Evaluating the different control policies for COVID-19 between mainland China and European countries by a mathematical model in the confirmed cases

Feng Lin1, Yi Huang2, Huifang Zhang, Xu He, Yonghua Yin3, Jiaxin Liu4

Institute of blood transfusion
Chinese academy of medical science and Peking union medical college
No.26 Huacai Road, Chengdu, 610052, People’s Republic of China

Abstract: This study focuses on evaluating the different policies of controlling the outbreak of COVID-19 in mainland China and in some European countries. The study is based on mathematical model which is a modified susceptible-infected-recovered (SIR) model. The model takes death and recovery into consideration which in convenience is called the susceptible-infected-recovered-death (SIRD) model. The criterion for the recovered patients is assumed by COVID-19 nucleic acid testing negative. The mathematical model is constructed by retrospective study. Determination of the parameters in the model is based on the epidemic bulletin supplied by the Chinese Center for Disease Control and Prevention (CDC) and National Health Commission of the People’s Republic of China (NHC) from Jan 16 2020 to Mar 5 2020. The data cover the date when the epidemic situation is reported and the data showed that the epidemic situation is almost under control in China. The mathematical model mainly simulates the active cases and the deaths during the outbreak of COVID-19. Then apply the mathematical model to simulate the epidemic situations in Italy and Spain, which are suffering the outbreak of COVID-19 in Europe. The determination of the parameters for the 2 European countries is based on the data supplied by Worldometers. By comparing the difference of the parameters based on the same mathematical model, it is possible to evaluate the different policies in different countries. It turns out that the relatively easing control policies might lead to rapid spread of the disease.

Key words: COVID-19; Dynamics; Mathematical modeling;

1E-mail: linfeng@ibt.pumc.edu.cn
2E-mail: huangy@ibt.pumc.edu.cn, equal contribution with the first author
3E-mail: yhyin.ibt@gmail.com
4E-mail: liujiaxin@ibt.pumc.edu.cn, corresponding author
1 Introduction

The current outbreak of coronavirus disease (COVID-19) was first reported from Wuhan, China, on 31 December 2019 [1]. When the confirmed cases in more than 20 countries worldwide raised concern on an international scale, it is closely monitored by researchers and public alike [2–5]. According to the infection, the data of COVID-19 supplied by the Chinese Center for Disease Control and Prevention (CDC) and National Health Commission of the People’s Republic of China (NHC) in Mainland China, the outbreak of COVID-19 in mainland China is almost under control, while in European countries the epidemic situation is severe as the number of deaths increases rapidly [6].

In January the disease spread very fast in China. According to the information supplied by NHC and the Chinese CDC, the number of confirmed cases was 45 on Jan 16 2020, which increased to more than 80000 on Mar 5 2020. As the increasing rate of the active cases is close to 0 in March, it could be deduced that the epidemic situation is almost under control. However the epidemic situation in European countries was severe in March. As there was a relatively complete process from epidemic outbreak to control in China, it could supply a model by retrospective study. A mathematical model is proposed by modified susceptible-infected-recovered (SIR) model, which for convenience is named as susceptible-infected-recovered-death (SIRD) model. The determination of parameters relies on case number data provided by the Chinese CDC which updates the number of active cases both in China and the global per day [7] and NHC which mainly updates the number of the active cases, the close contacts and the recovered cases in China [8]. Using the mathematical model proposed by the epidemic situation in China, the epidemic situation in European countries could also be estimated in the same way. The determination of the parameters in European countries relies on the data from Worldometer [6]. In this study 2 European countries Italy and Spain where the epidemic situation is quite severe are chosen.

The SIR model is an effective model to predict the growth of the confirmed cases [9]. The parameters in the SIR model are usually constant [2–5]. And the mortality is not taken into consideration. So far, the researchers have already taken banning of traffic into consideration. By retrospective study, it is found that the parameters are actually time-dependent. And the mortality may be an important variable during the outbreak of the disease. Even though the death is much less than the infected population, it needs to be considered. In the traditional SIR model, this kind of information is not taken into consideration. With these modifications, a new model proposed in this study is called the Susceptible-Infected-Recovered-Death (SIRD) model.

