



**Colorado State University**

**A mathematical attempt to understand, analyze and  
predict the behavior of COVID-19 virus using SIR  
model**

Sohaib Habibollah, Parsa Ghadermazi

## Introduction

Coronavirus disease, or COVID-19, is an infectious disease caused by a newly discovered coronavirus[1], first identified in Wuhan, China, in December 2019[2], and declared as a pandemic in March 11, 2020. Up to this date, November 26 2020, 12.9M cases have been identified as COVID patients in the United States and 60.6M worldwide[3]. This makes the epidemiology of COVID-19 an interesting topic for further investigation.

Epidemiological compartmental models are useful and powerful tools that could provide an insight to the behavior of infectious diseases. These models intent to describe the relationship among sub-population where each sub-population is being represented by a compartment. The simplest model is the SIR model where the population is divided into **S**usceptible, **I**nfected and **R**ecovered. Even though SIR model is the simplest model, one could still gain useful information and assess the gravity of the disease and the efficacy of some protocols.

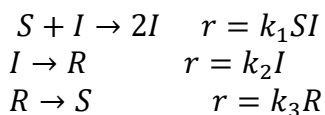
In this project COVID-19 outbreak is modeled using the simple SIR model [4] and a slightly modified version of it [5]. This work will try to simulate distinct scenarios in order to anticipate the behavior of this virus under distinct circumstances. One scenario that could result particularly interesting is the introduction of vaccines. The world is at a critical point since there are two great vaccine candidates that are in phase three of clinical trials. This work aims to answer, how this can affect the behavior of the outbreak, and more importantly, if there is an optimal vaccination rate. In addition, it tries to see if it would be possible to predict when everyone could resume a normal lifestyle.

Answering all these inquiries requires implementation of a broad range of topics discussed during the semester, and the final goal is to infer some real world conclusions from the results of the models.

## SIR model (no vaccine no quarantine)

Although SIR model has its limitations (e.g closed homogeneously mixed community)[6], this model could still provide some insight to what everyone has endured this year. This part of the project intends to shed some light on some of the currently implemented protocols (quarantine/lock down) as well as explaining some of the widely used terminologies such as the concept of Flattening the curve.

In this scenario, we are trying to understand and analyze the behavior of COVID-19 virus using a simple SIR model in the absence of a clinically proven vaccine. In order to do this work in a systemic fashion, one should start by constructing a system of reactions that describes different events :



Where S,I and R respectively represent susceptible, infected and recovered individuals and k's represent the rate in which individuals travel across the three sub-population groups.

Now a system of reactions was constructed, a SIR dynamic model corresponding to this system will have the following form :

$$\frac{dS}{dt} = -k_1SI + k_3R$$

$$\frac{dI}{dt} = k_1SI - k_2I$$

$$\frac{dR}{dt} = k_2I - k_3R$$

To efficiently utilize SIR model to predict a behavior of a virus and to try to predict what could have influenced the behavior of this virus in certain places. with the knowledge that all values must be  $\geq 0$ , one could employ various techniques to determine the optimum values for the three reaction rates constant. In this part of the project we resorted to using data regarding COVID-19 first outbreak in Wuhan [7]. it is worth mentioning that an optimized values using the US data was the group first choice. However, due to the fact that the US data is going wild and following no specific pattern it seemed like a poor candidate for this application. (it hasn't completed a full cycle thus is difficult to fit to a SIR model). A Python code was developed [8] to perform this task using a package called lmfit[9]. The results of this work is shown in figure 1.

