

# A fatality data based on an optimized SEIR Model for Epidemic: A study about the testing and quarantining

Liwei Yang (✉ [loveblairsky@gmail.com](mailto:loveblairsky@gmail.com))

Bayview Secondary School <https://orcid.org/0000-0001-7115-8049>

Bochuan Chen

Beihang University

Siyao Du

China Medical University First Hospital

Lang Zhao

Fuwai Hospital

Yu Zhang

Beihang University

---

## Article

**Keywords:** SIR Modeling, COVID-19, testing rate, fatality data, quarantining ratio, estimation of the time of outbreak, future epidemic prediction

**Posted Date:** June 30th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-38431/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

## Title: A fatality data based on an optimized SEIR Model for Epidemic: A study about the testing and quarantining

Keywords: SIR Modeling; COVID-19; testing rate; fatality data; quarantining ratio; estimation of the time of outbreak; future epidemic prediction

### Abstract

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), caused by a newly discovered coronavirus, has been announced as a global pandemic by WHO on March 12, 2020. We aimed to provide a more accurate, object-oriented SEIR model for Epidemic as well as a thorough analysis for different intervention strategies and critical parameters. Based on the object-oriented mindset in computer science, we assigned properties to every independent patient, which includes age, likelihood of catching the virus, independent mortality rate, social activeness, awareness of the danger of the virus. We also considered macro level parameters including testing rate, positive rate. Based on the fatality data obtained, we estimated the occurrence time of the first case in China (excluding Hubei), Russia, Germany, and the United States using the model. Eight groups of virus-transmission simulation with different testing rate and quarantine ratio were conducted. The mortality data was found to be an effective and crucial parameter while the number cases can be decisive reflecting the severity of the pandemic.

### Introduction

Novel coronavirus pneumonia is an acute infectious disease that transmits rapidly from person to person which has a relatively high mortality rate. The initial symptoms of the virus are fatigue, fever, dry cough, and dyspnea.<sup>1</sup> Most patients can develop antibody and cured without medical intervention; however, 19.1% patients with sever symptoms may develop acute respiratory distress syndrome, which may lead to death.<sup>2</sup> Soon after it was detected in late December 2019, from Wuhan, China, the number of cases grew exponentially worldwide and became a global pandemic.<sup>3</sup> According to the data from the World Health Organization, on April 18,2020, there are 2,160,207 confirmed Covid-19 cases worldwide, with more than 146,008 deaths caused by the virus.<sup>4</sup> Since the current situation of the epidemic is growing worse as the virus spreads silently, it is vital to predict the future situation in the affected regions, and analyze the strategies taken, which would help to find the more reasonable and effective intervention to reduce the damage of the virus. Based on the SEIR model, we constructed the optimized OO-SIR model, which considered age distribution of the patients, dynamic mortality rate, social activeness of individuals, positive rate, and testing rate. Some studies have already discussed containment strategies in UK, Singapore and Hongkong. However, we believe a detailed discussion is required to analyze the significance of testing and quarantining patients.<sup>5,6,7</sup>

### Method

#### Parameter

##### Age

The age of patients of COVID-19 disease and their mortality ratio demonstrates a strong gradient pattern. In the OO-SIR model, there are 3 parameters related to age: patient age distribution of the target population ( $I_a$ ), age distribution of the target population ( $N_a$ ), and constant likelihood of patients in a certain age interval to catch the virus ( $\beta_0$ ).  $\beta_0$  is a critical parameter in the model - which determines the predicted patient age distribution in the target population – will influence the mortality rate directly, has to be calculated with a reliable source of data.

China is the first country that controls the community transmission of the coronavirus, which has a relatively low calculated mortality rate which is less than 2%(Calculated mortality rate) and a large number of infections (over 12,000 reported cases). Therefore, it is reasonable to believe that the number of tests conducted in China is large enough, which make its data reliable enough to calculate the  $\beta_0$ . Thus, the data of China (excluding Hubei) is used to calculate the  $\beta_0$ .  $I_0$  (patient age distribution in China excluding Hubei),  $N_0$  (population age distribution in China), and  $\beta_0$  (likelihood of patients in a certain age interval to catch the virus) is shown (Table 1).

In particular, the relationship between  $\beta_0$ ,  $I_0$ , and  $N_0$  is shown (Equation(1)).<sup>2,8</sup>

$$\beta_0 = \frac{I_0}{N_0} \quad (1)$$

**Table 1 Age distribution of patient population age distribution, and likelihood of patients in a certain age interval to catch the virus in China (excluding Hubei)**

Age	$I_0$	$N_0$	$\beta_0$
$\leq 20$	0.021	0.24	0.088
20~	0.081	0.17	0.476
30~	0.17	0.16	1.06
40~	0.19	0.17	1.12
50~	0.22	0.12	1.83
60~	0.19	0.07	2.74
70~	0.099	0.04	2.20
$\geq 80$	0.032	0.016	2.00

After obtaining the  $\beta_0$  and age distribution of the target population ( $N_a$ ), the patient age distribution of the target population ( $I_a$ ) can be obtained (Equation(2)).

$$I_a = \beta_0 N_a \quad (2)$$

#### Mortality Rate

In conventional SIR models, the mortality rate is considered as a constant with a certain age distribution<sup>9</sup>. According to Rajgor, the mortality rate of COVID-19 is 1.38%.<sup>10</sup> Nevertheless, at normal circumstance, the calculated mortality rate of Covid-19 is way higher than 1.38%. Zhang revealed the relationship between mortality rate of Covid-19 and the number of health workers - the mortality rates in Hubei and Wuhan demonstrated an exponential decay pattern as medical workers increase.<sup>11</sup> Thus, in OO-SIR model, the mortality rate is defined as a variable: the function about population mortality rate( $\mathcal{D}$ ), initial population mortality rate( $\mathcal{D}_0$ ), medical system coefficient( $\varpi$ ) is shown (Equation (3)).

$$\mathcal{D} = \mathcal{D}_0 \cdot \varpi \quad (3)$$

In the real-world scenario, the mortality rate is affected by many factors, including comorbidities, immune system, random mutation, medical intervention, etc. In OO-SIR model, the level of medical intervention of a patient is described as medical system coefficient ( $\varpi$ ). It is not easy to quantify the health condition and basic disease of every patient during mathematical modeling; however, the positive relationship between age of the patient and the risk of death of the patient is revealed. In other words, a patient with greater age is tend to have poorer health condition and more comorbidities. A study from Chinese Epidemiology Working Group for NCIP Epidemic Response revealed the relationship between age and mortality rate.<sup>2</sup> Thus, OO-SIR model modeled the relationship between the age of patient and the individual mortality rate ( $d_I$ ), assuming the age as the only factor affecting the risk of death, while considering the intervention of the medical system.

The individual mortality rate of a patient in a certain age interval will increase as the age increases. By modeling the data with exponential function, the following equation can be obtained. The relationship between  $d_I$ , age of the patient ( $A$ ), and  $\varpi$  can be written (Equation (4)).

$$d_I = \varpi(I_H, C_H) \mathbf{0} \cdot \mathbf{01}e^{0.07A} \quad (4)$$

On a greater scale, the population mortality rate ( $\mathcal{D}$ ) represents the average individual mortality rate ( $d_I$ ) within the population. Thus, the relationship between population mortality rate ( $\mathcal{D}$ ), individual mortality rate ( $d_I$ ), and number of infected patients ( $I$ ) can be written (Equation (5)).

$$\mathcal{D} = \frac{\sum d_I(A)}{I} \quad (5)$$

In the model, the initial population mortality rate ( $\mathcal{D}_0$ ) is defined when the country has sufficient medical supply. Since the age distribution within a region/country is relatively stable during the spread of virus, the

initial population mortality rate ( $\mathcal{D}_0$ ) is stable as well, it is defined as a constant. The relationship between initial population mortality rate ( $\mathcal{D}_0$ ), initial individual mortality rate ( $d_0$ ), and the age of the individual patient ( $A$ ) is shown (Equation (6)).

$$\mathcal{D}_0 = \frac{\sum d_0(A)}{I} = \frac{\sum I_a 0.01 e^{0.07A}}{I} \quad (6)$$