The mathematical model in this study only contains 4 kinds of variables: susceptible group $S(t)$, infected group $I(t)$ (the confirmed cases), recovered group $R(t)$ (the nucleic testing negative after infection) and the mortality group $D(t)$. The transportation control is not involved in this model. The mechanism only reflects the spread of disease in the healthy population. The result in this model could be compared with the transportation control which is the actual result.
2 Epidemiological modeling

2.1 Epidemiological modeling based on retrospective study of China

Before introducing the mathematical model, the observed results are shown in figure 1. The sub-figures are (a), (b), (c), and (d). The observation of the confirmed cases, the death, the close contacts and the recovered cases lasts for 50 days. It began from Jan 16 2020 to Mar 5 2020. Figure (a) and figure (b) imply that the growth of the close contacts and the confirmed cases slowed down on the 30th day. And figure (d) implies that the growth of the deaths slowed down on the 40th day. According to the data, rapid increase of the confirmed cases implies that the growth of it might be in the exponential form [2].

Comparing figure (a) with figure (b), it implies that the close contacts are actually the susceptible population rather than the total population in China or in Hubei province. The transportation policies in China [10–15] is proposed for controlling the outbreak of the disease. If some one was traveling by train as was diagnosed, all the people in that train should be confined in an isolated area. This means that all the people in that train were assumed to be susceptible one, while people who are not in that trip are not regarded as the close contacts. Other policies should also been taken into consideration, such as active cases were either quarantined or put under a form of self-quarantine at home [2]. The suspicious cases were also either quarantined or put under a form of self-quarantine [2]. The policies imply that the number of the susceptible population only depends on the cumulated number of the close contacts.

In this work, the variable $S(t)$ which denotes the number of the susceptible cases in SIR model is related to the number of the close contacts. By retrospective study, the total number of the close contacts is $6.9 \times 10^5$. However, as the epidemic situation is changeable, it is difficult to give the exact number of close contacts, which are the susceptible population. Moreover the number of the close contacts is increasing day
by day, while the infected number of population is growing. So it could be deduced that in fact the exact number of the close contacts is unknown. For these reasons, the mechanism of infection-recovering-death is given by (1)-(4):

\[
\begin{align*}
\frac{dS(t)}{dt} &= -r_1 I(t)S(t) + r_2 I(t) \quad (1) \\
\frac{dI(t)}{dt} &= r_1 I(t)S(t) - \mu I(t) - \delta I(t) + \rho R(t) \quad (2) \\
\frac{dR(t)}{dt} &= \mu I(t) - \rho R(t) \quad (3) \\
\frac{dD(t)}{dt} &= \delta I(t) \quad (4)
\end{align*}
\]

formula (1) means that the number of the susceptible population consists of 2 parts: the number of confirmed cases would increase the number of the susceptible population by the rate of \(r_2\), and the susceptible population is converted into the confirmed cases when they are in contact with the confirmed cases with the rate \(r_1\). When the confirmed cases were in contact with those healthy ones, the healthy ones might be transmitted into the infected cases. The confirmed cases are described as variable \(I(t)\), by formula (2). According to the nucleic testing, the recovered cases could be defined by the infected cases whose nucleic testing negative. In this model, it is assumed that the recovered population with nucleic acid testing positive again might change into the active cases, as it is reported that some recovered COVID-19 patients appeared nucleic acid testing positive \([16]\). The increase of \(I(t)\) depends on changing the susceptible cases into the active cases and the recovered population with nucleic acid testing positive. The recovery rate \(\mu\) and the death rate \(\delta\) eliminate the number of the active cases. The number of the recovered patients depends on the variable \(I(t)\) in equation (3). The number of the death is described in equation (4). The transmission coefficients would be estimated by the observed results.

