SIR model for COVID-19

David Hwang 11/20/2024

- Abstract

The COVID-19 pandemic, caused by a novel respiratory virus, brought substantial challenges to global public health systems. In this study, we apply the Susceptible-Infectious-Removed (SIR) model with a difference equations to see how the spread of COVID-19 varies under different conditions, including the basic reproductive number (r_0), the average infectious period (τ), the initial number of infections (I_0), and pre-existing immunity levels achieved through vaccination. The simulations and tests reveal that the dynamics and peak intensity of an epidemic are strongly influenced by r_0, which can be lowered through public health measures such as mask mandates, social distancing, and vaccination campaigns. Lowering r_0 and increasing immunity within the population work together to delay or even prevent outbreaks. The findings also show that having "herd immunity" requires vaccination coverage proportional to $1 - \frac{1}{r_0}$. For instance, with a modestly reduced r_0 to 1.5, the proportion of the population that must be vaccinated to halt the epidemic decreases significantly. This result highlights the combined effectiveness of behavioral changes and immunization efforts. Through all the results, this study emphasizes the critical role of integrated strategies, involving both public health interventions and vaccination, in reducing the burden of infectious diseases like COVID-19.

- Introduction

“Viral respiratory infections are a particularly common class of infectious disease that jump between species relatively often, leading to epidemics in populations with minimal immunity to the novel form. COVID-19 was the most recent of these. With this class of infectious agent, individuals generally develop some degree of immunity after

an infection (although the immunity may wane over time), and thus is it common to model these using a SIR model."(Green 206p) Based on this information, creating a mathematical model, i.e. SIR model, will be an ideal approach to understanding the dynamics of the infectious disease, COVID-19.

SIR model described by a difference equation model:

$$\begin{aligned} \hat{S}_{i+1} &= \hat{S}_i - \beta\hat{S}\hat{I} \\ \hat{I}_{i+1} &= \hat{I}_i + \beta\hat{S}\hat{I} - \gamma\hat{I} \\ \hat{R}_{i+1} &= \hat{R}_i + \gamma\hat{I} \end{aligned}$$

Here, $\widehat{S}$, $\widehat{I}$ and $\widehat{R}$ refer to fractional populations, most commonly reported in the epidemiological literature in units of cases per 100,000. β is a parameter that describes the likelihood of new infections, and γ is the recovery rate. We can express β and γ in terms of the basic reproductive factor, r_0, and the average length of infectivity, τ, as: $\gamma = \frac{1}{\tau}$ and $\beta = \frac{r_0}{\tau}$. For COVID-19, estimates early in the pandemic suggested a basic reproductive factor of about 3.0, with an average length of infectivity of about 10 days.

