Assessing the Global Tendency of COVID-19 Outbreak

Qinghe Liu[1], Zhicheng Liu[1], Junkai Zhu[1], Yuhao Zhu[2],

Deqiang Li[1], Zefei Gao[1], Liuling Zhou[1], Junyan Yang[3], Qiao Wang[1,4]

[1] School of Information Science and Engineering, Southeast University, Nanjing, China

[2] Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, the Netherlands

[3] School of Architecture, Southeast University, Nanjing, China

[4] Shing-Tung Yau Center, Southeast University, Nanjing, China

*Correspondence to Q. Wang (Email: qiaowang@seu.edu.cn)

March 24th, 2020

ABSTRACT

COVID-19 is now widely spreading around the world as a global pandemic, whereas it had been effecitvely under control in China by March 15. In this report, we estimate the global tendency of COVID-19 and analyze the associated global epidemic risk, given that the status quo is continued without further measures being taken.

The results show that the global R_0, excluding China, is estimated to be 2.48 (95% CI: 1.60 – 4.33). The United States (The U.S.), Germany, Italy, Spain, have peak values over 50,000. And the peak arrival time of all these countries is after Apr. 15, 2020. We predict that there are four regional clusters of the outbreak: Southeast Asia extending southward to Oceania, the Middle East, Western Europe and North America. Among them, Western Europe will become the major center of the outbreak. The peak values in Germany, Italy and Spain are estimated to be 98,893, 156,857 and 65,082, respectively. The U.S. is the country with the most serious outbreak trend. Based on the current situation of intervention measures, the peak value in the U.S. will reach 267,324. Above all, if the current policy environment is maintained, the cumulative number of patients worldwide will be 1,225,059 (95% CI: 486,934 – 4,533,392).

Keywords: COVID-19; SARS-Cov-2; SEIR model, Global epidemic.

1. INTRODUCTION

At the end of 2019, the first case of the novel coronavirus pneumonia (NCP) was confirmed in Wuhan, Hubei Province, China, caused by a new type of coronavirus named "2019 new coronavirus [1] (SARS-CoV-19)" [2]. The disease spread across China during the traditional Chinese Spring Festival. As of 24:00 on March 14, China had passed the peak of the epidemic and started to recover gradually, with a total of 80,844 confirmed cases and 3,199 deaths reported [3]. However, new cases of disease began to appear in other parts of the world and increases rapidly. Because of the rapid spread of the disease, the World Heath Organization (WHO) has assessed COVID-19 as a pandemic [4], which is the first pandemic caused by a coronavirus. As of March 14, 2020, the cases of COVID-19 have been reported in 135 countries and territories worldwide, with a total of 142,539 confirmed cases (61,518 outside of China) and 5,393 deaths (2,199 outside of China) [5]. Among them, a total of 17,660 cases were confirmed in Italy, 11,364 in Iran, 8,086 in Republic of Korea, 1,678 in the U.S., furthermore, Europe became the epicenter of the pandemic, with more reported cases and deaths than the rest of the world combined, apart from China [6].

With the outbreak of COVID-19, countries or territories have been taking different measures to cope with the spread of the pandemic, but the number of infectious people is still increasing. Meanwhile, the development of the pandemic has also caused a huge impact on the trade flows and economic development of the world. These situations raised many urgent problems. How will the epidemics develop in countries or territories around the world? When will the spread of epidemic arrive the peak or turn to stabilize? How many people will be at risk of infectious? Therefore, it is of great significance to analyze the trend of COVID-19 and predict the arrival of peaks for the prevention and control of the pandemic all over the world .

Based on the global data [7] reported by national authorities, we perform an analysis on epidemic status in 118 countries. Section 2 briefly discusses the most up-to-date literature related to COVID-19 epidemic status. Section 3 describes SEIR model in details. Section 4 presents the simulation and analysis of the model. Section 5 concludes.

2 RELATED WORK

On March 1, 2020, Li Y *et al* [8], used the outside-China diagnosis number released by WHO and built a mathematical model to capture the global trend of epidemics outside China, and they found that 34 founder patient outside China were not found, the worldwide epidemic trend is approximately exponential, and may grow 10 folds every 19 days.

On March 2, 2020, Zhuang Z *et al* [9], used a stochastic model to simulate the transmission process of South Korea and Italy under two corresponding assumptions of exponential growth periods, the results indicated that the reproductive number of the Republic of Korea and Italy are 2.6 (95% CI: 2.3-2.9) and 3.3 (95% CI: 3.0-3.6), respectively. In addition, the estimation of dispersion term were conducted, which indicated that there are fewer super-spreading events.

On March 8, 2020, Zhang Z *et al* [10], applied a data-driven coding method for the prediction of the global spreading epidemic of COVID-19. Based on the historical epidemiological data and the sets of parameters of augmented SEIR model in 367 cities of China, they selected the best fit profiles to predict the trend of epidemic with any population by comparing the given an early epidemic and the historical profiles. In their study, the peaks of infectious cases in South Korea, Italy, and Iran are expected to occur at the end of March, and the percentages of population infectious will less than 0.01%, 0.05% and 0.02%, respectively.