2.2 Determination of the transmission coefficient

The determination of the transmission coefficients mainly reported in appendix. In this section we only give an introduction of determining the transmission rate. The transmission rate \(r_1\) calculated by the observed results is shown in figure 2(a) and the death transmission rate \(\delta\) calculated by observed results in figure 2(b).
The infection coefficient is figured out in this way: divide the number of the diagnosed people by the number of the close contacts every day. The death coefficient is figured out in this way: divide the number of the death by the number of the recovered patients every day.

According to the statistical character of $r_1$, its mean value is 0.143. The shape of its distribution function skew-right (skewness $> 0$), and the distribution function is "thinner" than normal distribution (kurtosis $> 3$). So the most probable number of $r_1$ is larger than 0.143. In this work we choose $r_1 = 0.145$. The detailed approach of determining the formula of the transmission rates would be shown in appendix. Similarly, the recovery parameter is also time-dependent. The details of determination of the time-dependent function of $\delta(t)$ and $\mu(t)$ are shown in appendix. There only show the results. $\delta(t)$ is chosen in the formula as (5).

$$\delta(t) = 0.0012 + 0.00250 \times 1.1^{-t}$$

$\mu(t)$ is given in the formula (6)

$$\mu(t) = 0.00003t^2$$

The parameter $\rho = 0.005$, as the incidence of the recovered patients with nucleic acid testing positive is very low. $r_2$ is to be set according to the observed number. $r_2$ must let the peak number of the confirmed cases similar to the observed ones.

2.3 Prediction of epidemic situation in European countries

As the epidemic situations in European countries is changeable, the model (1)-(4) is also used for prediction. The curing rate is assumed to be the same, as there is no reason to assume the medical resources in European countries are better or worse than it in China. The difference stems from the medical resources management, which are concerned with $\delta$, $r_1$ and $r_2$. The data have been in collection since Feb 15 2020 [6]. According to the observed data, the parameters $r_1$ and $r_2$ should be adjusted. The death rate needs to be re-estimated as it reflects the management of medical resources in European countries.
Lack of medical resources is also reflected in European countries. Comparing with the cases in China, it might be more severe in European countries. The death rate of Italy and Spain is estimated as in the formula (7):

\[ \delta(t) = 0.010 + 0.00250 \times 1.1^{-t} \] (7)

the formula (7) is estimated in the same way as that was done for China. It is obvious that the death rate is larger in European countries than that in China.

3 Results

3.1 Death rate in China, in Hubei province China, in Italy and in Spain

Hubei province contains the largest number of the infected population in China. And most of the confirmed cases are in Hubei province. This is the result of the policies that confining the migration of the people during the outbreak of the COVID-19. The tendency of the death in Hubei province is consistent with the death in mainland China (figure 3).

Figure 3: The red points are the death coefficients in Hubei province and the dark blue points are the death coefficients in China.

Figure 3 implies that the evolution of the SIRD model might show the same results in Hubei province. And it could be deduced that if the epidemic situation shows the outbreak, the tendency of the death coefficients are the same in both China and Hubei province China. As lack of experience and medical resources at the beginning, he death coefficient might be high. Then the death coefficient tends to a stable constant.

Lack of medical resources is also reflected in European countries. Comparing with the case in China, it might be more severe in European countries. The death rate of Italy and Spain is estimated as in the formula (7). It is obvious that the death rate is larger in European countries than in China. The fitting line is shown in figures 4(a)-4(b).
Figure 4: The death rate figured from the observed data was in dots, and the simulated death rate is drawn in blue line. As at the beginning of the 18 days, there were no reported deaths in Spain, the estimation of Spain began from 3th Mar 2020.

3.2 The simulated result of China

At initial time $t_0$, the number of the recovered patients is $R(t_0) = 2$, the number of active cases is $I(t_0) = 45$. The number of the susceptible population is $S(t_0) = 763$. By numerical solving equations (1)-(4), the simulation of the number of $S(t), I(t), R(t)$ and $D(t)$ is shown in figure 5. $S(t)$ is the black dashed line in figure 5(a). The line is not clear in figure 5(a). As the number of the susceptible population is changeable because of the number of the active cases, and it is not important in this study, it is not shown independently in figure 5. $R(t)$ is the brown dashed line in figure 5(a). The reason why it is much larger than the active cases is that the number might be recounted. To show $I(t)$ and $D(t)$ clearly, $I(t)$ is shown in figure 5(b) and $D(t)$ is shown in figure 5(c).