- Results and Discussion

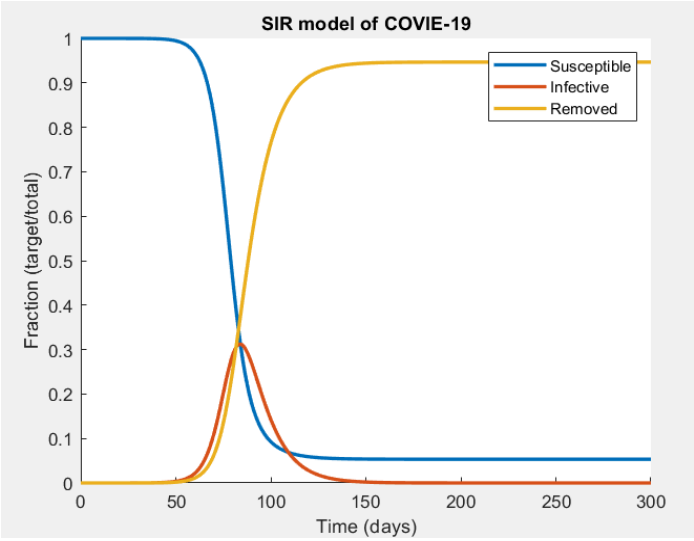
Now we implement our SIR model of COVID-19 in Matlab, using parameters consistent with our understanding early during the COVID-19 pandemic mentioned in the introduction. Now, we simulate the expected disease dynamics for 300 days in the condition of 4 infected individuals arriving in a city of 10 million with no prior immunity, plot $\widehat{S}$, $\widehat{I}$ and $\widehat{R}$ as a function of time on a single graph.

```
%%%%%%%%%%%%%%%%%%%%%%%%
%% SIR model of COVID-19
%%%%%%%%%%%%%%%%%%%%%%%%
% main parameters
r_0 = 3; % basic reproductive factor
tau = 10; % average length of infectivity (days)
gamma = 1/tau; % rate of recovery
```

```
beta = r_0/tau; % likelihood of new infection
inf_0 = 4; % 4 infected individuals were arrived in a city
total_pop = 10000000+inf_0; % total number of population in a city
R_hat0 = 0; % initial fraction of removed individuals
I_hat0 = inf_0/total_pop; % initial fraction of infective individuals
S_hat0 = 1-I_hat0-R_hat0; % initial fraction of susceptible individuals
I_hat = I_hat0; % I_hat is a number
S_hat = S_hat0; % S_hat is a number
R_hat = R_hat0; % R_hat is a number
% simulation
dt=1; % time step: one day
timev=0:dt:300; % time vector for 40 years
sim_S_h=zeros(length(timev),1); % vector to store the fraction of susceptible
sim_I_h=zeros(length(timev),1); % vector to store the fraction of infective
sim_R_h=zeros(length(timev),1); % vector to store the fraction of removed
k=0; %counter
sim_S_h(1)=I_hat;
sim_I_h(1)=I_hat;
sim_R_h(1)=R_hat;
for t=timev
   k=k+1;
   % main equation
   if k~=1
       S_hat = sim_S_h(k-1) - beta*sim_S_h(k-1)*sim_I_h(k-1);
       I_hat = I_hat + beta*sim_S_h(k-1)*sim_I_h(k-1) - gamma*sim_I_h(k-1);
       R_hat = R_hat + gamma*sim_I_h(k-1);
   end
   if S_hat<0 S_hat=0; end;
   if R_hat>1 R_hat=1; end;

   % store value of S&I&R for plotting
   sim_S_h(k)=S_hat;
   sim_I_h(k)=I_hat;
   sim_R_h(k)=R_hat;
end
figure(1); clf
hold on
plot(timev,sim_S_h,'LineWidth', 2)
```

```
plot(timev,sim_I_h,'LineWidth', 2)
plot(timev,sim_R_h,'LineWidth', 2)
hold off
legend('Susceptible','Infective','Removed')
xlabel('Time (days)');ylabel('Fraction (target/total)')
title('SIR model of COVIE-19')
```



We test our model to get more significant insights about the peak (highest total number of currently infected individuals), epidemic, and mortality.

```
% peak time of infected individuals
[max_I,idx_m] = max(sim_I_h);
% active cases per 100,000 people at the peak
max_I*100000;
```

From the graph and our Matlab code above, we can find that the peak (highest total number of currently infected individuals) occurs 84 days after the arrival of the first infected individuals. Multiplying the peak fraction of infected individuals by 100,000, we can get the value for the number of active cases per 100,000 people, around 3.1235e+04.

```
% the highest number of new daily infections
Daily_I = beta*sim_S_h.*sim_I_h; % daily fraction difference
[max_d,idx_d] = max(Daily_I)
max_d*100000
```

From our SIR model, we can notice that $\beta\widehat{S}\widehat{I}$ is the inflow term for the infected fractional population which means the fractional population of the number of new daily infections. Thus, we calculate only this inflow term, $\beta\widehat{S}\widehat{I}$ to find out when the highest number of new daily infections occur. From the list for the daily infection occurrence (beta*S_hat*I_hat), we can get the maximum value, 0.0421 at the 78th index. Thus, the highest number of new daily infections occur between 77 and 78 days after the arrival of the first infected individuals. Also, around 4.2108e+03 people were infected per 100,000 people on that day.

```
% Fraction of infected population by the end of the epidemic
sim_I_h(end)
sim_R_h(end)
sim_S_h(end)
1-sim_S_h(end)
```

We can say the epidemic is finished when the value $\widehat{I}$ is close enough to zero. From the graph of dynamics above, we can roughly say that the epidemic is finished after 200 days. Through the code above we can find 391.75e-09% of the population are currently infected when around the end of the epidemic. If we calculate the total fraction for the number of infections from the beginning to the end of the epidemic including the people in $\widehat{R}$, 94.67% of the population has been infected.

```
% need of hospital care at the peak of the epidemic assumed 10% of
Hospitalization
max(sim_I_h)*0.1*100000 % needed of hospital beds at peak demand per
100,000people
max_d*0.1*100000 % expected hospital admissions at peak growth per
100,000people
```

