# Mathematical Modelling of SIR for COVID-19 Forecasting

# Muhammad Imran<sup>a,b</sup>, Mengyun Wu<sup>a</sup>, Yun Zhao<sup>c\*</sup>, Emrah Beşe<sup>d</sup>, Wadi Alonazi<sup>e</sup>, Muhammad J. Khan<sup>f</sup>

### Abstract

This paper uses a mathematical model for COVID-19 pandemic forecasting estimation. The continuum of mathematical and statistical models on communicable and noncommunicable diseases has shown a great concern for risk to human lives. The three terms of the SIR model, S (susceptible), I (infectious) and R (recovered), are the main factors of any disease model. This SIR model was introduced in 1927 for forecasting communicable diseases. The SIR model is a simple disease technique by which we can explain mathematically the spread of a virus through a population using mathematical models. The SIR model answers the main three question of this study under specific assumptions. The key parameter in the derivative equations is the value of Q, which is the ratio of contact and the proportion of the total population that comes into contact with a contaminated individual. In the COVID-19 outbreak, this value is very high and the virus is spreading fast. What we see is that, if the value of Q is high, then the disease will spread widely and will result in an epidemic (in our case, it is already at pandemic level). Hence what can we do to reduce the value of Q? This is why currently we are told to wash our hands, because, if we wash our hands, even if we have been in a contact with somebody with the disease, we are much less likely to then become infected. Similarly, social distancing tells us to keep away from people, because if we stay away from other people we are reducing our probability of coming into contact with someone who has the disease and we are therefore contributing to reducing the value Q and controlling the spread. Keywords: COVID-19; SIR Model; Communicable Disease; Pandemic; Social Distancing

### Introduction

The Coronavirus infectious disease known as COVID-19 started in the city of Wuhan, Mainland China, in December 2019. It has now already spread to more than 150 countries across the world. Initially the worst-affected countries have been in the European Union (EU), the United States of America (USA) and mainland China. As a result, normal social and economic life cycles, both in the domestic economy of those affected countries and in their foreign economic relations, have been severely damaged.

\*Corresponding Author: zy570256@gmail.com

We have also seen signs of the exponential growth of COVID-19 cases in South Asian countries, including Pakistan and India. To understand the increase in and sensitivity to COVID-19 cases in the region, this study discusses the estimation of forecasting equations of SIR mathematical models.

Mathematics has long been considered the pillar of sciences. Its applications range from mathematical sciences to biological and social sciences. Recently, statistical and mathematical models have become essential techniques in helping us maximize our understanding of disease spread, transmission and forecasting of infectiousness of communicable and noncommunicable diseases. Mathematical models have been of great interest across the world for diseases such as COVID-19. Chen, Yang and Dai (2020) assess the spread of COVID-19 by estimating the spread of the epidemic and providing recommendations on its prevention and control in the future. The dynamic mathematical modelling of infectious diseases is broadly based on their

<sup>&</sup>lt;sup>a</sup>School of Finance and Economics, Jiangsu University, Zhenjiang, 212013, Jiangsu China

b Department of Management Sciences, Bacha Khan University, Charsadda, KPK Pakistan

<sup>&</sup>lt;sup>c</sup>Department of Economics and Management, Yuncheng University, Yuncheng 44000 Shanxi, , China

<sup>&</sup>lt;sup>d</sup>Institute of Social Sciences, Near East University, Mersin, Turkey <sup>e</sup>College of Business Administration King Saud University, Riyadh 11451, Saudi Arabia

fschool of Economics and Management, Chang'an University, Xi'an Shaanxi China

Muhammad Imran, Mengyun Wu, Yun Zhao, Emrah Beşe, Muhammad J. Khan

compartmental composition. The SIR model consists of susceptibility (S) -infectious (I)-recovered (R) characteristics for infectious disease. Initially, this mathematical technique/model for communicable disease was proposed by Kermack

and McKendrick (1927), who named it SIR. In many instances in an infected population individual recovered or became part of a susceptible population or lost immunity and died (Rajasekar and Pitchaimani, 2019).