Zhang Indicated that the population mortality rate decayed exponentially as more hospital beds and doctors (medical system capacity) fluxed into Wuhan; thus, on the contrary, the population mortality rate would grow exponentially as the medical system gets overwhelmed. However, the growing pattern will not be unlimited. Due to the reason that the mortality rate without medical intervention is certain, the growth will converge to the mortality rate when there is no medical intervention. Therefore, the medical system occupancy rate has significant impact on the overall population mortality rate ( $\mathcal{D}$ ) OO-SIR model described this phenomenon mathematically. A function with such a pattern can be described using the error function.<sup>12</sup> The model defined medical system coefficient ( $\varpi$ ) to help calculating the theoretical mortality rate, which will grow as more medical resource is occupied. The function about medical system coefficient ( $\varpi$ ), hospitalized cases ( $I_H$ ), and medical system capacity ( $C_H$ ) is shown (Equation (7)).

$$\varpi = \left( \int_0^{\frac{I_H}{C_H}-2} e^{-\eta^2} d\eta + 1 \cdot 5 \right) - 1 \quad (7)$$

#### Calculated mortality rate

It is almost impossible to detect every patient, which means there has to be untested, undocumented cases. This phenomenon may cause the calculated mortality rate to be higher than the population mortality rate. The calculated mortality rate ( $\mathcal{D}_c$ ) is defined as the mortality rate calculated by the data provided by the health official in a country, province, or region. The relationship between calculated mortality rate, confirmed cases ( $I_A$ ), and confirmed deaths ( $F_A$ ) was defined (Equation (8)).

$$\mathcal{D}_c = \frac{I_A}{F_A} \quad (8)$$

#### Social activeness and public awareness

The spreading speed of the virus ( $\beta$ ) is proportional with contact frequency between objects.<sup>13</sup> To better describe the contact frequency, the OO-SIR model defined effective contact frequency ( $f$ ). In the model, there are four critical parameters defined that affect the effective contact frequency: social activeness  $\varepsilon$  (Equation(10)), public awareness  $\alpha$  (Equation(11)), and the immunity coefficient  $\sigma$  (Equation(12)). The relationship between  $f$ ,  $\varepsilon$ ,  $\alpha$ ,  $\sigma$ ,  $\mu$ , and number of active cases ( $I_c$ ) is shown (Equation(9)).

$$f = \sigma \frac{(1 - \alpha) \sum \varepsilon}{I_c} \quad (9)$$

In the model,  $\varepsilon$  is assigned using normal distribution; object with higher social activeness tends to make more contacts per day. The social activeness distribution function about constant t, coefficient  $C_2$ ,  $c_2$  are written (Equation(10)).

$$\varepsilon = C_2 + e^{-c_2 \eta^2} \quad (10)$$

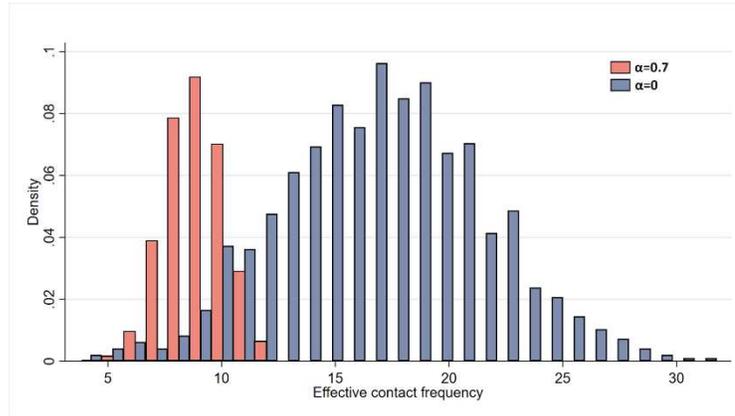
$\alpha$  represents the population's awareness of the danger of the virus and the strength of government enforcements. Intuitively, the population will decrease the frequency of social contact and wear PPE intentionally when they are conscious about the virus. Thus, public awareness is going to increase as more patients died due to the COVID-19 or the strength of city lockdown increases. As the  $\alpha$  increases, the effective contact frequency ( $f$ ) will decrease.

The relationship between public awareness ( $\alpha$ ), fatality ( $F$ ), days since city lockdown ( $T_A$ ), and constant  $C_3$  is shown (Equation(11)).

$$\alpha = \frac{2}{\sqrt{\pi}} \int_0^{\text{Log}_{10} F - 1} e^{-\eta^2} d\eta + C_3 T_A \quad (\alpha > 0) \quad (11)$$

The distribution of effective contact frequency made by objects will shift when  $\alpha$  changes (Figure 1); the distribution of effective

contact frequency will shift to the left if the public awareness increases.



**Figure 1** The effect of social awareness on effective contact frequency

$\sigma$  is defined as immunity coefficient, which represents the proportion of people who has immunity within a population.<sup>14</sup> The relationship between immunity coefficient ( $\sigma$ ), number of infected patients ( $I$ ), recovered cases ( $R$ ), and population ( $N$ ) is shown (Equation(12)).

$$\sigma = \frac{N - (R + I)}{N} \quad (12)$$

#### Quarantining the patients

In practice, most confirmed coronavirus cases will be quarantined; however, there are still unconfirmed and asymptomatic cases. To make the model closer to the real-world scenario, the number of infected patients ( $I$ ) is divided into two groups: patients who are able to transmit the virus ( $I_c$ ) and isolated cases ( $I_s$ ). The relationship between them is shown in (Equation(13)).

$$I = I_c + I_s \quad (13)$$

The model also included the concept of quarantine ratio ( $i_s$ ), representing the proportion of isolated cases in all coronavirus cases. Due to the fact that quarantined patients are isolated from the susceptibles, they are theoretically less likely to transmit the virus. The relationship between quarantine ratio ( $i_s$ ), isolated cases ( $I_s$ ), and number of infected patients ( $I$ ) is written as (Equation(14)).

$$i_s = \frac{I_s}{I} \quad (14)$$

However, the  $i_s$  is extremely difficult to calculate due to the insufficient coronavirus tests conducted in some regions, the  $i_s$  calculated in the model using (Equation(17)).

#### Positive rate

The positive rate ( $r_t$ ) represents the ratio of confirmed cases and total tests conducted. Although  $r_t$  has no actual impact on the computation in the model, a region with low  $r_t$  tend to have lower testing rate. Lower testing rate means less patients confirmed and less cases quarantined, resulting in lower  $i_s$ . Therefore, the model assumes that  $r_t$  is inversely proportional with  $i_s$ , as shown in (Equation (15)).

$$r_t \propto \frac{1}{i_s} \quad (15)$$

Countries with high  $r_t$  tend to have higher  $\mathcal{D}_c$  (Figure 2). This phenomenon is potentially because a higher positive rate means a higher proportion of untested patients. With a same number of total tests conducted (when the number is large enough), more patients will test positive and the  $r_t$  will be higher.

A higher number of patients tend to lead to a higher number of  $\mathcal{D}_c$ , so here concludes that the  $\mathcal{D}_c$  is inversely proportional with  $r_t$ , their relationship is shown in (Equation(16)).