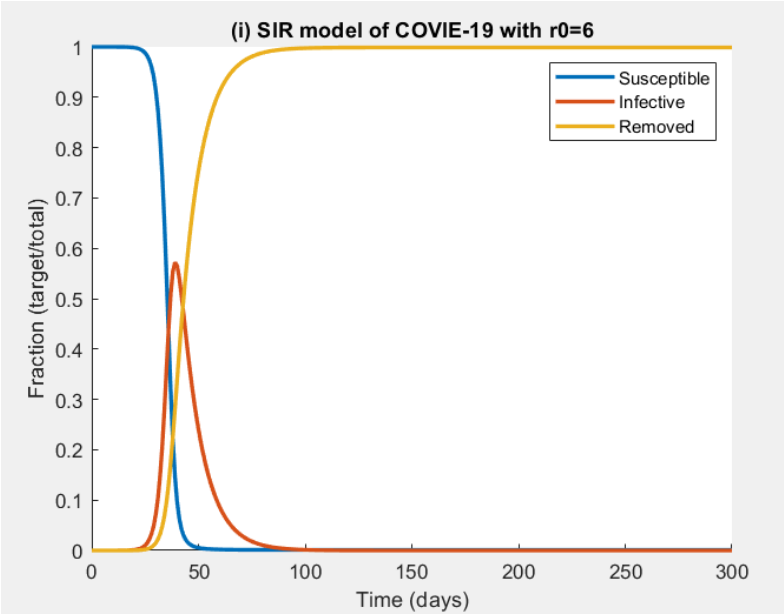
From the result of the code above, we know that around 3.1235e+03 beds would be needed at peak demand(the highest total number of currently infected individuals). When the growth peaked(the highest number of new daily infections), we can expect that 421.0761 people would get hospital admission on that day.

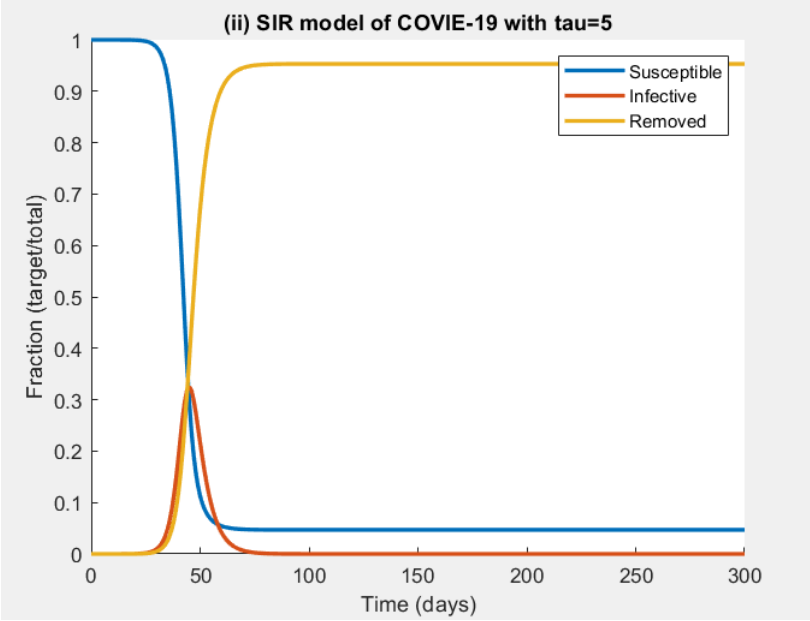
```
% # of people who died by the end of the epidemic assumed that the mortality of
the illness was 1.5%
Death_rate = (I_hat0+sum(Daily_I))*0.015;
Death_rate*total_pop
```