On March 10, 2020, Zhang Z *et al* [11], used the state transition matrix model to predict the epidemic trend of South Korea, Italy and Iran. By matching and fitting different scenarios, the inflection point arrival time will be March 6-12 for South Korea, March 10-24 for Italy and March 10-24 for Iran, and the cumulative number of cases will reach 20k in South Korea, 209k in Italy and 226k in Iran, respectively.

On March 14, 2020, Li L *et al* [12], made a propagation analysis and the wordwide prediction of COVID-19, and they also realized the backward inference of the epidemic starting day. The main results show that the epidemic in South Korea will be basically under control at the end of March, and the inflection day is January 7, before the control the reproductive number is 4.2 and 0.1 after the control. Epidemic size in Italy and Iran will reach 200,000 and 20,000, respectively at the end of March, and the inflection starting day are both on January 13, the reproductive number are in decline from 4.2, 4.0 to 0.1, 0.2. And their work could be a reference to our prediction results.

On March 17, DJ Darwin R *et al* [13], used the progression of the epidemic curve and the defined frame work to estimate the instantaneous reproductive number combined with the whole genome sequencing, wich can provide the genomic evolution and variation in the context of the outbreak dynamics. They divided the epidemic into different serial interval scenarios, and estimated the R values. The R value in Japan, Germany, Spain, Kuwait and France are over 2, in Italy, Iran and South Korea are even over 10, however, the R will be low after the social distancing intervention.

In our previous work [14], we predicted the time and value of inflection point and peak point in both Hubei and outside Hubei of China. Inspired by the recent studies and our precious works, we conduct a worldwide prediction of the time and value of the peak point, which is essential for the containment of the epidemic spread of COVID-19. The predicted results are based on the current medical situation, if there is specific medicine for COVID-19 in the future, the time and value of peak point of all countries will be advanced and shrinking compared with none-medicine situation.

3 COVID-19 TRANSMISSION MODEL: SEIR

3.1. SEIR Model

The SEIR model was firstly proposed by Aron and Schwartz in 1984 [15], where S, E, I and R referring to four different populations in epidemiology: the Susceptibles, Exposed, Infectious, and Recovered, respectively [16][17]. As the research object of this paper is more than 100 countries or regions, it is not the same as the previous study on China's epidemic [14]. In the study of China's epidemic situation, many factors such as Spring Festival, control policies and so on are taken into account, but in contrast, this paper will analyze the heterogeneity of each country or region on the basis of the common spread of infectious diseases. Although these countries have different policies and customs, their basic spread pattern of the epidemic still meets the principle of dynamics [18]. The epidemic impact of current attitudes and measures on the future trend can also be seen from the data. In this paper, SEIR model which has a strong generalization ability is used. The main assumptions of this model are as follows:

(1) In the predicted time scale, the existing influences of policies or culture of the target object will be unchanged.
(2) During the transmission of virus, the probability of infection in contact with each person is equal in the same target group.

$$\begin{cases} \frac{dS}{dt} = -\beta IS \\ \frac{dE}{dt} = \beta IS - (\alpha + \gamma_1) E \\ \frac{dI}{dt} = \alpha E - \gamma_2 I \\ \frac{dR}{dt} = \gamma_1 E + \gamma_2 I \end{cases} \tag{3.1}$$

where t means time, and S, E, I, R represent 4 different state variables, respectively. Formula (3.1) can be represented by the state transition diagram in Figure 1. All parameters and physical interpretations of the model are shown in Table 1.

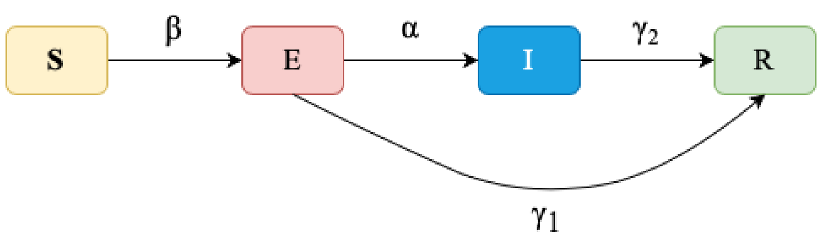

Figure 1. SEIR system state transition diagram.

Table 1. List of Symbols

State variable	S	Number of susceptible people
	E	Number of exposed people
	I	Number of infectious people
	R	Number of recovered people
Transfer variables	$\boldsymbol{\beta}$	Rate of those susceptible and being infected.
	$\boldsymbol{\alpha}$	Rate of those exposed and being infectious.
		Transmission rate to the recovered from the exposed.
		Transmission rate to the recovered from the infectious.

3.2 Practical Consideration

We divided the spread of the COVID-19 into three main phases according to the time and the size of the confirmed cases, and the schematic diagram of each stage for the number of patients is shown in Figure 2.

Phase one: the time range of this phase starts from the beginning of SARS-CoV-2 transmission and ends when the number of patients begins to increase significantly. COVID-19 spreads during incubation period at this phase, but the confirmed cases officially counted remains at a low level, and the public's awareness of defense is light.