Figure 5: The simulated COVID-19 infected population in China. The blue line (fig. 5(a) and fig 5(b)) is the number of the population that infected by the disease. The brown line (fig. 5(a)) is the population that are recovered. And the dashed red line is the number of the death (fig. 5(a) and fig 5(c)). The parameters are $r_1 = 0.145, r_2 = 0.107, \rho = 0.005$. $\mu(t)$ and $\delta(t)$ are referred to formular (6) and formula (5). The initial conditions are $S(t_0) = 763, I(t_0) = 45, R(t_0) = 2, D(t_0) = 1.
3.2.1 Comparing with the actual results

Comparing the simulated results with the actual results would reflect the infection under 2 kinds of cases: one is the ad libitum transportation, the other is the transportation under control. It has already been mentioned in the previous introduction, the mechanism of this model did not take the transportation control into consideration. So this result could reflect how the infection would evolve without transportation control. Figure 6 shows the comparison of the simulated number with the actual number supplied by the Chinese CDC. Figure 6(a) gives the comparison of the number of the infected population every day in mainland China with the simulated result. And figure 6(b) shows the comparison of the number of the death in mainland China (blue dots), Hubei province China (red dots) and the simulated result. In fig.6(a), the data of the active cases in Hubei province China is limited, so the number of active cases is not available.

Figure 6: The blue dots denotes the number of the active cases (fig.6(a)) / death (fig.6(b)) by COVID-19 in China every day. The red dots denotes the number of the death (fig.6(b)) by COVID-19 in Hubei province China every day. The number of the active cases in Hubei province China is not available.

Generally, the increase of the active cases and the death simulated from the mathematical model is not as large as that in the actual data, and the key performances such as the peak number of the active cases, the number of the death are consistent with the actual information.

3.3 Simulation result of Italy and Spain

In European countries, the mechanism is the same as that in China. The time-dependent parameter the recovery rate is the same as that in China. The difference stems from the parameters like $r_2, \rho$ in the model. The death rate might be high at the beginning, because of lack of medical resources and curing experience. Later the death rate is slowed down, while the recovery rate would increase slowly. The policies in European country is different from that in China. The prediction results of Italy are shown in figure 7(a-7(b), fig.7(a))
Comparing with the actual infection and the predicted infection

Comparing with the actual death and the predicted death

Figure 7: The predicted COVID-19 infected population in Italy. The dots in figures 7(a) 7(b) are the actual results. Figure 7(a) is the simulated number of the active cases. Figure 7(b) is the simulated number of the death. The parameters that are not shown in figures 7(a) 7(b) are $\mu(t)$ and $\delta(t)$ which could be referred to formula (6) and (5). The initial conditions are $S(t_0) = 0$, $I(t_0) = 3$, $R(t_0) = 0$, $D(t_0) = 0$

Similarly, the prediction result simulated for Spain is shown in figure 8(a) 8(b). The simulation for Spain began on Feb 24 2020 when the first case was confirmed.

Similarly to the simulation in China, the increase rate in simulation is not as large as that in the actual data. As it has been reported that the susceptible cases might be very inexplicit [17–19], the initial number of the susceptible cases $S(t_0)$ is not accurate. It would be shown in the section of discussion.

4 Discussion

4.1 Comparing the control policies between China and European countries

The model simulation for China and European countries are with the same mechanism. The only difference are the parameters $r_2$ and $\rho$. And it results in the difference of $\delta$. According to equations (1)
$r_2$ is related to the increasing of the susceptible cases and $\rho$ is related to the number of the recovered patients with nucleic acid testing positive. Both of the parameters are related to the transportation policies. The parameters are determined by the actual data, so the simulation in figures 6-8 are close to the actual data. The parameter $r_2$ and $\rho$ are adjusted to fit the observed data. Though the simulation result is different from the observed data, mainly in the increase rate, the relative quantity of $r_2$ and $\rho$ in China and European countries implies whether the epidemic control policies is rigorous.