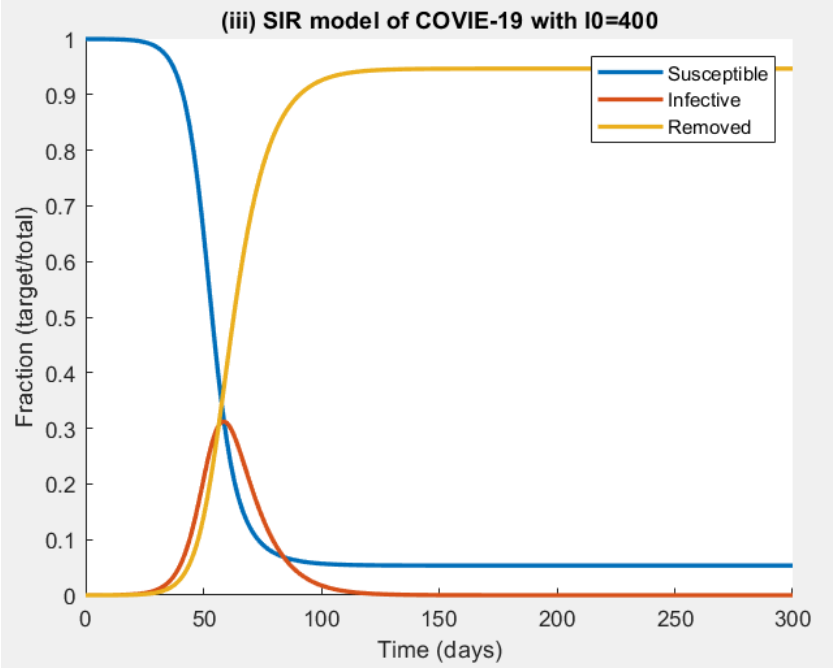
First, we set the mortality of the illness as 1.5% which is consistent with early data for COVID-19. Then from the result of the code, we get that 1.4200e+05 people in the city of 10 million would have died by the end of the epidemic.

From all of our simulations and tests, we know that if we set 10% of COVID-19 cases to require in-hospital care with the same initial conditions, 3123.5 beds would be needed when the total number of currently infected individuals is the highest. However, through the statistical information, even if we ignore the bed occupancy due to existing annual mortality(837 per 100,000), we know that most countries do not have over 1500 hospital beds per 100,000. This means when the total number of currently infected individuals is high, many infected people cannot be accepted to their local hospital because of the lack of beds, so they are likely to go to another hospital in other locations, states, or countries. This will increase the possibility of introducing the disease to the new region. Considering the results of our model and these realities, we have major worries about COVID-19 spreading to new places. Also, the mortality rates that we set for epidemics are generally based on the assumption that adequate treatment will be available. The mortality rates will increase dramatically if health facilities cannot meet the demand. Thus, for the people who eventually cannot go to any hospitals due to the lack of beds and get any adequate treatment, the higher population rate can die than what we expected in our simulation which assumes adequate treatment. This will have a major impact on the increase of mortality in developing countries, in order of the number of hospital beds: the United States, Canada, Japan, and South Korea. Another thing we must consider is that mortality rates existed even before this disease emerged for a variety of reasons. Because there are not enough hospital beds, people who would otherwise need hospital care will not be able to receive it, which will increase the existing annual mortality rate.

At this time, we want to know how r_0, τ, and I_0 affect the course of the epidemic. We repeat our simulation by changing the value of parameters, r_0 and τ, which will effect the change in β and γ. (i) with the basic reproductive factor (r_0) doubled; (ii) with the average infective time (τ) halved; and (iii) with the initial number of cases increased to 400 (I_0= 400).