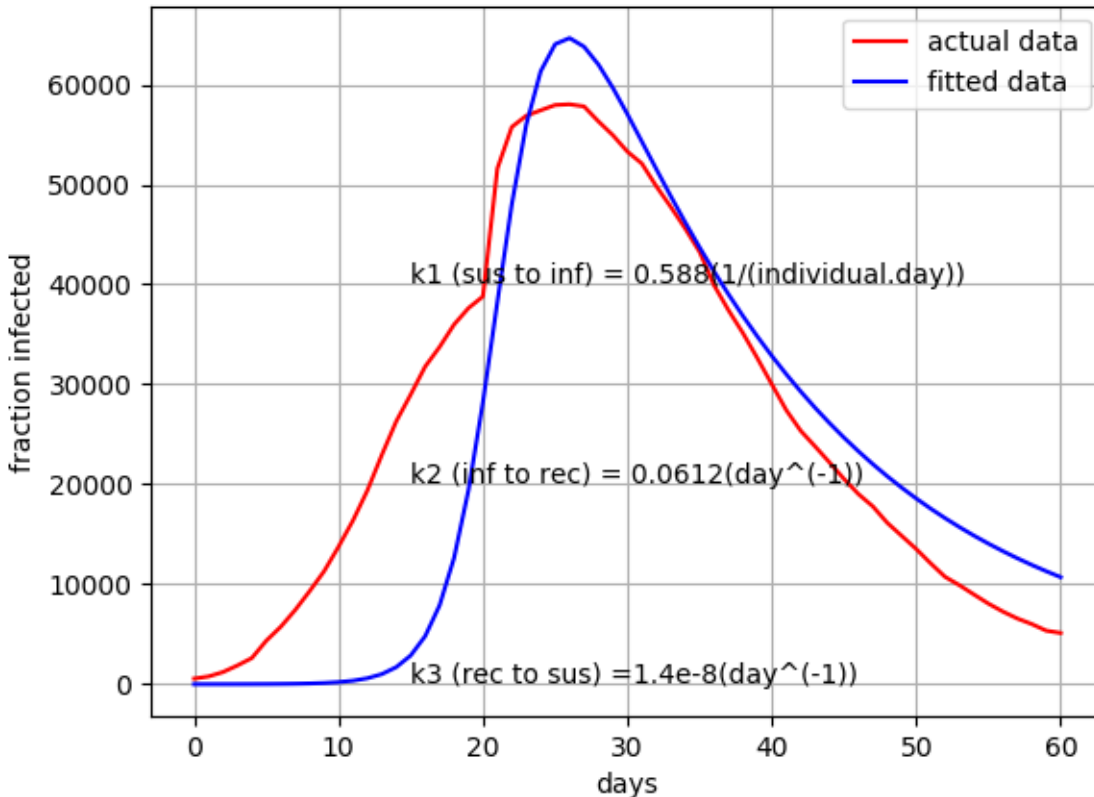


Figure 1 : SIR model parameterization

This set of optimized parameters for COVID-19 (  $k_1=0.58$  ,  $k_2=0.0612$  and  $k_3= 1.4e-8$  almost 0) makes a general since keeping in mind how rapid this virus can spread. it is true that using this set this model was able to recapitulate the peak of the active cases however, the model tends to under-predict

the actual number of active(infected) cases. This could be attributed to the fragility of this model that stems from the long list of assumptions that was made to simplify this model. Another factor that worth mentioning is the effect of human behavior. SIR model treats individuals as a homogeneously mixed system which goes without saying that it is not applicable to humans. Even with all of the holes that were poked in the reliability of the SIR model, one could still use the model as a screaming tool to examine the stability of the steady states of this virus(if any existed) and to assess the efficacy of certain spread controlling protocols.

### Steady states and Stability of steady states:

A crucial piece of information that one need when analyzing the behavior of a virus is to study the stability of the steady states that correspond to the describing system-simple SIR model. To spare readers the tedious lengthy details associated with mathematically finding steady states and assessing their stability, a summary of the solution is shown in Table 1.

Table 1 : Stability of steady states(SIR model no vaccination)

Steady state	Susceptible	Infected	Stability
1	Entire population	No one	Unstable
2	$\frac{k_2}{k_1 N_0}$	$\frac{k_2 k_3 - k_1 k_3 N_0}{k_1 (k_2 + k_3)}$	Stable

### SIR model (quarantine no vaccine):

The first protocol to asses using SIR model is the policy of mass quarantine or lock downs. It is safe to assume that everyone has questioned the benefits of such protocol. Two distinct quarantine scenarios were simulated in order to help establish a protocol that allows for a better control of the virus while minimizing cultural damages (e.g, economy, lifestyles,.etc.) The first scenario is applicable when an approved reliable testing method is not available or is not feasible. This scenario dictates that a random percentage of the population will be placed in quarantine. The goal of such practice is to flatten active cases' curve. Flattening the curve implies squashing the peak of the curve(lower maximum number of infected individuals) and delaying such peak which would reduce the chance of overwhelming the health system of the region in which the outbreak is taking place. The results of this simulation is shown in figure 2. A quick look at figure 1 and figure 2 suggests that ,for the Wuhan's outbreak(peak at 25 days) , if such practice was followed then a partial lock down of approximately 15%(peak at 27 days) was most likely implemented.