The SIR model is a simple disease forecast model and can mathematically explain the spread of a virus in a target population. The history of mathematical modeling of infectious and noninfectious diseases has attracted great attention. The dynamic models of infectious diseases are based mainly on their composition and structure, such as for the SIR characteristics of well-known infectious diseases, such as influenza, swine flu, plague, HIV/AIDS, malaria, etc.

### **Objectives of the Research**

219

In this paper we identify the factors in the SIR model which need to be controlled to stop infection spread. This model can also be used to forecast the expected number of infections for a given period of time. We address the following three research questions:

- 1. How do we calculate COVID-19 spread?
- 2. What equations or model can give us a result for the expected number of individuals who will have COVID-19 at a particular time?
- 3. How do we devise and interpret an equation about how many people will be affected by COVID-19 in total?

To answer these questions, the following process is discussed in the context of the ongoing COVID-19 pandemic. The model tells us that, to minimize the impact of the spread, we need to lower the 'contact ratio' as much as possible – which is exactly what the current social distancing

measures are designed to do. The SIR model can help us estimate different outputs in different

### regions.

According to real-time data (Dong, Du and Gardner, 2020), positive COVID-19 cases are increasing exponentially in most countries around the world. Wuhan was the initial epicenter of COVID-19 which then further spread into Europe and America. At present the USA is the most infected country globally. The forecasting peak of any disease such as the flu, malaria or COVID-19 dissemination plays a key role in policy making efforts to control the spread (Li et al., 2020; Zhou et al., 2019; Zhao et al., 2020; Liu, Gayle and Smith, 2020). The epidemic has not yet ceased. However, effective measures have helped medical services provide passive medical care and helped decrease the death rate. Forecasting the peak can help us understand the growth rate. Perc et al. (2020) report from their research that the growth of COVID-19 should be kept to less than 5% to control the pandemic.

### The SIR Model

The SIR model is based on a differential equation. In the first step of the model, we consider the dependent and independent variables. The independent variable is time t, measured in days. Similarly, we consider two related sets of dependent variables. The first set of dependent variables is the number of people in each of population group. Each of these populations is a function of time, t In conducting this study we are proposing a general model which can be used for any country or region. The population of our target countries is considered those who are 'susceptible'. Mathematically, we express each of these population groups using the parameters Susceptible (S), Infective (I), and Recovered (R), which can be expressed as function of *t* as follows:



Figure 1. SIR Model flow and Rate of Change

Using Susceptible S(t), Infectious I(t), and Recovered R(t), we can develop a simple equation: S(t) + I(t) + R(t) = N

where N is the total population.

For t=0, we need to set up conditions: for example, at time t=0, at the beginning of the epidemic, this can be written as follows:

 $S(0) = S_0$ 

 $I(0) = I_0$ 

R(0) = 0, that is, when there is no infection, the R(0) is Zero (0).

The mathematical equations can be expressed as follows:

$$\frac{dS}{dt} = -\beta SI$$
$$\frac{dI}{dt} = \beta SI - \gamma I$$
$$\frac{dR}{dt} = \gamma I$$

where  $\beta$  is the transmission rate of infection, and <sup>v</sup> is the recovery rate of COVID-19. Those who are recovered once again become part of the susceptible class (S).

#### Assumptions of the SIR model for COVID-19

For all statistical and mathematical models, we must present propositions that simplify real phenomena, because they may be too complicated to express with certain set assumptions. Everything is expressed in a set of simple equations and assumptions. Thus our first assumption is that the epidemic is short enough and does not last for too long. Therefore we assume that the total population remains same.

The second assumption in the SIR model is related to the way the disease spreads: we assume that the incidence of infectious diseases is proportional to the susceptibility and contact between infectious diseases, and we assume that this happens at a constant rate.

Our final assumption relates to the discharge rate. In this category we are going to assume again that there is a constant rate: this could be a death rate or a recovery rate, but again we are going to assume it is constant. The recovered will again become part of the susceptible group (as described in Figure 2) and death can negate the new-born.