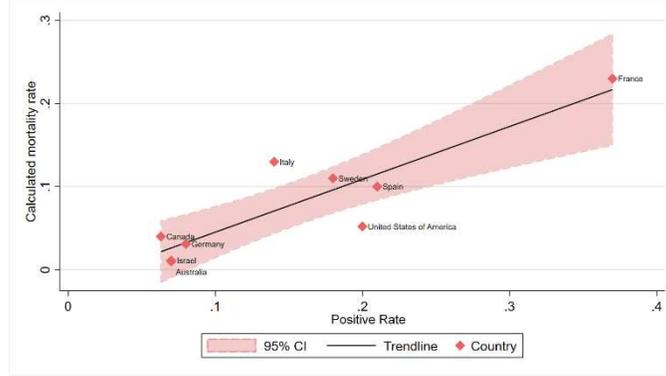
$$r_t \propto \frac{1}{D_c} \quad (16)$$

Thus, the following equation can be drawn.

$$r_t \propto \frac{1}{i_s} \propto \frac{1}{D_c}$$

By applying the hospitalized cases and positive rate, the  $i_s$  can be estimated (Equation(17)) ( $C_4$  is a constant).

$$i_s = \frac{i_a C_4}{r_t} \quad (17)$$



**Figure 2** The trend of calculated mortality rate and positive rate in Israel, Australia, Canada, Germany, Italy, Sweden, the United States of America, Spain, and France

#### Testing rate

In the model, the testing rate ( $r_s$ ) is defined as the percentage of confirmed patients as a proportion of the number of patients that have symptom and can be tested. This parameter can reflect the accuracy and reliability of the data of confirmed cases in a country. The relationship between testing rate ( $r_s$ ), tested and confirmed cases  $I_t$ , and number of infected patients ( $I$ ) is written (Equation(18)).

$$r_s = \frac{I_t}{I} \propto \frac{1}{r_t} \quad (18)$$

#### Model

A computer model is constructed based on the model mentioned above; the simulation results were calculated using the model. The data that need to input is shown (Table 2), all the data are gathered from the health officials and Johns Hopkins university CSSE.<sup>15</sup>

The OO-SIR model can be described as following equations:

$$\begin{cases} \frac{dS(t)}{dt} = N - (dI + dR) \\ \frac{dI(t)}{dt} = \beta f I_c \\ \frac{dR(t)}{dt} = \gamma I = (\delta + D)I \\ \frac{dD(t)}{dt} = DI \\ \frac{dr(t)}{dt} = \delta I \\ \beta = \frac{\mathcal{R}_0}{C_a} f \end{cases}$$

$$f = \sigma \frac{(1 - \alpha) \Sigma \varepsilon}{I_c}$$

$$\varepsilon = C_2 + e^{-c_2 t^2}$$

$$\sigma = \frac{N - (R + I)}{N}$$

$$\alpha = \frac{2}{\sqrt{\pi}} \int_0^{\text{Log}_{10} F^{-1}} e^{-\eta^2} d\eta + C_3 T_A$$

$$\gamma = \mathcal{D} + \delta$$

$$\mathcal{D} = \mathcal{D}_0 \cdot \varpi = \frac{2\mathcal{D}_0}{\sqrt{\pi}} \cdot \left( \int_0^{\frac{I_H}{C_H} - 2} e^{-\eta^2} d\eta + 1 \cdot 5 \right) - 1$$

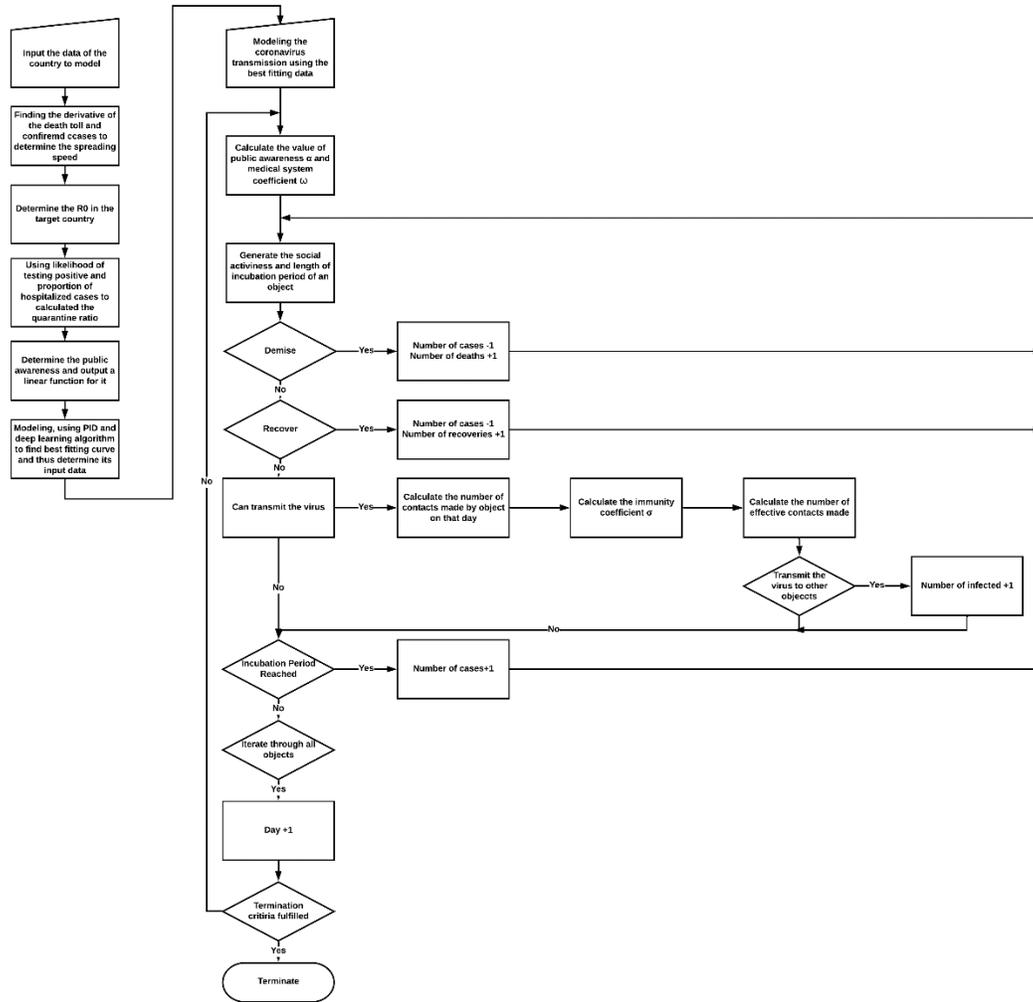
$$\delta = \frac{\int_{21}^t e^{-\ln(C_1(\eta-21))^2} d\eta}{\int_{21}^{29} e^{-\ln(C_1(\eta-21))^2} d\eta}$$

$$I_c = I - I_s$$

**Table 2 The parameters included in the OO-SIR model**

Parameter	Meaning
$C_1, C_2, C_3$	Constant
$S$	Susceptible
$I$	Infected
$t$	Time
$R$	Removed
$D$	Death
$r$	Resolved
$\alpha$	Public awareness
$\sigma$	Immunity coefficient
$\gamma$	Rate which the infected cases being removed
$\mathcal{D}_0$	Initial population moratity rate
$\mathcal{D}$	Population moratity rate
$\delta$	Rate which recovered cases are discharged
$I_H$	Hospitalized cases
$C_H$	Hospital capacity
$f$	Effective contact frequency
$I_c$	Patients who are able to transmit the virus
$I_s$	Isolated cases
$T_A$	Time after lockdown

The procedure of OO-SIR Model is shown (Figure 3).



**Figure 3 The flow chart of the OO-SIR model**

By deploying machine learning and PID algorithm, the model can find out the undermined coefficients in the model: testing rate, quarantine ratio, date of lockdown, etc. The model will calculate the derivative and integral to do the regression analysis, and try to find the best fitting scenario and to give reference for later modeling.

## Result

### Simulation result

Based on the model discussed above, by inputting the data of each country and region<sup>15</sup>, the coronavirus transmission and deaths simulations of China (Excluding Hubei), Russia, Germany, and the United States were conducted.  $\mathcal{R}_0$  and  $\alpha$  were recalculated by the model and redressed manually;  $\mathcal{P}$ ,  $N$ ,  $C_H$ ,  $i_s$  were entered manually (Table 3).