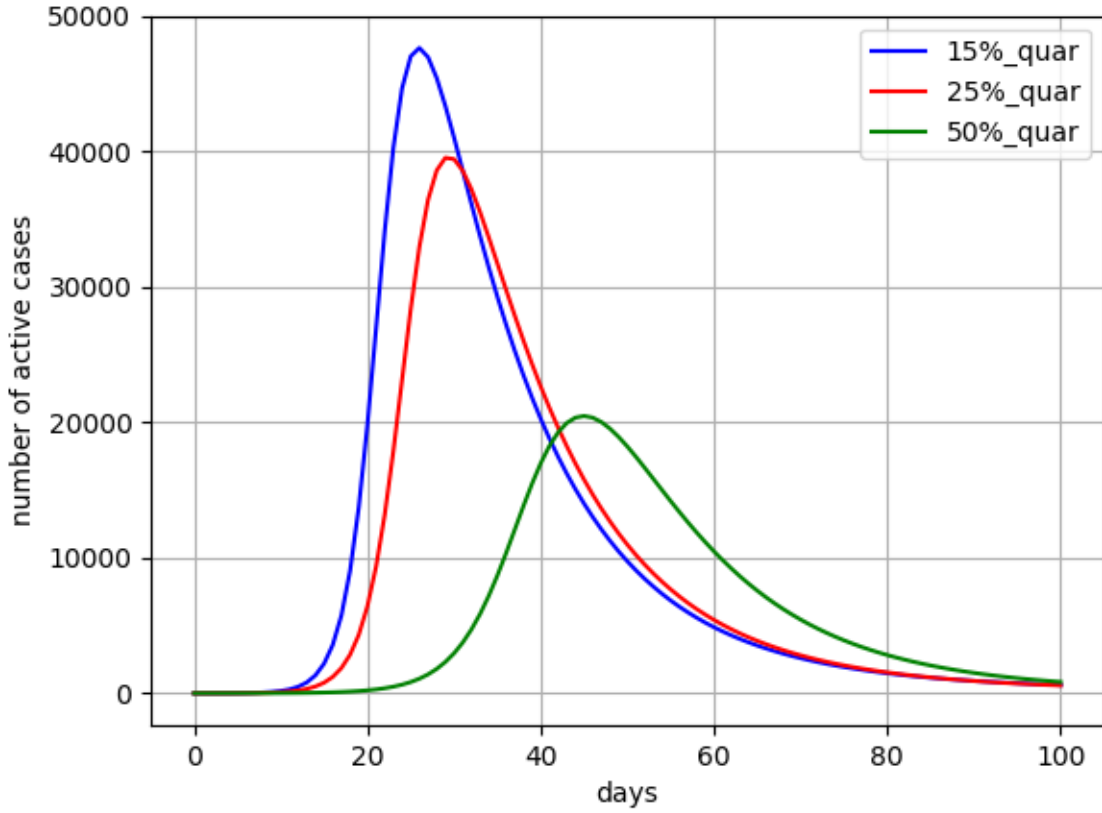


Figure 2 :Effect of random quarantine protocol

The second quarantining option comes with the challenge of having a reliable , affordable and easy to execute method of testing(it also depends on human behavior). It seems like currently there exist a good variation of affordable and accurate testing techniques. Taking this in consideration, SIR model could be modified such that  $\frac{dI}{dt} = \frac{dQ}{dt} + \frac{dUQ}{dt}$  where Q and UQ represent quarantined and unquarantined individuals. SIR system of ODE's will become :

$$\frac{dS}{dt} = -k_1 S U Q + k_3 R$$

$$\frac{dQ}{dt} = f t * k_1 S U Q - k_2 Q$$

$$\frac{dUQ}{dt} = (1 - f t) * k_1 S U Q - k_2 U Q$$

$$\frac{dR}{dt} = k_2 (Q + U Q) - k_3 R$$

where  $f t$  is the fraction of individuals who were tested and then quarantined until recover. Simulating this slightly modified SIR model could provide an insight concerning the efficacy of such protocol especially when compared to random quarantining protocol. The results for this simulation is shown in figure 3