220

### Figure 2. Recovered are added into Susceptibles

In our study we are interested in S, I and R, the estimated calculation of the numbers of susceptibles, those infected and the recovered, respectively:

# d(s)

### **SIR Model Derivation**

If we begin with the susceptibles, in Equation 1, according to our assumption, for the number of susceptible people over time we expect from Assumption 2 that this is going to decrease as people become infective and so the rate of change of the number of susceptibles is going to be negative, because it is decreasing the rate of contact (R), which is Assumption 2, and we said that this was proportional to the number of infectives (I) and the number of susceptible, so I and S together symbolize the contact between the number of infectives and number of susceptibles and the (R) is the contact rate between them.

Now for the infectives we have a similar differential equation to consider the rate of change: d(I)

$\overline{\overline{d(t)}}$		
= R I S		
– <i>a I</i>	 	(2)

In Equation 2, we need to understand the rate of change of I over time t and that this will increase as people move from susceptible to infective so that now we have (R, I, S). Hence the term is the same as we had for the first equation for the change rate of susceptibles but now it is an addition because some susceptibles become infected and we also have, under Assumption 3, a situation where infectives recover or die at a constant rate. Thus if someone is an infective, then he/she moves into the third category of recovered (R) or the removed category (death). Here we have minus this constant rate which we are going to equal (a) times the number of infectives (I), and that just leaves the final equation below, which is to say that the rate of change of those removed in the population in Equation 3 must be equal to the gains (al) in Equation 2:

# d(R)

# d(t)

As individuals are removed from the infective class in Equation 2, they move into the recovered (or died) category of Equation 3, and we now have three differential equations for three groups of individuals within the total population. We are familiar with these equations and we need some initial data before we can solve the system of equations. The way we do this is that we define the initial number of susceptible people in the population and say that this is going to be equal to  $\mathbf{S} = \mathbf{S}_{\mathbf{0}}$  (initially).

Then we say that the initial number of infectives will also be specified; we call this (I), and at the very start of the outbreak we do not expect there to be anybody in this removed section: hence

 $I = I_0$  (initially)

Because none of the individuals has yet recovered (or died) from the outbreak, the value of (R) is considered to be zero (0):

### $\mathbf{R} = \mathbf{0}$

Now we have yet to discuss Assumption 1 in the context of our model in the context of our equations. If we return to our model, this says that the population must stay invariable during the pandemic, which in fact means that the rate of change of (susceptibles + infectives + removed) combined must be zero, because our total target population is given by these three factors (S+I+R):

$$\frac{d(s)}{d(t)} = -RIS + \frac{d(I)}{d(t)} = RIS - aI + \frac{d(R)}{d(t)}$$
$$= aI = 0$$

$$\frac{d}{dt}(S+I+R)=0$$

Now, going one step further with this first

assumption, we can solve this equation because we already know the first conditions for the population. If the total population does not change with respect to time, that says that it is the original constant number for all possible values over time. Hence we just take the initial value to be a starting point, which is the value of the population at the beginning. But then, as time passes, this cannot change, because we assume that it has a constant value for the pandemic period. Thus it is always equal to that initial value now that we have formed our differential equations that together make up the SIR model.

In the following section, we need to answer our set of questions.

### How to calculate COVID-19's further spread?

Let us consider our initial value of the infected population given by  $I = I_0$  at the start of the pandemic and we wish to investigate how it will grow, because, if the number of infective people starts to grow, then we have a spread of the disease through a population. Therefore we are interested

to calculate the change in infective numbers over time, as given below:

 $\frac{d(I)}{d(t)} = R I S - a I = 0$ 

But before we do that, we in fact want to start with the  $\frac{d(s)}{d(t)}$  issue, because this indicates that the changing rate of number of susceptibles is equal to a negative value -R I S, because (R) is some positive constant and (I) is a number of infected population as well as of (S). All of these three are positive and the rate of change of (S) is always negative, which tells us that (S) must always be smaller than its initial value and this of course makes complete sense in the context of a disease, because at the beginning of the outbreak everyone in the population in theory is susceptible to the disease, especially with something new like COVID-19 that is unprecedented. Hence susceptibles (S) always decline because their change rate is negative, which tells us that (S) must be less than or equal to its initial value (So):

S≤ So.