**Table 3 The modeling input parameters of China (Excluding Hubei), Russia Germany, and the United States**

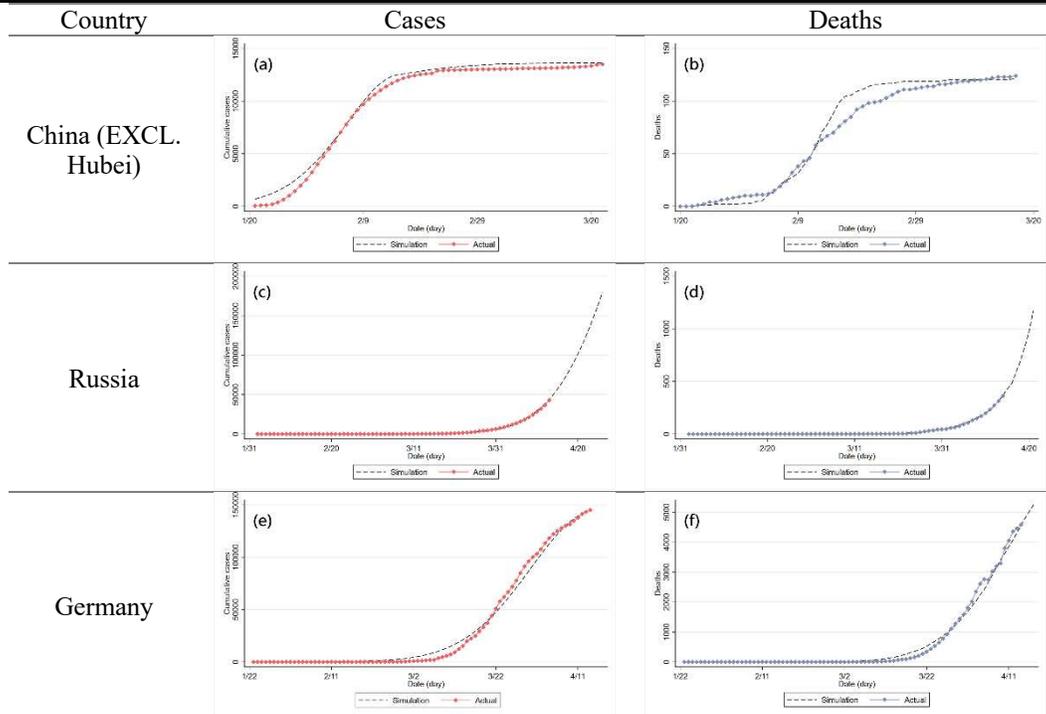
Parameter	China (EXCL. Hubei)	Russia	Germany	United States
-----------	---------------------	--------	---------	---------------

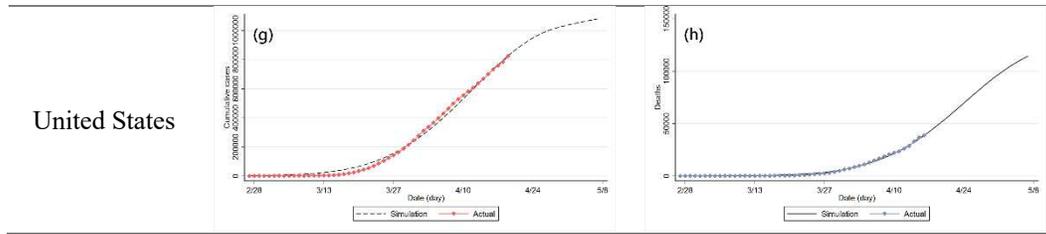
$\mathcal{P}$	13127	5934	3400	7,799
$N$	1,400,000,000	144,500,000	88,000,000	328,200,000
$\mathcal{R}_0$	2.79	2.34	2.02	2.49
$C_H$	$\infty$	25,000	100,000	400,000
$i_s$	0.9	0.5	0.6	0.2
$\alpha$	0-0.95	0-0.9	0-0.9	0-0.85

Throughout the simulation, the simulated cumulative case curve and actual cumulative case curve of China (Excluding Hubei), Russia, Germany, and the United States demonstrated high degrees of fitting (Table 4) (determination coefficients of 0.978, 0.999, 0.995, and 0.997). The simulated death toll curves and actual death toll curves of China (Excluding Hubei), Russia, Germany, and the United States have determination coefficients of 0.978, 0.998, 0.994, and 0.995. The occurrence time of the first case, occurrence of first death, the testing rate, final number of cases, the final number of reported cases, and testing rate were estimated

**Table 4 The result of modeling, curve of cumulative cases, and curve of deaths tolls of China (Excluding Hubei), Russia Germany, and the United States**

Parameter	China (EXCL. Hubei)	Russia	Germany	United States
Occurrence of first case	Dec.29, 2019	Feb.5, 2020	Jan.3,2020	Dec.23, 2019
Occurrence of death case	Jan.28, 2020	Mar.9, 2020	Feb.14,2020	Feb.3, 2020
Number of cases	13,500-16,667	100,000-720,000	145,000-360,000	6,660,000-25,600,000
Number of reported cases	9,450-15,000	80,000-648,000	140,000-360,000	1,160,000-25,600,000
Testing rate	0.70-0.90	0.80-0.90	0.60-0.80	0.05-0.15





Noticeably, Germany and the United States have relatively low  $r_s$ , which may cause a higher  $D_c$  (Figure 2); however, the model demonstrated a high uncertainty of the  $r_s$  of Germany, after multiple trails, the model offered 2 potential explanation: the patient age distribution in Germany skewed towards elderly, or the  $r_s$  in Germany is fairly low (0.60-0.80). In the United States, the growth of the number of cumulative cases demonstrated linear pattern, which very likely means that the testing capacity has reached. Therefore, the number of daily new cases would change drastically unless the  $r_t$  changes. The high  $D_c$  of the United States is mainly due to its low  $r_s$

### Testing rate

The testing rate plays a crucial role affecting the health officials to make decision and respond to contain the virus by influencing the number of cases reported. We analyzed the effect of testing rate on the overall virus transmission – death toll, cumulative cases, active cases, and reported cases – with different quarantine ratio.

#### Effect of testing rate on the number of cases and deaths with low quarantine ratio

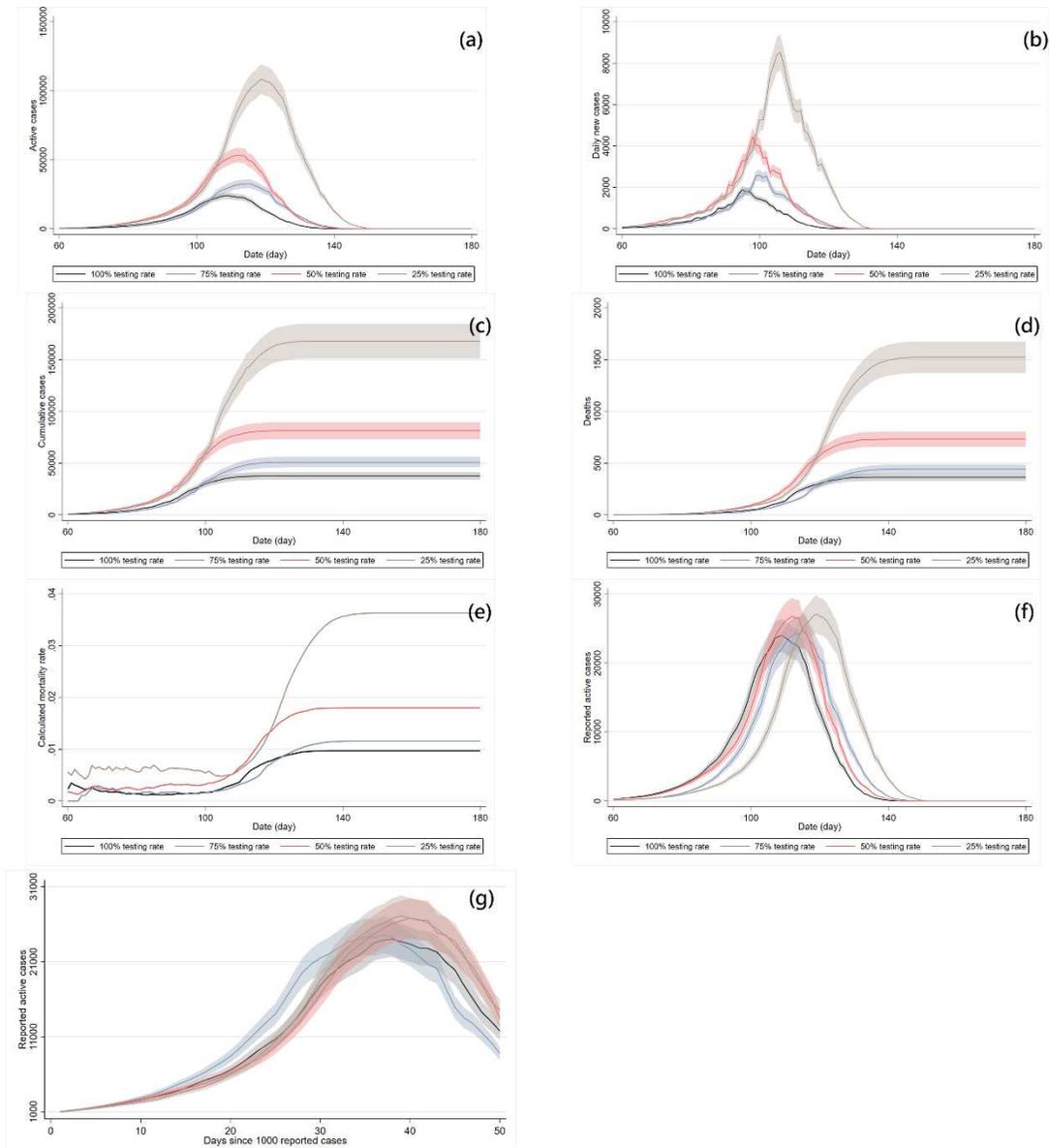
By conducting more coronavirus tests, the governments and health officials can acquire a more accurate view of the transmission of the virus. The model was used to investigate the effect of testing rate on the overall number of deaths and cases. Four trails with 100%, 75%, 50%, and 25% testing rate were conducted, city lockdown will trigger when 10,000 cases are confirmed, assuming same speed of responding.