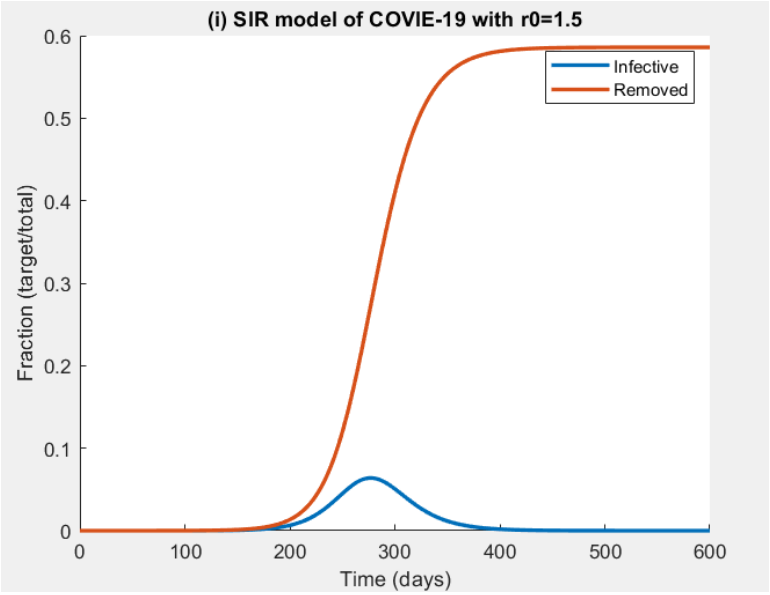


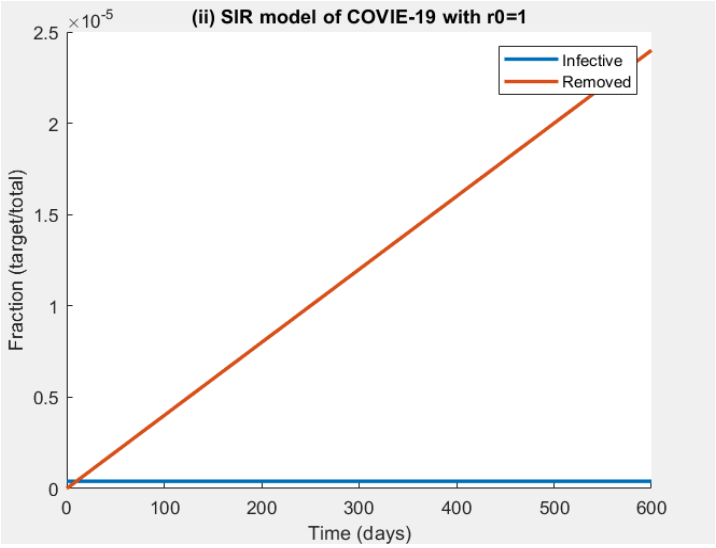


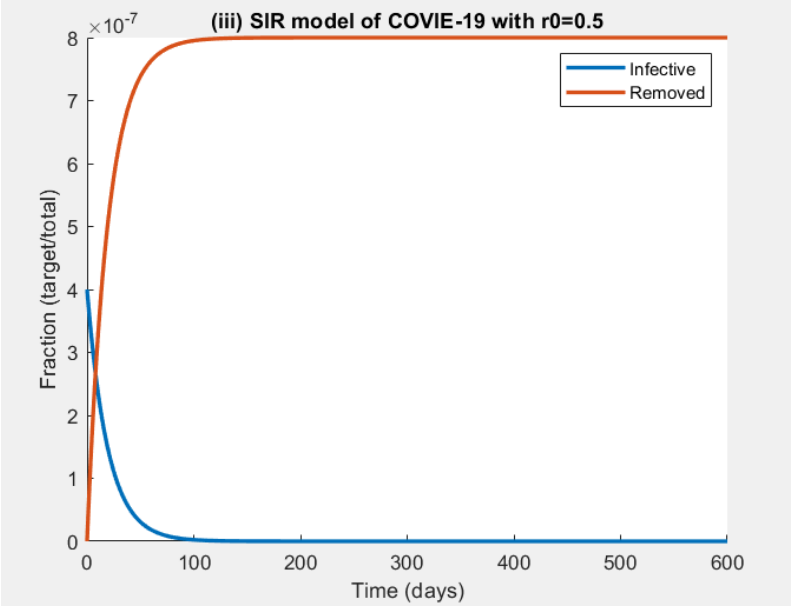
We know that r_0 is the basic reproductive factor and tau is the average length of infectivity. r_0 is proportional to beta. beta is a parameter that describes the likelihood of new infections. Therefore, the probability of new infections is calculated in proportion to r_0, and when r_0 is high, people are more likely to be infected, so the rate of infection spreads faster. Also, through the SIR model, we can see that beta is proportional to the amount of inflow of people newly infected with the epidemic. Therefore, as shown in graph (i), when we increase r_0, it increases the inflow of $\widehat{I}$, which ultimately affects the increase in the maximum value of the highest total number of currently infected individuals during the epidemic. τ is inversely proportional to β and γ. As defined before, β is a parameter that describes the likelihood of new infections, and γ the rate of recovery. Therefore, as τ decreases in (ii), β and γ values increase, which eventually accelerates the course of the epidemic and eventually quickly reaches the end of the epidemic. Since not only β increases but also γ increases, the increase in the maximum value of the highest total number of currently infected individuals is not greatly affected. I_0 is the initial number of infected cases. Since high I_0 accelerates the change of $\widehat{S}$ and $\widehat{R}$, this accelerates the course of the epidemic and eventually quickly reaches the end of

the epidemic. Not only $\beta\widetilde{SI}$ increases but also $\gamma\widehat{I}$ increases, so the increase in the maximum value of the highest total number of currently infected individuals is not greatly affected.

Many responses to managing a potential epidemic effectively reduce the basic reproductive number. We repeat the simulation three more times with different r_0 values to find any information. (i) r_0 = 1.5 (reduced by a factor of 2); (ii) r_0= 1.0 (reduced by a factor of 3); (iii) r_0 = 0.5 (reduced by a factor of 6).