Therefore, we can take this value  $(S_0)$  and insert it into our equation:

 $\frac{d(I)}{d(t)} = R I S - a I$ 

Now we have an inequality in our rate of change for the number of infected and we can say that an epidemic will occur if the size of (I) increases from the initial value of  $(I_0)$ .

Hence, to answer our Q1 (will the COVID-19

pandemic spread?), this just comes down to the sign of this constant (r  $S_0$  -a): if this constant is positive, then there will be a spread of the disease (COVID-19):

$$\frac{dI}{dr}$$
 < I (rS0 - a)

This means that if S<sub>0</sub> is greater than  $\frac{a}{r'}$ , then the disease will indeed spread, so the equation can be written as follows:

 $\frac{dI}{dx} < I(rS0 - a) S0 > \frac{a}{r}$ This ratio  $\frac{a}{r}$  is a little easier to consider if we

This ratio  $\frac{1}{r}$  is a little easier to consider if we reverse it and consider what we call Q, where Q is equal to R/a and is called the contact ratio:

$$Q = \frac{\pi}{a}$$

This is the proportion of the population that meets an infected person during the outbreak period and causes an increase in the infection of others. We can rearrange this inequality to get a slightly new version of the same condition for expecting whether or not an epidemic will occur.

If we multiply it by (R) and then divide by (a), this

gives us a new parameter as shown below:

$$R_0 = \frac{r S_0}{r}$$

This is called the basic reproductive value and this condition indicates that we will have an epidemic if this ratio is greater than 1, as shown below:

$$R_{o} = \frac{rS_{o}}{a} > 1$$

In this equation, the number **Ro** (or the basic reproductive ratio) is something that we may have heard about during the prevalence of COVID-19. This number represents the number of secondary infections in the population caused by initial infection.

In other words, if one person has the infection, then the **Ro** value will tell us how many infections on average that person will cause to other community members. It also tells us how many other people will become infected within the target population. During seasonal flu, for example, the value of R<sub>0</sub> is somewhere between 1.5 and 2, whereas for COVID-19 it is estimated to be more like 3 to 4. The exact numbers are obviously still being determined, because this is ongoing, unprecedented outbreak, but the number is certainly much higher than the 1.5 or 2 that we have seen for the seasonal flu in the past. This means that for everyone person infected with COVID-19, they passing it on to three to four people, which is why it is spreading so rapidly.

### What equation or model can tell us the maximum number of people who will have COVID-19 at any one time?

We need to investigate the maximum number of infectives at any given time, because knowledge of such a number is very helpful when it comes to planning how to distribute health resources. That is why we want to create an equation for infectious group (I) that is in terms of various parameters. We know this from within our system of equations, and what we do time in fact is to combine these:

$$\frac{d(s)}{d(t)} = -RIS, \frac{d(I)}{d(t)} = RIS - aI$$

We take these two equations together, because, if we take the differential of I w.r.t. 't'  $\frac{d(I)}{d(t)}$  divided by  $\frac{d(s)}{d(t)}$ , we end up with an equation  $\frac{d(I)}{d(s)}$ . If we simplify this, then the two terms below are going to cancel perfectly, because they are exactly the same: d(I) = R I S, a I = r I S

$$\frac{d(I)}{d(s)} = R I S - a I / -r IS$$

Hence we get a -1 term and then the second

### 222

term here, the (I), will cancel in both terms:  $\frac{d(I)}{d(s)} = \frac{RIS - aI}{-rIS} = -1$ 

Thus we get + (a) over (R) and with an (S) on the bottom, returning to our answer to Question 1:  $\frac{d(I)}{d(s)} = \frac{RIS - aI}{-rIS} = -1 + \frac{a}{rS}$ 

For the spread of the disease we introduced this parameter (Q), which was equal to (R) divided by (a), and so if we rephrase that final term in terms of (Q) we have -1 + (1 divided by (Q) times (S)):

$$\frac{a(I)}{d(s)} = \frac{RIS - aI}{-rIS} = -1 + \frac{1}{qS}$$

And the equation below is something that we can now integrate directly and solve:

$$\frac{d(I)}{d(s)} = -1 + \frac{1}{qS}$$

223

Now we also of course have initial conditions and that is what is going to form the right-hand side of the equation below; hence our final equation is:

$$\mathbf{I} + \mathbf{S} - \frac{1}{q} In\mathbf{S} = \mathbf{I}_{o} + \mathbf{S}_{o} - \frac{1}{q} In\mathbf{S}_{o}$$