**Table 5 The input parameters of four groups with 100%, 75%, 50%, and 25% testing rate with a relative same quarantine ratio and trigger of city lockdown**

Parameter	100% Testing rate	75% Testing rate	50% Testing rate	25% Testing rate
$\mathcal{P}$	N/A	N/A	N/A	N/A
$N$	10,000,000	10,000,000	10,000,000	10,000,000
$\mathcal{R}_0$	2.5	2.5	2.5	2.5
$C_H$	$\infty$	$\infty$	$\infty$	$\infty$
$i_s$	0	0	0	0
$\alpha$	0-0.95 (lockdown trigger after 10,000 cases)			

When there are no patients quarantine, the curves of reported active cases converges despite the fact that the difference in actual number of cases is large (up to 335.5%) (Figure 4). The four groups have remarkable difference in the number of deaths, the calculated mortality rate demonstrated is inversely proportional with the testing rate as predicted before. Thus, when the quarantine ratio is 0 (or ignorable compare to the number of actual cases), difference between the testing rate cannot result in a significant shift in the number of confirmed cases. The testing rate has a proportional relationship with the number of reported cases, which cannot be estimated solely through the curve of reported cases and daily new cases, meaning the severity of the case cannot be estimated solely through the curve of reported cases.

The calculated mortality rate raised as the spreading speed gradually decreases; however, the calculated mortality rate may decrease due to more tests conducted, which is not expected by the model. This could be attributed to the improvement of testing capacity, which brings a higher testing rate thus a lower calculated mortality rate.



**Figure 4 Demonstrative trails – spreading speed of the virus (a) number of active cases (b) number of cumulative cases (c) number of reported active cases (d) calculated mortality rate (e) number of reported active cases (f) convergence of reported active cases (g) with 100% testing rate, 75% testing rate, 50% testing rate, 25% testing rate without quarantine and same trigger of city lockdown**

**Effect of testing rate on the number of cases and deaths with high quarantine ratio**

Another group of trails with higher quarantine ratio were also done to investigate the effect of testing rate on the transmission of the virus. With the population of 10,000,000 and quarantine ratio of 0.75 (relative to reported cases), the following 4 trails were conducted (Table 6).

**Table 6 The input parameters of four groups with 100%, 75%, 50%, and 25% testing rate with a relative same quarantine ratio and trigger o city lockdown**

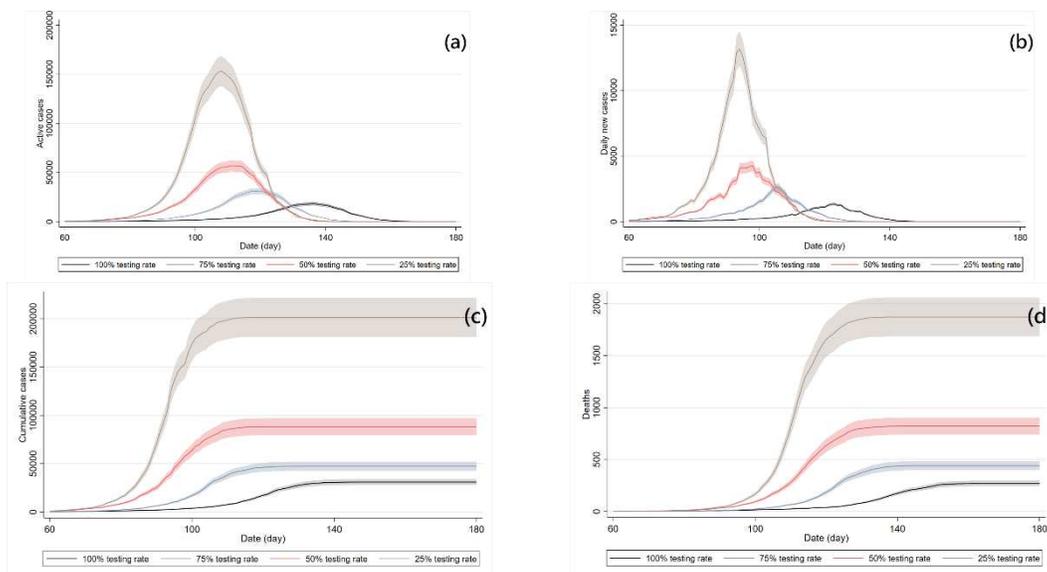
Parameter	100% Testing rate	75% Testing rate	50% Testing rate	25% Testing rate
-----------	-------------------	------------------	------------------	------------------

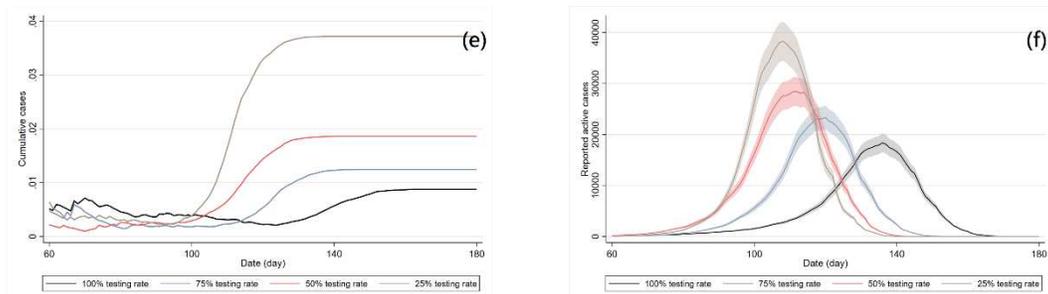
$\mathcal{P}$	N/A	N/A	N/A	N/A
$N$	10,000,000	10,000,000	10,000,000	10,000,000
$\mathcal{R}_0$	2.5	2.5	2.5	2.5
$C_H$	$\infty$	$\infty$	$\infty$	$\infty$
$i_s$	0.75(0.75 relative to reported cases)	0.56 (0.75 relative to reported cases)	0.38 (0.75 relative to reported cases)	0.19(0.75 relative to reported cases)
$\alpha$	0-0.95 (lockdown trigger after 10,000 cases)			

With a higher quarantine ratio (relative to reported cases), the pattern of convergence of the reported active cases was not observed here (Figure 5). Furthermore, trails with lower testing rate tend to take shorter amount of time to reach peak, which may be because of an earlier triggering of the lockdown. In this case, higher spreading speed of the virus would indicate a lower testing rate. Nevertheless, the calculated mortality rate would not change noticeably when the quarantine ratio is changed compare to the previous trails. In a same population, the only factor affecting the calculated mortality rate is the testing rate.