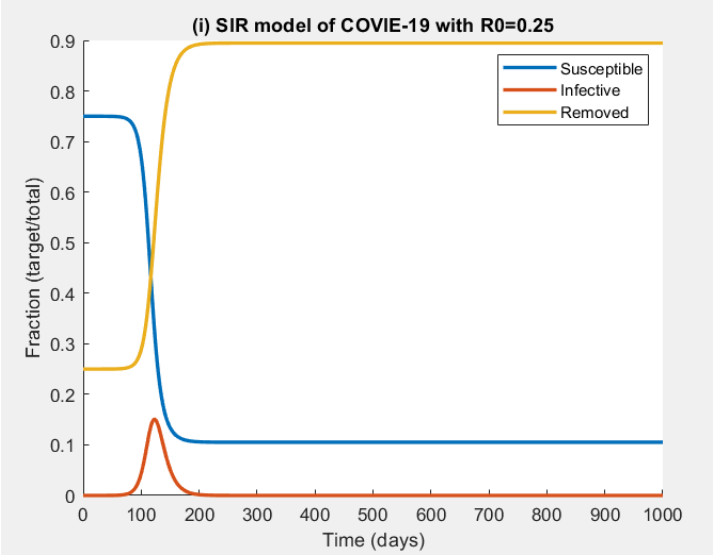


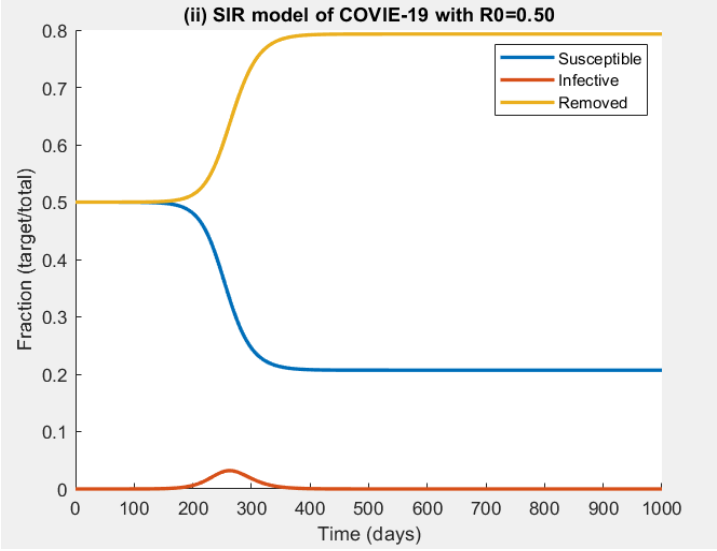


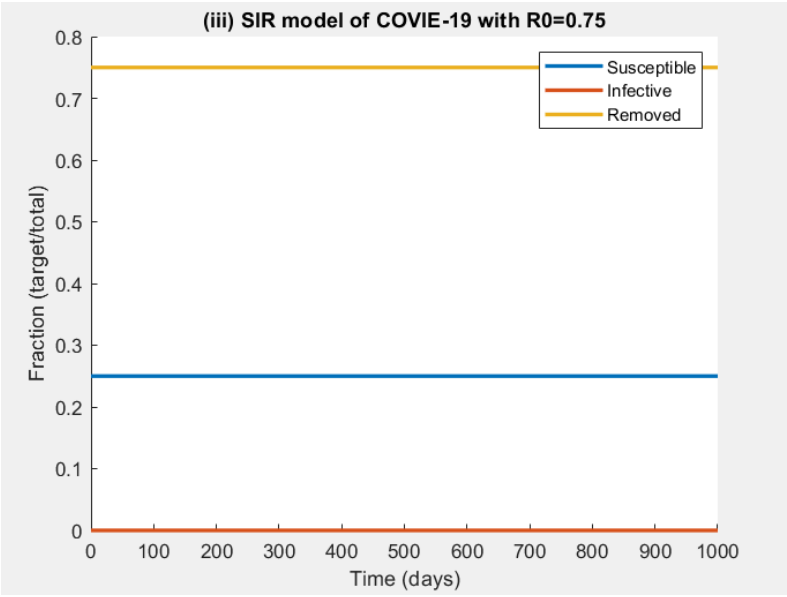
In the previous study questions, the disease dynamics we saw all had an epidemic. The presence of an epidemic means that there is an increase of $\widehat{I}$ followed by a decrease. This time, we conducted an experiment by gradually decreasing r_0 from (i) to (iii), and unlike the previous model, we confirmed that there was no epidemic occurred in models (ii) and (iii) except for (i). In model (i) with r_0 = 1.5, the basic reproductive factor decreased compared to the previous model with r_0 = 3, so β also decreased accordingly, which ultimately decreased the maximum value of the highest total number of currently infected individuals and the speed of the course of the epidemic. However, we can still confirm that an epidemic exists through its graph. On the other hand, in model (ii) with a lower r_0 value than (i), the graph of the dynamics of $\widehat{I}$ is stable, the slope is negative and close to 0, and the graph of $\widehat{R}$ becomes a straight line with a significant slope. Therefore, an epidemic does not exist in this model. In model (iii), the graph of the dynamics of $\widehat{I}$ decreases sharply to 0 due to the lower value of r_0 than in (ii). Accordingly, the graph of $\widehat{R}$ increases sharply. Because $\widehat{I}$ only decreases, there is no epidemic. Through these experiments, we can see that reducing r_0 can reduce the impact of an epidemic (by delaying the peak or reducing the degree of the peak) or