Whilst we have this equation for (I) in terms of (S) and the parameters of our model, we have not yet found (I<sub>Max</sub>), the maximum numbers of infected at any given time, which is what we want to know to answer our second question. Now that we are normally thinking about maximums and minimums of functions, we would differentiate the function, but fortunately we already have the derivative from combining Equation (1) and Equation (2).



Thus we can see that this is zero (0) when (S) is equal to 1 or Q, because if (S) is 1 over Q we get minus 1 plus 1 = 0:

 $\frac{d(I)}{d(s)} = -1 + \frac{1}{qS} = 0$ 

So the maximum value of (I  $_{Max}$ ) is in fact in the following equation:

$$\mathbf{I} + \mathbf{S} - \frac{1}{q}In\mathbf{S} = \mathbf{I}_{o} + \mathbf{S}_{o} - \frac{1}{q}In\mathbf{S}_{o}$$

When (S) is equal to 1 over Q, substituting this value into our equation and rearranging for (I), we get the value of (I  $_{Max}$ ), so that below is our final expression for (I  $_{Max}$ ):

$$(I \text{ Max}) = I_0 + S_0 - \frac{1}{q}(1 + In(qS_0))$$

And by writing it this way, we can say specify the maximum number of infectives, so that the answer to the question that we are interested in is that the maximum number of people who will have the disease at a given time is equal to the total target population. Hence, to begin it amounts to everybody, but then we take away something in the value given in the following equation (represented in the box), which turns out to be positive:

$$(I \text{ Max}) = I_0 + S_0 - \frac{1}{q}(1 + In (qS_0))$$

Parameter Q is the contact ratio and  $S_0$  is the total population for a disease like COIVD-19. An outbreak like COVID-19 has happened for the first time and the whole population is considered to be susceptible to COVID-19. Initially,  $S_0$  is some very large, fixed number, but the interesting thing here is what happens as (Q) varies. If we consider this as a function f(x):

$$(I \text{ Max}) = \mathbf{I}_{o} + \mathbf{S}_{o} - \frac{1}{q}(1 + \ln(q\mathbf{S}_{o}))$$

We want to plot 1 over x times 1 plus log of x times some constant value  $S_o$ , and below what this is going to look like is as follows: so (f) of (x) and (x), we have 1 over (X) 1 plus log, and thus we have this increase, initially, and then the function starts to do something like the graph below and this peaks at

around  $S_o$  .

The important factor here is the measure of Q, which is the contact ratio and the fraction of population that comes into contact with an infected individual. Therefore, in the current COVID-19 outbreak, this value of Q is indeed very high, because the disease is very easy to transmit with the result that many people become infected. Finally, for our model, this means that the contact ratio Q is high for COVID-19. Looking at our graph, if Q is large or X on our graph, then F(x) is actually very small, so we are down at the far end of the graph and the value of (F) is indeed very small.



What this means for our maximum number of infective is

f(x)

 $(I \text{ Max}) = I_0 + S_0 - \frac{1}{q}(1 + In(qS_0))$ 

224

F(x) is the maximum number of people infected at any given time and is equal to the total susceptible population minus this function, where our function in fact is now quite small. Therefore, this is an alarming situation for an outbreak that has a large Q value such as with COVID-19. This means that the maximum number of people will become infected.



How to devise and interpret an equation for how many people will be affected by the COVID-19?