**Table 7 The result of modeling of four groups - control, social distancing, quarantining patients, and social distancing and quarantining patients**

Output	100% Testing rate	75% Testing rate	50% Testing rate	25% Testing rate
$N$	10,00,000	10,00,000	10,00,000	10,00,000
Cumulative cases	30,892	47,251	88,722	201,426
Active cases at peak	18,387	31,098	56,992	153,169
Reported cases at peak	18,387	23,324	28,496	38,292
Death	271	441	882	1873
Decreased by	84.6%	76.5%	56.0%	0%
Time until peak	137 days	121 days	112 days	109 days





**Figure 5 Demonstrative trails – spreading speed of the virus (a) number of active cases (b) number of cumulative cases (c) number of reported active cases (d) calculated mortality rate (e) number of reported active cases (f) with 100% testing rate, 75% testing rate, 50% testing rate, 25% testing rate without quarantine and same trigger of city lockdown**

## Discussion

With the object-oriented mindset, we brought the SIR model to the individual level, and included the medical system coefficient, testing rate, and positive rate, to provide an insight for the transmission and fatality of COVID-19. We used the model to simulate the COVID-19 transmission scenario in China (excluding Hubei), Russia, Germany, and the United States and effect of different strategies, testing rate, and quarantine ratio on the transmission of the virus. The date of the first cases in China (excluding Hubei), Russia, Germany, and the United States was estimated to be Dec.29,2019 (95% CI Dec.26,2019-Jan.1,2020), Feb.5,2020 (95% CI Feb.2,2020-Feb.8,2020), Jan.3,2020 (95% CI Dec.31,2019-Jan.6,2020), Dec.23,2020(95% CI Dec.20,2019-Dec.26,2019); the testing rate to be 0.9-1.0, 0.95-1.0, 0.60-0.90, and 0.05-0.15 with 95% confidence interval.

According to our modeling, enforcing city lockdown when 5% of the population are infected can decrease the final number of cases by 58.6% (95% CI 54%-63%), quarantining 75% patients can decrease the final number of cases by 17.5% (95% CI 10-26%), and using both them and decrease the final number of cases by 93.0% (95% CI 92%-95%). With no intervention to mitigate the transmission of the virus, 96.5% of the population are infected, which is substantially higher than previous estimations. We believe this due to a relatively high proportion of people infected despite the fact that the  $R_0$  is lower than 1. It is extremely important topic the right approach to balance the economy and the curve, fast and strict response is required to contain the virus and avoid fatality.

We found the phenomenon that when the relative quarantine ratio is low, the number of reported cases would converge regardless the significant difference in testing rate. The relationship between testing rate and calculated mortality rate was proven, which may provide a new method for future strategy making and intervention. The model indicates that the decrease of daily new cases may be merely due to the decrease of positive rate when the testing capacity is limited. The severity of the virus must be underestimated, countries have to keep enforcing social distancing to protect the risky population.

We reserved port for later optimizations of the model, which will help us keep advancing the model and improve the accuracy. Also, we provided extension classes for potential revisions.

Also, our study has some limitations. OO-SIR model, as a fatality-based model, assumed a certain distribution of the age of patients, which may cause age of patients to deviated from the real-world scenario, which may cause the number of actual cases to be overestimated. Also, the effect of social distancing may also be overestimated since not everybody is wearing PPE at current time. We also did not consider the difference between ability of transmission of asymptomatic cases and symptomatic cases. Yu revealed that COVID-19 has airborne transmission, which would bring more uncertainty to transmission inside a condo or business establishment.<sup>16</sup>

Due to the reason that the population in the model is closed, the travellers from other countries and regions are ignored, the import cases from other countries will make the scenario more complex and difficult to predict. Secondly, the country is considered as a cohesive whole, which may result the prediction to deviate as higher proportion of the population getting infected.

The estimated occurrence of the first case can be influenced significantly by the living style of the persons, which will bring down the confidence interval of the occurrence of the first cases by 40% if the first case is twice as socially active as estimated.

Overall, our modeling added up uncertainty to the previous estimations of the severity of COVID-19 in some countries. The reliability of this model was proven through simulation of various countries. The model suggest that the social distancing and cases quarantining are the most effective ways to control the virus, and, the herd immunity may not be as useful as previous thought. After multiple simulations, we believe lockdown and social distancing has to be continued to flatten the curve.

## References

---

<sup>1</sup> Pan L, Mu M, Yang P, Sun Y, Wang R, Yan J, et al. Clinical Characteristics of COVID-19 Patients With Digestive Symptoms in Hubei, China. *The American Journal of Gastroenterology*. 2020;115(5):766–73.

<sup>2</sup> Epidemiology Working Group for NCIP Epidemic Response. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. *Chinese Journal of Epidemiology*, 2020, 41(2): 145-151.

<sup>3</sup> WHO. WHO announces COVID-19 outbreak a pandemic. March 12, 2020. <http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic> (accessed Apr 20, 2020).

<sup>4</sup> WHO. Coronavirus disease 2019 (COVID-19) situation report—89. March 3, 2020. [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200418-sitrep-89-covid-19.pdf?sfvrsn=3643dd38\\_2](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200418-sitrep-89-covid-19.pdf?sfvrsn=3643dd38_2) (accessed Apr 20, 2020).

<sup>5</sup> Davies NG, Kucharski AJ, Eggo RM, Gimma A, Edmunds WJ, Jombart T, et al. Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. *The Lancet Public Health*. 2020;

<sup>6</sup> Koo JR, Cook AR, Park M, Sun Y, Sun H, Lim JT, et al. Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study. *The Lancet Infectious Diseases*. 2020;20(6):678–88.

<sup>7</sup> Cowling BJ, Ali ST, Ng TWY, Tsang TK, Li JCM, Fong MW, et al. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. *The Lancet Public Health*. 2020;5(5).

<sup>8</sup> National Bureau of Statistics of China. Report of sixth national population census of the people's republic of china. April 28,2011. <http://www.stats.gov.cn/tjsj/tjgb/rkpcgb/> (accessed Apr 20, 2020)

<sup>9</sup> Maki Y, Hirose H. Infectious Disease Spread Analysis Using Stochastic Differential Equations for SIR Model. 2013 4th International Conference on Intelligent Systems, Modelling and Simulation. 2013;

<sup>10</sup> Rajgor DD, Lee MH, Archuleta S, Bagdasarian N, Quek SC. The many estimates of the COVID-19 case fatality rate. *The Lancet Infectious Diseases*. 2020;

<sup>11</sup> Zhang Z, Yao W, Wang Y, Long C, Fu X. Wuhan and Hubei COVID-19 mortality analysis reveals the critical role of timely supply of medical resources. *Journal of Infection*. 2020;

<sup>12</sup> Nandagopal M, Sen S, Rawat A. A Note on the Error Function. *Computing in Science & Engineering*. 2010;12(4):84–8.

<sup>13</sup> Hu H, Nigmatulina K, Eckhoff P. The scaling of contact rates with population density for the infectious disease models. *Mathematical Biosciences*. 2013;244(2):125–34.

<sup>14</sup> Fine P, Eames K, Heymann DL. "Herd Immunity": A Rough Guide. *Clinical Infectious Diseases*. 2011;52(7):911–6.

---

<sup>15</sup> Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*. 2020;20(5):533–4.

<sup>16</sup> Yu IT, Li Y, Wong TW, Tam W, Chan AT, Lee JH, et al. Evidence of Airborne Transmission of the Severe Acute Respiratory Syndrome Virus. *New England Journal of Medicine*. 2004;350(17):1731–9.