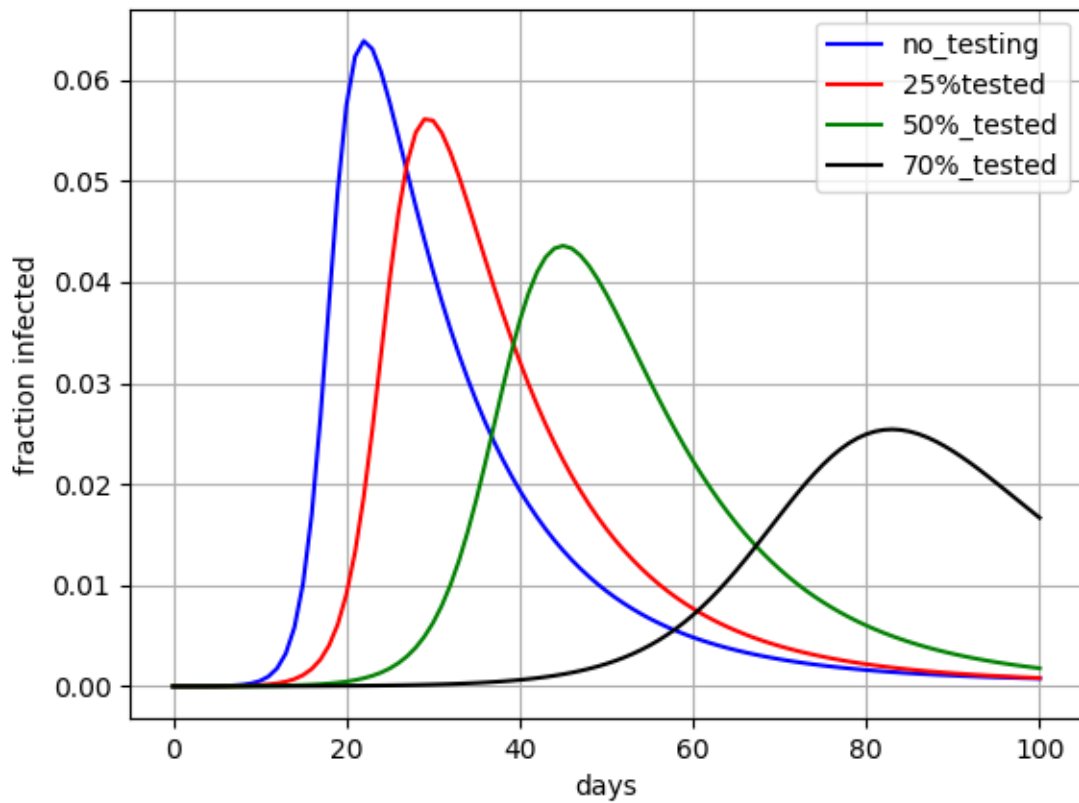


Figure 3: Effect of test\_quarantine protocol

The results shown in figure 3 suggest that, active case's curve could be flattened by implementing test-quarantine protocol. This protocol will serve the purpose of effectively controlling the spread of the virus while minimizing its impact on economy and lifestyles. The biggest challenge associated with this protocol is to come up with an effective yet not overwhelming testing technique and plan.

## Vaccination

Another strategy that can significantly change the behavior of the curve is vaccination. Taking vaccination into account is the same as adding another reaction to the system of reactions which converts the susceptible individuals directly to recovered. The case of China is an interesting one to study for vaccination. First, let's take a look at what the curve looks like right now:

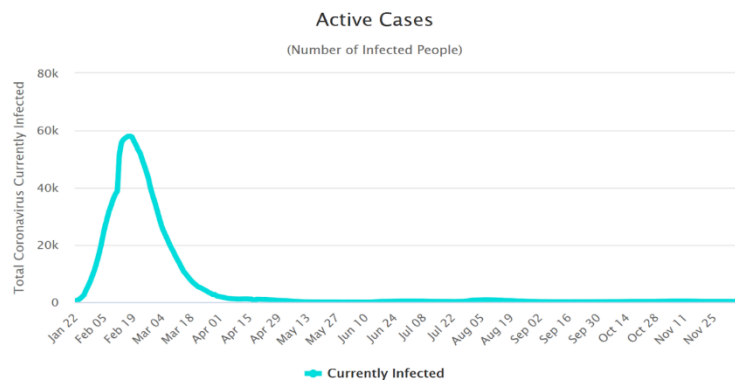
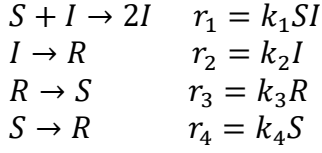