To answer this question we need to return to Assumption 1: that the total population in a particular country or region is considered constant. First, we must consider what it means for the end of this disease, because if we want to know the total number of people infected with the disease, we need to end the spread of the disease, which means that the number of infectious diseases must be reduced to zero. So, if we call this point in the future just the end of the outbreak, then what we can do is to look at our total (susceptibles, infectives and recovered) population equation as follows:  $S + I + R = I_0 + S_0$ 

Rearrange to find the size of R and the removed component at the very end of the outbreak, because the number of people who have either caught the disease or died of the disease (i.e., all the people shifted to removed section or R component of the model) will in fact then give the total number of individuals who were infected. What we can do is now to rewrite this equation for what it looks like at the end of the epidemic and this can be presented as follows:

 $R(end) = -S(end)) + I_o + S_o$ 

Thus we have the total population  $I_{o}\,+\,S_{o}$  minus the number of susceptible people left at the

end of the epidemic. To find the value of (S (end)) what we do is in fact return to our equation from Q2, the one that came from integrating  $\frac{d(I)}{d(S)}$  and now we allow time to progress to the end of the epidemic:

$$\mathbf{I} + \mathbf{S} - \frac{1}{q}In\mathbf{S} = \mathbf{I}_{o} + \mathbf{S}_{o} - \frac{1}{q}In\mathbf{S}_{o}$$

Hence we solve the equation below to get the value of the susceptibles left at the end of the pandemic **S** (end):

$$S(end) - \frac{1}{q}InS(end) = I_o + S_o - \frac{1}{q}InS_o$$

And then we replace the value into the equation below to get the number of removed people:

 $\mathbf{R}\left(\mathbf{end}\right) = -\mathbf{S}\left(\mathbf{end}\right) + \mathbf{I}_{\mathbf{o}} + \mathbf{S}_{\mathbf{o}}$ 

This is the number of the removed/recovered population at the end of the pandemic and this is exactly the answer to our third question as to how many people were infected during the pandemic. Now we are going to consider the graph of this function, so that if we plot ( $S_{(end)}$ ) in place of ( $S_0$ ) as the (Y-axis) value in the graph below, and then to the (X-axis), we are going to put (Q), our contact ratio, because as we saw in Q2 the maximum:



The number of infectives was key to controlling the behavior of the disease outbreak, and so in the graph above we replace X by Q, so that what we are plotting in fact is Y minus 1 over X times the log of Y which is equal to a load of constant values minus 1 over X times (again) some other constant. This looks approximately like the following:

It comes through (S (end)) and then flattens at the



(Q).,

Thinking again about the context of the ongoing COVID-19 pandemic, we explain that in Q2 the

value of Q is the contact ratio which is very high, and so for a large value of Q this is going to have a very small value of  $(S_{(end)})$ . This is indeed going to be quite small and that is bad news once again in terms of answering our question, because  $(R_{(end)})$ : the total number of people who catch the disease, remember, is equal to the total population  $I_o + S_o$ and then subtracting of this  $(S_{(end)})$  but for a large value of Q,  $(S_{(end)})$  is small and we are not really subtracting much from the total population,

In summary, the vast majority of the population will catch the disease if the value of Q is sufficiently large.

### SIR Model and Discussion for COVID-19

The contact ratio Q is really key to determining this behavior and we can see this very clearly in the three answers to our important questions about this disease spread.

Q1: Spread  $R_o = qS_o > 1$ 

- Q2: Imax = Total population f(q)
- Q3: Total = total population g(q)

Here we see that if Q is large, then first of all the disease spread translates into a pandemic1. Question 2 tells us that the maximum number of infectives at a given time is equal to everybody minus this function of Q and we saw that the function of Q was actually small for a large value of Q. So, in the case of COVID-19, the maximum number of infectives at any given time is almost equal to the whole population. Then for Question 3, we are interested to investigate how many people will catch the disease again. Q is the contact ratio and so this tells us that basically again most of the population or the vast majority will catch the disease for a large value of Q. Now of course we all know that there is a pandemic and that most people

become infected, but this model tells us the extent, and we should stress that this is one of the most basic disease models. These mathematical models not only tell what may seem obvious, they also tell us how to alter and control events. This can guide us to bringing the situation under control and in our favor. Here what we can see from our simple model is the importance of the contact ratio Q: it appears in the answers to all three of our key questions, so that, while we cannot stop the spread anymore, because that has already happened, what we can do is to look at Q2 and Q3, because we can see that, if we want to reduce the number of people that have the disease at a given time ( $I_{max}$ ), then what we need to do is to make the value of (F) as small as possible, and we saw earlier on our graph that this happened when Q was small and, similarly for Q3, for the total number of people catching the disease we want to make (G) of (Q) as large as possible, so that the total number of people becomes much smaller – and again that happened for a smaller value of Q.