# Figures

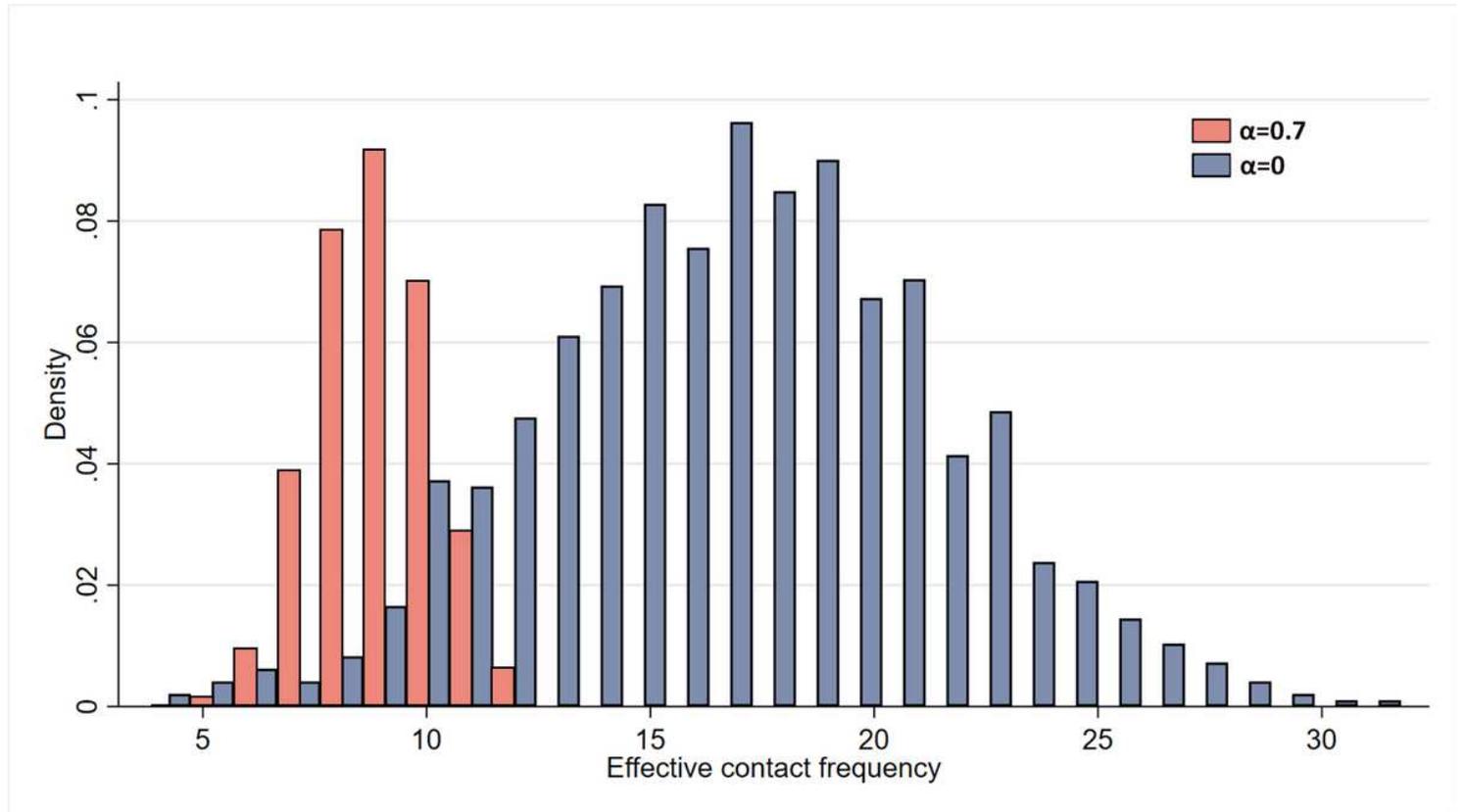
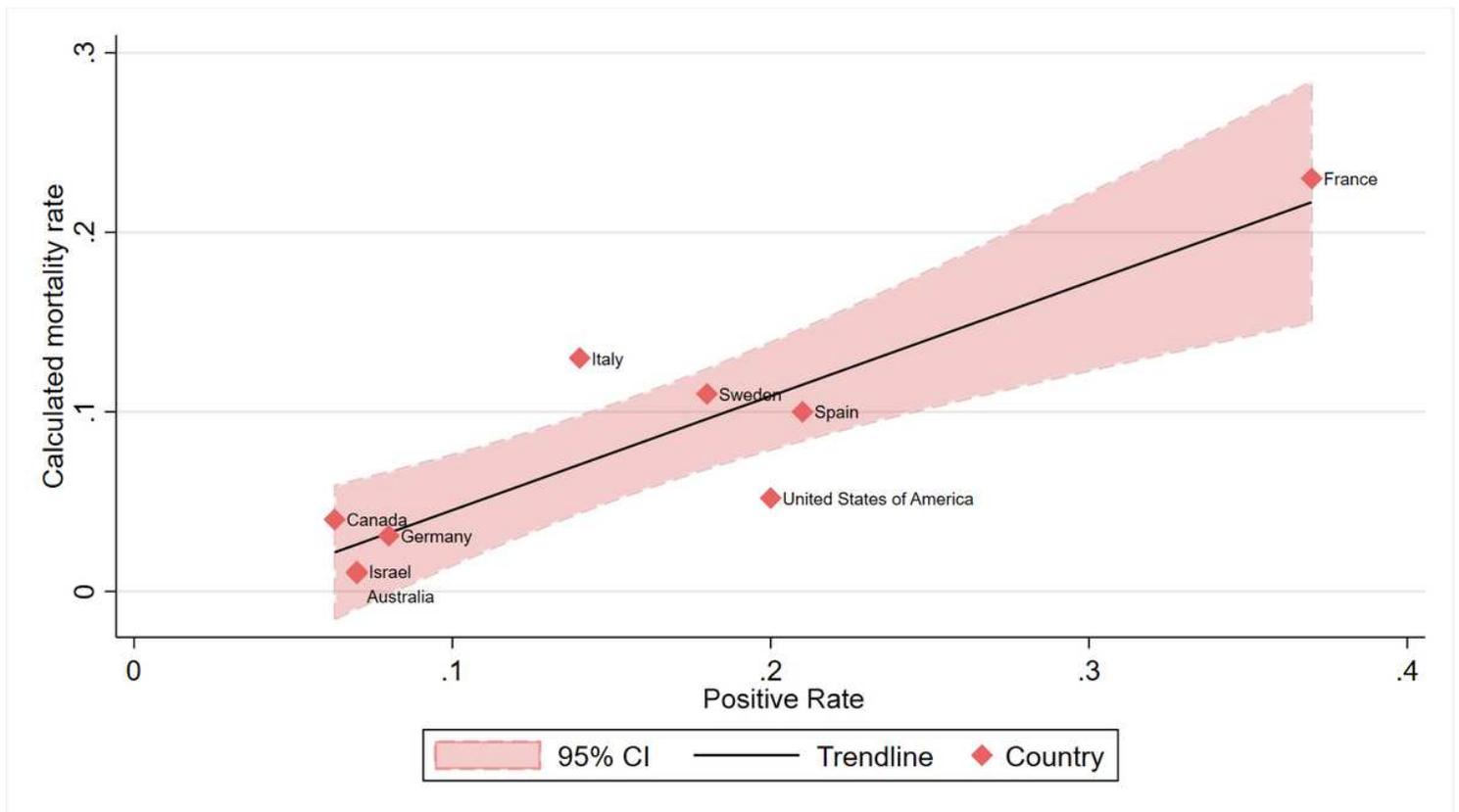


Figure 1

The effect of social awareness on effective contact frequency



**Figure 2**

The trend of calculated mortality rate and positive rate in Israel, Australia, Canada, Germany, Italy, Sweden, the United States of America, Spain, and France

