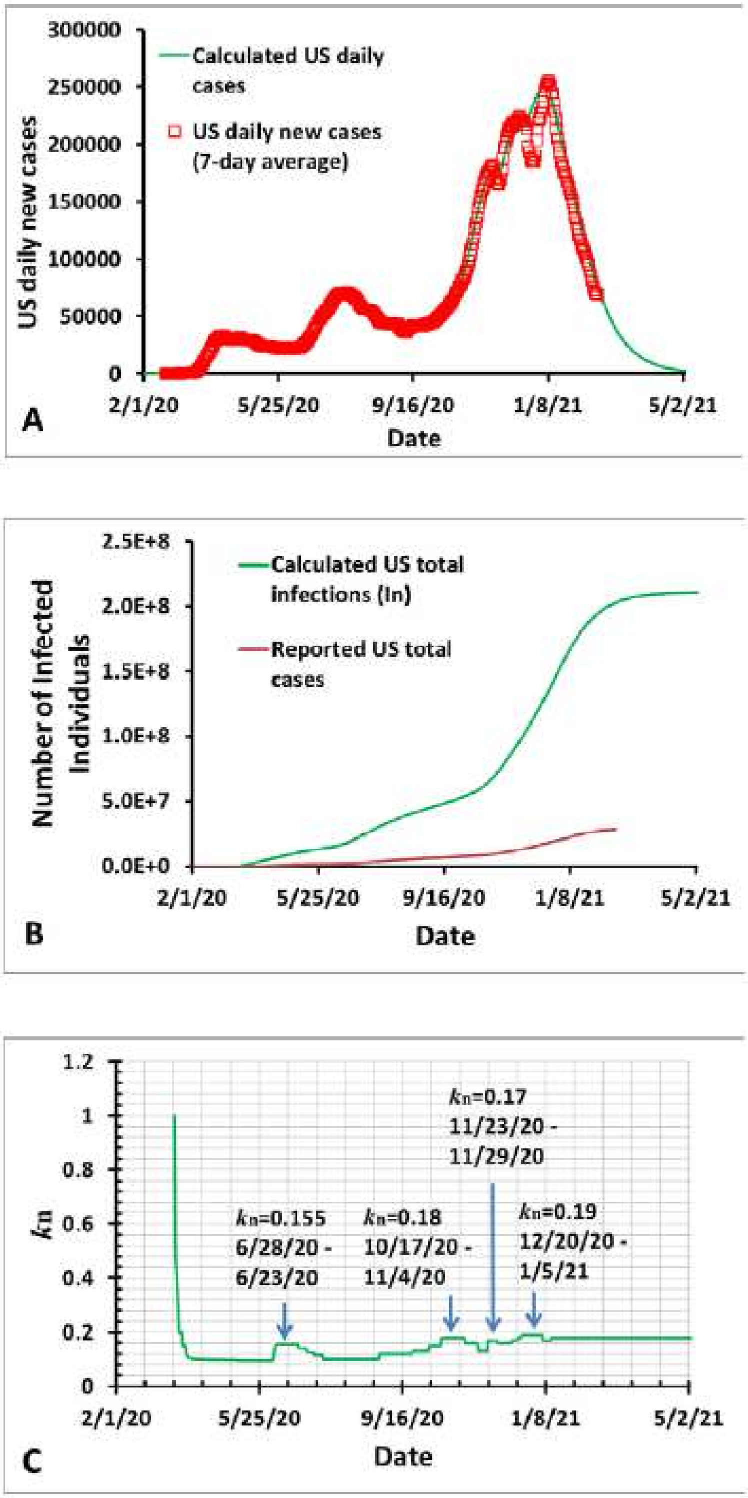


Figure 3

See manuscript for full figure caption.

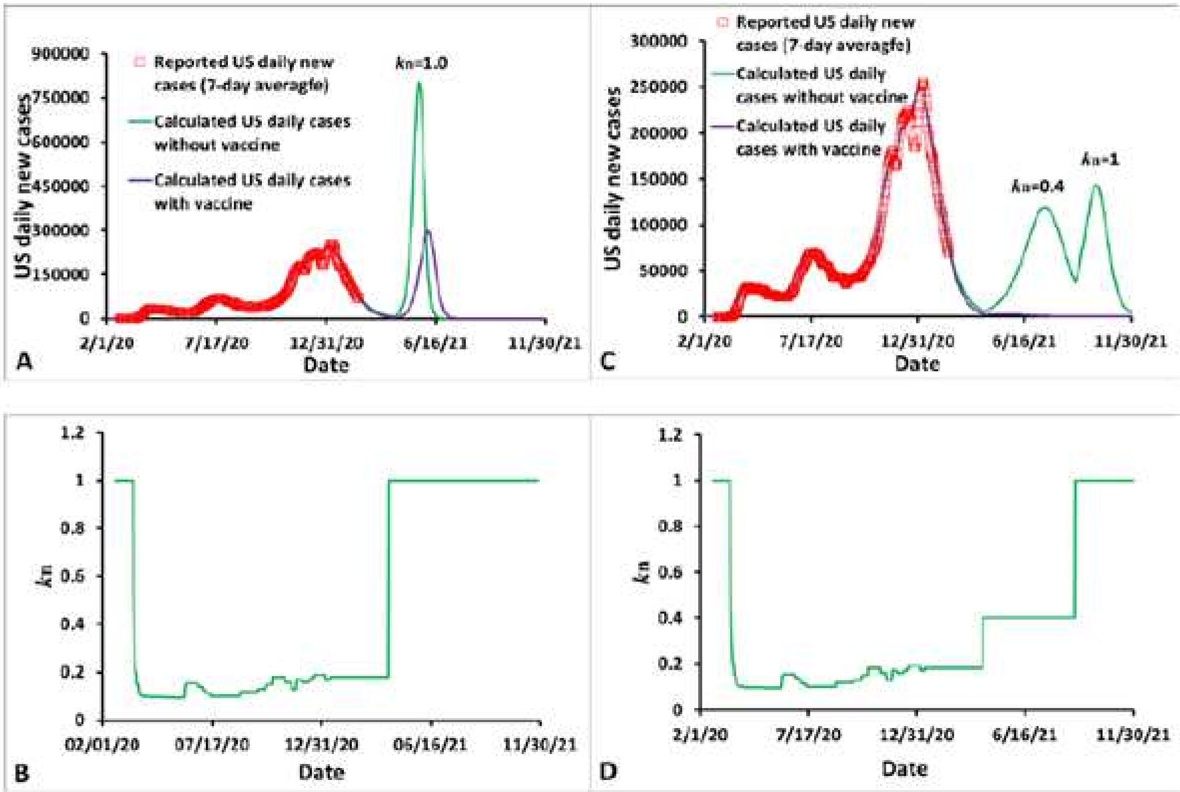


Figure 4

See manuscript for full figure caption.