4.2 The initial number of the susceptible cases might not be accurately determined

In figures 6-8 it is clear that the simulation result is delayed comparing with the actual result. There might be 2 reasons: one is that the mechanism does not explicitly contain banning of the transportation, the other is that the initial number of susceptible cases is not accurate due to lack of access to COVID-19 testing in the early weeks. For the second reason, we change the initial number of $S(t_0)$ to check whether this deduction is right.

If in fact the initial number of the susceptible cases was 3000 rather than 763 in China, and was 300 rather than 0 in Italy and Spain, then the active cases and the death is shown in figures 9-10.

(a) The hypothesised infection in China is shown in dashed red line
(b) The hypothesised infection in Italy is shown in dashed red line
(c) The hypothesised infection in Spain is shown in dashed red line

Figure 9: All the parameters and initial conditions except those denoted in figures 9(a)-9(c) remain the same as those in the previous simulation of number of the active cases. The dashed red lines show the simulation with hypothesised $S(t_0)$.

(a) The hypothesised deaths in China is shown in dashed red line
(b) The hypothesised deaths in Italy is shown in dashed red line
(c) The hypothesised death in Spain is shown in dashed red line

Figure 10: All the parameters and initial conditions except those denoted in figures 10(a)-10(c) remain the same as those in the previous simulation of number of the death. The dashed red lines show the simulation with hypothesised $S(t_0)$.

The number of the active cases were the same as those in the previous simulation, the only change is
the initial number of the susceptible population and the coefficient $r_2$. It turns out that $r_2$ for European countries is still larger than that for China. It implies that the banning of transportation for China is more rigorous. Moreover, the initial number of the susceptible population is larger than that was reported regardless of the country due to lack of access to COVID-19 test kits in the early weeks.

In view of the simulation by mathematical model, the initial number of the susceptible cases might be difficult to be determined. This might be demonstrated by the report that 60% of active cases might have no symptoms but could infect other healthy people [17-19]. And it is reported that the infected cases are of strong infection when they appeared to have no symptoms [24]. So when the first infected person appears in some area, the infected ones with no symptoms might convert the healthy people into the susceptible ones without records. In the hypothesised simulation, the actual data seemed to be closer to the simulated curves. But the increasing rate of the simulation is not as large as that in the observed data. The reason might be that the transportation control was not considered in this mechanism, which has already been mentioned in the previous section. This implies that banning of the traffic is an effective way to prevent the spreading of disease.

**Competing Interest Statement**

The authors have declared no competing interest.

**Funding Statement**

This study was supported by Fundamental Research Funds for the Central Universities (3332019170), Research on the construction of monitoring information system for blood safety (2016ZX330057), Investigating the influencing factors on donation frequency of plasma donor based on geography information system (Q15086).

**Author contributions**

Feng Lin performed the research. Feng Lin, Jiaxin Liu and Yi Huang designed the research study. Feng Lin, Xu He and Yi Huang analyzed the data. Yonghua Yin, Feng Lin, Yi Huang and Huifang Zhang wrote the paper.

5 Appendix: determination of the coefficient

5.1 The infection coefficient in China

The infection coefficient is calculated in this way:

$$r_1(t) = \frac{I(t)}{S(t)}$$

(1)
The reason why we wrote \( \frac{1}{t} \) before the transition rate of infection is that it is assumed that the contact in fact occurs one by one quite often. The dimension of \( \frac{1}{t} \) is per person i.e. \( \frac{I(t)}{t} \) is dimensionless. For calculation convenience, we just figure out the variable part. According to the data given by Chinese CDC and NHC, the bound of infection coefficient is \( r_1 \in (0.02, 0.21) \). It is hard to give a formula to describe \( r_1 \). The best way is to choose an exact number in its 95% confidence interval to set \( r_1 \). The statistical character of infection coefficient is given by Table 1. The confidence interval is figured by the population mean estimated from the actual data:

<table>
<thead>
<tr>
<th>mean</th>
<th>variance</th>
<th>skewness</th>
<th>kurtosis</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.143738</td>
<td>0.0483431</td>
<td>4.82231</td>
<td>27.0367</td>
<td>(0.0791815, 0.208294)</td>
</tr>
</tbody>
</table>

5.2 The death coefficient in China

The death coefficient is calculated in this way:

\[
delta(t) = \frac{D(t)}{I(t)}
\]  

(2)

The death coefficient is like some exponential function of \( t \). By fitting the time to the \( \ln \delta(t) \). It changes from time to time. However, the tendency of the death coefficient is like some exponential function of \( t \). So by changing the bases of exponential function, it could turn out a series of fitted function. Some fitted function is shown in figure 11.

Figure 11: The fitted lines correspond to different base of exponential function. The blue dots are the death coefficient in the mainland of China, and the red dots are the death coefficient in Hubei province China. The data of the death and the patients could be checked on the website of Chinese CDC and NHC.

In figure 11, the black line is:

\[
\delta(t) = 0.0012 + 0.00250 \times 1.1^{-t}
\]

(3)

the brown line is
\[ \delta(t) = 0.0018 + 0.00301 \times 1.45^{-t} \]  

(4)

The fitted formulas are given below:

\[ \begin{align*}
\delta(t) &= 0.00058 + 0.00287 \times 1.05^{-t} \\
\delta(t) &= 0.0012 + 0.00250 \times 1.1^{-t} \\
\delta(t) &= 0.0015 + 0.00251 \times 1.15^{-t} \\
\delta(t) &= 0.0016 + 0.00258 \times 1.2^{-t} \\
\delta(t) &= 0.0016 + 0.00266 \times 1.25^{-t} \\
\delta(t) &= 0.0017 + 0.00275 \times 1.3^{-t} \\
\delta(t) &= 0.0017 + 0.00284 \times 1.35^{-t} \\
\delta(t) &= 0.0016 + 0.00293 \times 1.4^{-t} \\
\delta(t) &= 0.0018 + 0.00301 \times 1.45^{-t} \\
\delta(t) &= 0.0018 + 0.00310 \times 1.5^{-t} \\
\end{align*} \]

In this study we choose the formula fitted as black line:

\[ \delta(t) = 0.0012 + 0.00250 \times 1.1^{-t} \]  

(6)

the reason is that when \( t \) goes to infinity, \( \delta \rightarrow 0.0012 \).

5.3 The recovery coefficient in China

The recovery coefficient is calculated in this way:

\[ \mu(t) = \frac{R(t)}{I(t)} \]  

(7)

\( \mu(t) \) is like some power function of \( t \). By changing the power of \( t \), the result gives:

Figure 12: The recovery coefficient per day is shown in dots, and the simulated curing coefficients is shown in the line.
The fitted formula of black line is

\[ \mu(t) = 0.00070t^{1.2} \]  \quad (8)

The brown line is the fitted formula

\[ \mu(t) = 0.00004t^2 \]  \quad (9)

however, the blue line is the best:

\[ \mu(t) = 0.00003t^2 \]  \quad (10)

The fitted formulas are given below:

\[ \mu(t) = 0.00070t^{1.2} \]  \quad (11)
\[ \mu(t) = 0.00034t^{1.4} \]
\[ \mu(t) = 0.00017t^{1.6} \]
\[ \mu(t) = 0.000082t^{1.8} \]
\[ \mu(t) = 0.00004t^2 \]
\[ \mu(t) = 0.00002t^{2.2} \]
\[ \mu(t) = 9.15 \times 10^{-6}t^{2.4} \]
\[ \mu(t) = 4.37 \times 10^{-6}t^{2.6} \]
\[ \mu(t) = 2.08 \times 10^{-6}t^{2.8} \]
\[ \mu(t) = 9.89 \times 10^{-7}t^{3} \]

In this study we use the formula:

\[ \mu(t) = 0.00003t^2 \]  \quad (12)

References


[14] Yingqing Li. Yunnan adopts more measures to control virus outbreak. https://www.chinadaily.com.cn/a/202002/03/WS5e381eda310128217274872.html