prevent the occurrence of an epidemic altogether. Therefore, we can say that each factor of reduction of r_0 is really helpful in reducing the impact of the infectious disease.

Now we can think about increasing the initial value of $\widehat{R}$ through vaccination. Vaccination (or other sources of prior immunity) can be incredibly effective in helping to prevent infectious disease, but getting a large fraction of the population immunized for a novel infectious disease is a challenge. We repeat the simulation for 1000 days to consider the cases where (i) 25%, (ii) 50%, or (iii) 75% of the population is already immune (whether due to vaccination or past exposure to a similar virus).








```
Daily_I = beta*sim_S_h.*sim_I_h; % daily inflow fraction of I
Daily_R = gamma*sim_I_h; % daily inflow fraction of R
dynmaic_I = Daily_I-Daily_R; % derivative of I
idx_epidemic = find(dynmaic_I<=0, 1, 'first')
```

(i) model with $\widehat{R}$ = 0.25 assigns a smaller value to $\widehat{S}$ than the previous model with $\widehat{R}$ = 0. This is a clear statement since $\widehat{R}$ + $\widehat{S}$ + $\widehat{I}$ =1. This also results in a reduced inflow into $\widehat{I}$. As a result, the size of the peak (the highest total number of currently infected individuals) is smaller by the difference equation of the SIR model, and we can confirm this through the (i) graph. However, we can still confirm that an epidemic occurs in model (i) because the value of $\widehat{I}$ increases and then decreases. (ii) model with $\widehat{R}$ = 0.5 assigns a smaller value to $\widehat{S}$ because the value of $\widehat{R}$ is larger than in (ii). Here too, the inflow of $\widehat{I}$ decreases, and as a result, the size of the peak is smaller. However, we can still see that an epidemic exists because the value of $\widehat{I}$ increases and then decreases. (iii) In the model with $\widehat{R}$ =0.75, we confirmed through the Matlab code above that the value of $\widehat{I}$ continues to decrease without increasing, which indicates that an epidemic

does not occur in this model. Thus, we can say that $\widehat{R}$ =0.75 is the immunity rate that provides “herd immunity”.

We want to derive a minimal vaccination level required for herd immunity in the SIR model analytically. From the SIR model that was introduced in the introduction, we can induce that $\frac{dI}{dt}$ = $\beta\widehat{S}\widehat{I}$ - $\gamma\widehat{I}$. Then $\frac{dI}{dt}$=0= $\beta\widehat{S}\widehat{I}$ - $\gamma\widehat{I}$ = 0 implies $\widehat{I}(\beta\widehat{S}-\gamma)$=0 and this means that $\widehat{I}$=0 or $\beta\widehat{S}-\gamma$=0. If we set $\widehat{S_0} \leq \frac{\gamma}{\beta}$, then the value of $\widehat{S_1}$ will be less than $\frac{\gamma}{\beta}$ and $\widehat{S_i}$ become less than $\frac{\gamma}{\beta}$ for any i. Then this automatically leads to $\frac{dI}{dt}$ = $\beta\widehat{S}\widehat{I}$ - $\gamma\widehat{I}$ = $\widehat{I}(\beta\widehat{S}-\gamma)$ < 0 for all t. And we know that $\frac{dI}{dt}$<0 for all t means that no epidemic occurs in this model. Therefore we can say that $\widehat{S_0} = \frac{\gamma}{\beta}$ is the threshold. The vaccination level: $\widehat{R}$ = $1 - \widehat{S} - \widehat{I}$ and if $\widehat{R} \geq 1 - \frac{\gamma}{\beta}$ is true, $\widehat{S_0} \leq \frac{\gamma}{\beta}$ will always be true. Thus, we can get the threshold of a minimal vaccination level $\widehat{R}$ = $1 - \frac{\gamma}{\beta}$ = $\frac{2}{3}$ required for herd immunity where $\beta=\frac{3}{10}$ and $\gamma=\frac{1}{10}$. From our previous discussion on “herd immunity”, we concluded that when $\widehat{R}$ = 0.5, the model still has an epidemic, and when $\widehat{R}$ = 0.75, the model has no epidemic with “herd immunity“. Since $0.5<\frac{2}{3}$: obtained threshold<0.75 is true we can say that our simulation results above are consistent with this analysis.