What we can do is reduce the value of Q, the contact ratio, and the fraction of the susceptible population that meets an infective person. This is why currently medical advice is that we wash our hands frequently to avoid infection. Similarly, social distancing is a new measure to keep away from people, because if we stay away from other people we are reducing the probability of coming into contact with somebody who has the infection and this reduces the value of Q.

We need to do everything we can to reduce the value of Q and all the current measures are telling us to do exactly that. These measures ultimately need to lower the value of Q, because the SIR model indicates that the lower the value of Q, the fewer people who catch the disease.

### Conclusion

Mathematical modelling is a a pillar of science, with applications ranging from mathematical sciences to biological and social sciences. Mathematical and statistical modeling has become an essential technique in helping us to improve our understanding of disease transmission and forecasting of infectious communicable and noncommunicable diseases.

This study explains the basis of the SIR model and how we can forecast and initiate measures to control COVID-19. The SIR model is based on differential equations using three terms, Susceptibles (S), Infectious (I) and Recovered (R). This study focuses on answering three question about the possible peak of COVID-19 and how to

reduce the spread by examining it mathematically in the SIR model equation.

### Limitation

This study is limited only to the use of SIR model as forecasting technique for COVID-19. However, we did not estimate forecasting for any country or region during COVID-19 pandemic. In future studies, using the SIR model we can forecast the expected number of suspected and confirm cases for a particular period of time.

225

<sup>1</sup> It is a little late to say this now, because COVID-19 is already a pandemic!

Muhammad Imran, Mengyun Wu, Yun Zhao, Emrah Beşe, Muhammad J. Khan

## Acknowledgement

226

We take this opportunity to thank the editorial board members and anonymous reviewers for reviewing our paper and provide us more insightful comments to make it more beautiful. We also acknowledge the supported by the National Social Science Fund (19BGL127), the National Natural Science Fund (71572071), the China Postdoctoral Science Foundation funded project (2015M571708).

### **Conflict of Interest**

The authors declare no conflict of interest.

### References

- [1] Chen, Z. Yang, J. Dai, B. (2020). Forecast Possible Risk for COVID-19 Epidemic Dissemination under Current Control Strategies in Japan. Int. J. Environ. Res. Public Health, 17(11), 3872. https://doi.org/10.3390/ijerph17113872
- [2] Dong E, Du H, Gardner L. An interactive webbased dashboard to track COVID-19 in real time. Lancet Infect Dis. (2020). doi: 10.1016/S1473-3099(20)30120-1.
- [3] Kermack, W.O., McKendrick, A.G. (1927). Contributions to the mathematical theory of epidemics-I, Proc. R. Soc. Lond. Ser. A 115, 700-721.
- [4] Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. N Engl J Med. (2020) 382:1199–207. doi: 10.1056/NEJMoa20 01316
- [5] Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. J Travel Med. (2020) 27:taaa021. doi: 10.1093/jtm/taaa021
- [6] Perc M, Gorišek Miksic N, Slavinec M ' and Stožer A (2020) Forecasting COVID-19. Front. Phys. 8:127. doi: 10.3389/fphy.2020.00127
- [7] Rajasekar, S.P., Pitchaimani, M. (2019). Qualitative analysis of stochastically perturbed SIRS epidemic model with two viruses, Chaos,

Soliton. Fract. 118, 207-221.

- [8] Zhao S, Lin Q, Ran J, Musa SS, Yang G, Wang W, et al. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: a data-driven analysis in the early phase of the outbreak. Int J Infect Dis. (2020) 92:214–7. doi: 10.1016/j.ijid.2020.01.050
- [9] Zhou T, Liu Q, Yang Z, Liao J, Yang K, Bai W, et al. Preliminary prediction of the basic reproduction number of the Wuhan novel coronavirus 2019nCoV. J Evid Based Med. (2020) 13:3–7.

doi:10.1111/jebm.12376