Figure 4: Active cases in Wuhan, China

We see that the number of active cases are close to zero, and we are very close to a number of successful vaccines. So, how would a new outbreak look like in China? If there is no more quarantine in place the system of "reactions" will be as following:



And the governing system of differential equations as a result becomes:

$$\frac{dS}{dt} = -r_1 + r_3 - r_4 = -k_1 SI + k_3 R - k_4 S$$

$$\frac{dI}{dt} = r_1 - r_2 = k_1 SI - k_2 I$$

$$\frac{dR}{dt} = r_2 - r_3 + r_4 = k_2 I - k_3 R + k_4 S$$

$$\frac{d(S + I + R)}{dt} = 0 \rightarrow S + I + R = S_0 + R_0 + I_0$$

With the above relation we can eliminate R:

$$\frac{dS}{dt} = -k_1 SI + k_3(S_0 + R_0 + I_0 - S - I) - k_4 S$$

$$\frac{dI}{dt} = k_1 SI - k_2 I$$

One helpful assumption is that the number of initially resistant individuals is zero which is reasonable because the population of China is in the order of billions while total cases since the beginning is about 100000:

$$\frac{dS}{dt} = -k_1 SI + k_3(S_0 + I_0 - S - I) - k_4 S$$

$$\frac{dI}{dt} = k_1 SI - k_2 I$$

With same reasoning we can say  $S_0 \cong N$ , total population of Wuhan, China.

The value of  $k_4$  will be a parameter that will be changed to see the behavior of the curve.

Now everything is ready for analysis. The above system was solved in MATLAB, and the plot is shown in figure 5:

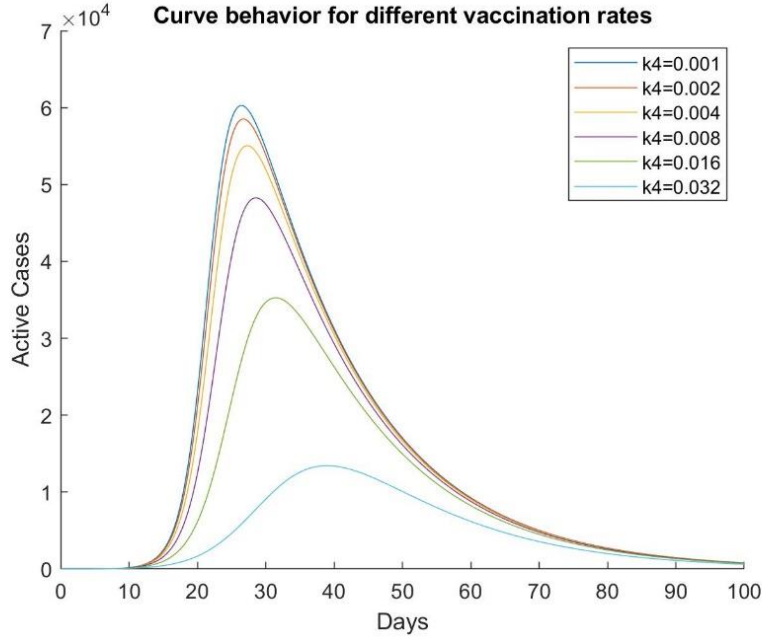


Figure 5: Effect of different vaccination rate constants

It seems from the plot that the steady state is always stable with the current parameters. We can make sure by doing the rigorous mathematical analysis on the matrix describing the system of differential equation:

$$\frac{ds}{d\tau} = -si + \frac{k_3}{k_1 N} (1 - s - i) - \frac{k_4}{k_1 N} s$$

$$\frac{di}{d\tau} = si - \frac{k_2}{k_1 N} i$$

$$\frac{s_0}{N} = K \quad \alpha = \frac{k_4}{k_1 N} \quad \beta = \frac{k_3}{k_1 N} \quad \gamma = \frac{k_2}{k_1 N}$$

Which  $\alpha$  can be considered as vaccination group.