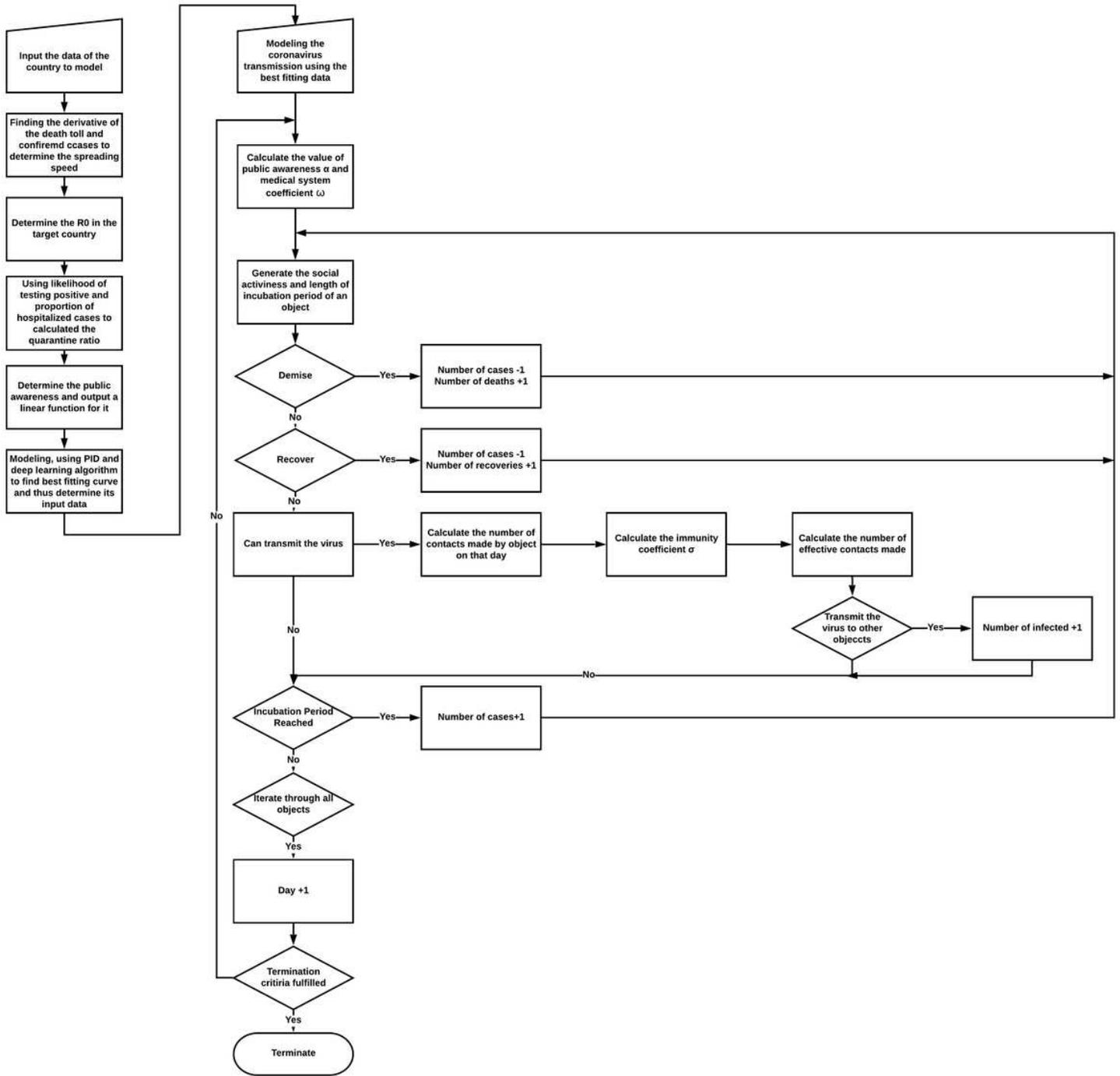
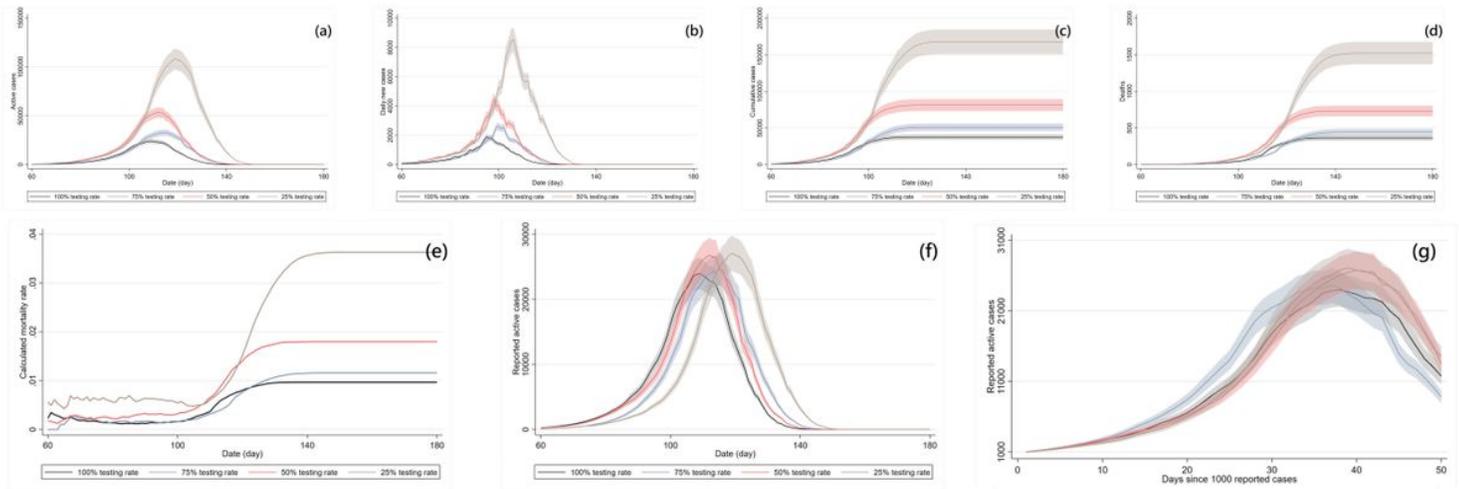


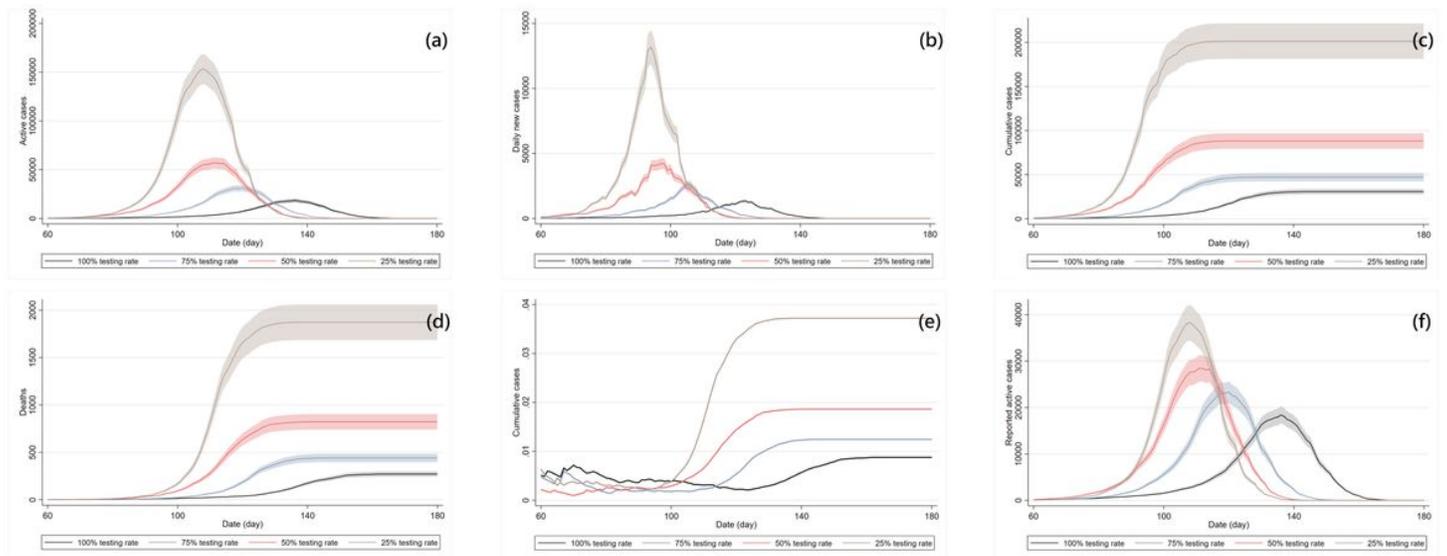
Figure 3

The flow chart of the OO-SIR model



**Figure 4**

Demonstrative trails – spreading speed of the virus (a) number of active cases (b) number of cumulative cases (c) number of reported active cases (d) calculated mortality rate (e) number of reported active cases (f) convergence of reported active cases (g) with 100% testing rate, 75% testing rate, 50% testing rate, 25% testing rate without quarantine and same trigger of city lockdown



**Figure 5**

Demonstrative trails – spreading speed of the virus (a) number of active cases (b) number of cumulative cases (c) number of reported active cases (d) calculated mortality rate (e) number of reported active cases (f) with 100% testing rate, 75% testing rate, 50% testing rate, 25% testing rate without quarantine and same trigger of city lockdown