If we change the parameter r_0 to 1.5, the threshold will also be changed. Through the same process above, we can get a new threshold of a minimal vaccination level $\widehat{R}$ = $1 - \frac{\gamma}{\beta}$ = $\frac{1}{3}$ required for herd immunity where $\beta=\frac{1.5}{10}$ and $\gamma=\frac{1}{10}$. The results above show that the minimum vaccination level required to prevent an epidemic has been significantly reduced by only slightly reducing r_0. This suggests that we should focus on reducing the reproductive factor r_0 and raising the vaccination level of people

to a critical point in order to prevent the outbreak of an epidemic or at least reduce its impact. The number of infected people initially introduced, I_0, or the average length of infectivity such as 'τ' are factors that are difficult for humans to control. However, r_0 is a controllable factor that can be lowered overall by collectively implementing mask-wearing, social distancing, and group restrictions in the case of COVID-19, a respiratory disease. Additionally, societies can actively encourage vaccination to ensure that populations reach the minimum vaccination levels required for "herd immunity". Therefore, although it may be thought that individual vaccination or mask-wearing is not enough to prevent the outbreak of an epidemic, if we create a culture that supports these behaviors together socially, we can significantly lower the statistical result of the group, r_0, or increase the vaccination level of the group, which will lead us to prevent an epidemic.

- Conclusion

In this study, we have made computational models for COVID-19 dynamics based on the SIR model. The model provides dynamics of COVID-19 over time and through this result we get insights about strategies for minimizing the impacts of epidemics. Initial simulations showed that a city of 10 million could experience a severe epidemic situation with a peak of approximately 31,235 active infections per 100,000 individuals in the absence of prior immunity or public health interventions. Furthermore, if we assume a 10% of cases need to be cared for in a hospital, the demand for hospital beds at the peak would exceed the healthcare capacity available in many regions, highlighting the critical need for proactive healthcare preparedness.

Parameter analysis revealed that increases in the basic reproductive number (r_0) lead to higher and faster peaks in infection rates, while reductions in r_0 slow transmission and, in some cases, prevent outbreaks altogether. Shortening the infectious period (τ) accelerates the epidemic's timeline but has minimal effect on peak infection levels. Conversely, higher initial infection levels make epidemic peaks earlier

without substantially increasing overall severity. These observations underline the importance of public health measures aimed at lowering r_0, such as widespread mask use, social distancing, and restrictions on large gatherings, particularly for respiratory diseases like COVID-19.

A model with initial immunity shows a powerful effect in reducing peak infection levels or preventing outbreaks of the epidemic. Lower r_0 values lead to reduced vaccination thresholds and make the result of combining behavioral interventions with immunization campaigns effective. These findings are consistent with analytical predictions that the vaccination threshold level for "herd immunity" is proportional to $1 - \frac{1}{r_0}$ and supports the importance of robust vaccination efforts.

In the instance of COVID-19, a respiratory condition, mask-wearing, social separation, and group limits can all be used to reduce r0, a controllable element. Societies can also take proactive measures to promote vaccination in order to guarantee that populations attain the bare minimum of vaccinations needed to achieve "herd immunity". Therefore, this study emphasizes that even though it may be believed that individual vaccination or mask-wearing is insufficient to stop an epidemic from spreading, we can prevent an epidemic by establishing a culture that socially supports these behaviors. This is because it will either significantly lower the group's statistical result, r0, or raise the group's vaccination level.

- References

David Green (2024). *Making predictions with a simple SIR model*

I have only used the information inside the homework paper