Steady states:

$$i = 0 \rightarrow s = \frac{\beta}{\alpha + \beta} \quad (1)$$

$$s = \gamma \rightarrow i = \frac{-(\alpha + \beta)\gamma + \beta}{\gamma + \beta} \quad (2)$$

We are interested in the first scenario, since our model can predict this close to day 100:

$$F = \begin{bmatrix} -si + \beta(1 - s - i) - \alpha s \\ si - \gamma i \end{bmatrix}, \quad A = \begin{bmatrix} i - \alpha - \beta & -s - \beta \\ i & s - \gamma \end{bmatrix}$$



$$(1) \quad A = \begin{bmatrix} -\alpha - \beta & -s - \beta \\ 0 & \frac{\beta}{\alpha + \beta} - \gamma \end{bmatrix}$$

$$\det(A - rI) = 0 \rightarrow (r + \alpha + \beta) \left( r + \gamma - \frac{\beta}{\alpha + \beta} \right) = 0 \rightarrow r = -(\alpha + \beta) \text{ or } r = \frac{\beta}{\alpha + \beta} - \gamma$$

$\frac{\beta}{\alpha + \beta} - \gamma < 0$  will guarantee the stability of this steady state. Plugging in gives:

$$\frac{\beta}{\alpha + \beta} - \gamma = \frac{k_3}{k_3 + k_4} - \frac{k_2}{k_1 N} \text{ Worst case scenario for stability is when } k_4 = 0 \text{ because it makes } r$$

more positive. As a result, if  $\frac{k_2}{k_1 N}$  is less than one this solution is always stable. Notice that the values reported in fitting section are for dimensionless parameters as a result  $\frac{k_2}{k_1 N}$  is in the order of  $N$ , we normalized by  $k_1 N^2$ , which is always greater than one and this shoes that the steady state is stable.

This is the steady state that our model predicts and we are interested in. Additionally we tried a numerical method to confirm the stability as  $k_4$  changes. For sake of brevity, we do not discuss this simple method here but it can be found on the GitHub repository for this project[8].

We can see the effect of different vaccination rates on the curve. The reason for monotonous behavior of increasing vaccination rate is that the ratio of  $S/(S+I)$  remains very close to one during the 100 days that we are considering. If we had half of the population as active cases, then vaccinating all of them would not be the best idea since it will increase recovered, resistant, individuals at the beginning but they will go back to susceptible state quickly again, and larger  $S$  means higher new cases per day again and can result in a not stable situation. However, for our case this is clearly not the case. One more analysis that would be interesting is to see if vaccine is introduced not in day zero but in other days, then how the curve will look like, with some changes to the script used for the plot in previous page, defining an event function for the ODE solver, we can see the behavior of the curve for different introduction days in figure 6.

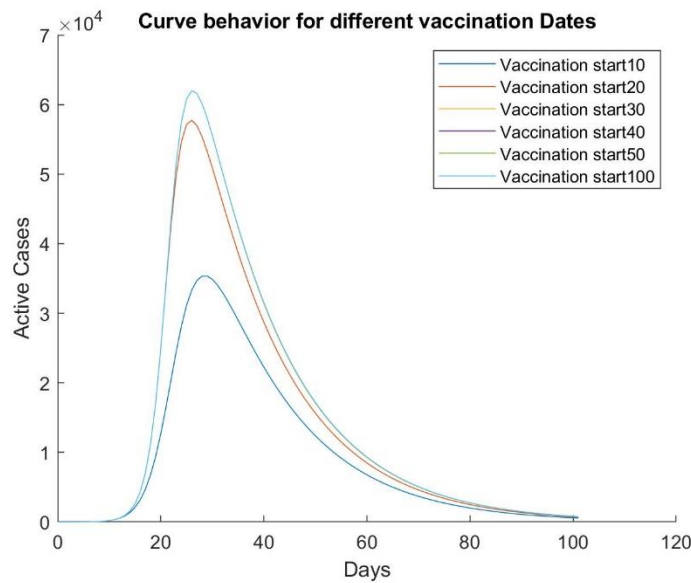


Figure 6: Effect of different vaccination start date

The resulting plot is very interesting. Let's say the peak happens at day 30 when we have no vaccination. This scenario is equivalent to starting the vaccination at day 100 since the active cases are already close to zero. If we are able to vaccinate individuals before this time, it would be effective as it is obvious for starting date 10 and 20. However, if we introduce the vaccine after the peak, 30 days here, it is not effective. As it can be seen in the plot, day 30, 40, 50, 100 all are completely on top of each other that the difference between them cannot be seen easily. This result can be extended to any community:

“As we get closer to the peak, vaccine effectiveness decreases, up to a point that it does not make a significant difference in the curve behavior.”

We postulate that the reason behind this behavior is what is called “herd immunity”. The peak means that maximum number of possible individuals have gotten sick. From this point on, the number of sick individuals is high and these individuals will not become susceptible immediately. So the dynamics of the system is overwhelmed by I and R, and changes in S population due to vaccination will not change the behavior of the curve significantly.

## Conclusion

In this project we tried to use a well behaved dataset, to see what we can learn by applying SIR model to COVID-19 active cases in Wuhan, China. Indeed, many factors are involved that make the analysis challenging. These factors include but not limited to partial lockdowns, change in the behavior of people during the disease, and so on. Despite all these variabilities, some general important conclusions can be made:

- Even partial quarantine can have big effect: It both delays and weakens the peak observed when there is no quarantine and vaccination.
- Quarantining individuals who tested positive can be even more effective. With this it is possible to continue on with normal life to some extent without necessarily facing an exponential increase in active cases. However, this requires that testing should be available immediately so that

quarantining can be done efficiently. This is similar to the policy of Larimer county and Colorado State University to keep campus open with low new daily cases which was successful.

- Vaccination can have the same effect as quarantine with the difference that, at least in this case study, it does not delay the peak, but it successfully dampens it. Another important factor to consider is vaccination start date. As we get closer to peak the effect of vaccination becomes less noticeable. So, given the risks of some candidate vaccines, there might be more effective strategies to lower the number of active cases without putting lives of individuals in more danger.

To apply the method to more complicated datasets, such as the United States, we can not simply use SIR model since there is not simply an exponential growth phase followed by a single peak. In case of the United States, there are three distinct noticeable exponential growth phase. To apply SIR model to US, we recommend splitting the three exponential growth phase and fitting SIR parameters to each and then taking an average. The results can be used to infer different effects such as vaccination, or an approximation of how close are we to the peak or whether the administration of vaccine is feasible or not. It should be noted that SIR model is one of the simplest models available, and it is almost impossible to describe the whole behavior of active cases accurately, but as an initial guess it can be still insightful.

## References

- [1] [https://www.who.int/health-topics/coronavirus#tab=tab\\_1](https://www.who.int/health-topics/coronavirus#tab=tab_1)
- [2]- <https://www.cdc.gov/coronavirus/2019-ncov/cdcresponse/about-COVID-19.html>
- [3]- <https://www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html>
- [4]- Adekola, Hafeez Aderinsayo, et al. “Mathematical Modeling for Infectious Viral Disease: The COVID-19 Perspective.” *Journal of Public Affairs*, 2020, doi:10.1002/pa.2306.
- [5] Anand, N., Sabarinath, A., Geetha, S. *et al.* Predicting the Spread of COVID-19 Using *SIR* Model Augmented to Incorporate Quarantine and Testing. *Trans Indian Natl. Acad. Eng.* **5**, 141–148 (2020). <https://doi.org/10.1007/s41403-020-00151-5>
- [6] Cooper, I., Mondal, A., & Antonopolous, C. G. (2020). A SIR model assumption for the spread of COVID-19 in different communities. *Elsevier Public Health Emergency Collection*. doi:10.1016/j.chaos.2020.110057.
- [7] China. (n.d.). Retrieved December 07, 2020, from <https://www.worldometers.info/coronavirus/country/china/>
- [8] [https://github.com/ParsaGhadermazi/COVID\\_Project.git](https://github.com/ParsaGhadermazi/COVID_Project.git)
- [9] Non-Linear Least-Squares Minimization and Curve-Fitting for Python¶. (n.d.). Retrieved December 07, 2020, from <https://lmfit.github.io/lmfit-py/>