Phase two: the main feature is that a large number of patients have been diagnosed, and the statistical data of confirmed cases increase significantly. But there still has no policy intervention and large-scale quarantine measures.

Phase three: as the intensity of policy intervention increases, the spread of COVID-19 is effectively suppressed, and growth rate of confirmed cases is gradually controlled until the end of the epidemic. The main feature of this phase is a strong human intervention.

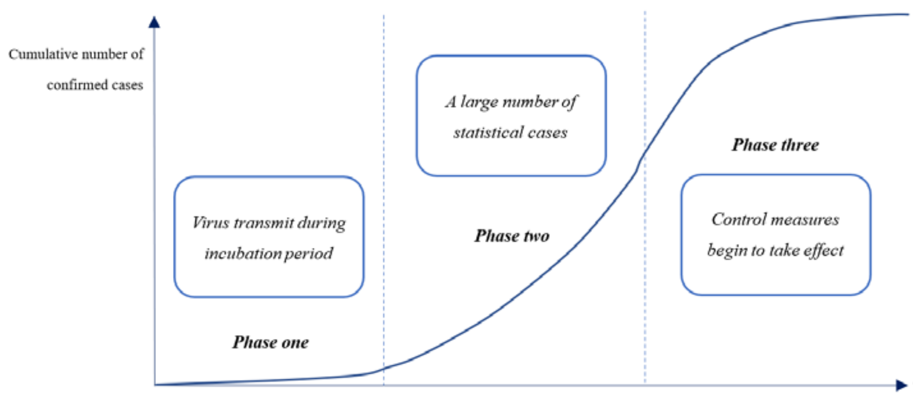

Figure 2. Phase diagram of confirmed cases changing with time in a certain region.

It is crucial to get as much information as possible from the data in three stages. Because this paper aims at the analysis of the peak size and its arrival time in each country or region, two latter transmission phases will be our focus in this paper. In phase two, SEIR will be used to analyze the trend of the epidemic, while in phase three, the simulation of policy intervention will be carried out in combination with the prior information of China studied in [14].

3.3. Estimation of R_0

The basic reproduction number, R_0, refers to the average value of how many people an infected person can transmit the virus to through natural transmission without external intervention [19]. From the second generator approach in [20], the equation (3.2) is derived as follows:

$$R_0 = \beta\left(\frac{1}{\alpha+\gamma_1} + \frac{\alpha}{(\alpha+\gamma_1)\gamma_2}\right) \tag{3.2}$$

Where α, β, γ_1, γ_2 are transfer variables of SEIR model. Considering that the calculation of R_0 is in the case of natural transmission, so we use the parameters of the starting time of epidemic spread in each country or region to estimate R_0 to avoid the impact of human intervention.

3.4. Data Source and Processing

In this study, we use open dataset from World Health Organization, BNO News, National Health Commission of China and Github. The details of data source are shown in Appendix 1.

World Health Organization (WHO). WHO provides timely official information of COVID-19, which includes official name, epidemic analysis, and situation reports. The number of confirmed and death cases are from daily situation reports.

BNO News. BNO News provides confirmed cases of coronavirus as well as recovered cases, and the data is updated throughout the day.

National Health Commission of the People's Republic of China. The national health commission of the People's Republic of China provides daily information of the epidemic situation in different parts of China, which is officially released as public data, including local real-time diagnosis data, suspected data, cure data *et al*.

Knowledge Centre on Migration and Demography of European Commission. European Commission provides global transnational migration data in Dynamic Data Hub. For a detailed description of data, please refer to the research of E. Recchi, *et al* [14].

Recovered cases data for each country or region was obtained from BNO News (https://bnonews.com/index.php/2020/02/the-latest-coronavirus-cases/). However, there are missing values. For an observation with a missing value, if the previous and the next data points are non-missing and the difference between the values is small than 10, we fill in the missing value by taking the average values. Otherwise, this observations is discarded.

4. MAIN RESULTS OF SEIR MODEL

4.1. Estimate of R_0

While estimating the R_0 of the epidemic globally, we ignore the countries whose ultimate cumulative number of confirmed cases are less than 100 by Mar. 22, 2020 and focus on 64 countries out of 118 (please refer to appendix). The average basic reproduction number of these countries is 2.48 (95% CI: 1.60 – 4.33). As we can notice, COVID-19 is lashing a great part of the world, and the sufferings are not likely to end soon. By estimating the average value of R_0 , we obtain the objective information of current situations as well as the importance of epidemic prevention.

Among these 64 mainly researched countries, the R_0 of five countries are higher than 3.60, suggesting the great potential of a destructive outburst. U.S. is one of these countries with a basic reproduction number of 6.25, much higher than the global average. In fact, the U.S. had a large outbreak of influenza before testing for SARS-CoV-2 (See Appendix 2 for more details). The basic reproduction number of the U.S. was followed by Germany, Spain and Italy, reflecting the severe epidemic status in European countries and confirming that they have become the epicenter of COVID-19 outbreaks. The lowest value of R_0 is 1.44 in Brazil, still greater than 1.

4.2. Estimate of the Peak of Confirmed Cases in Country-Level

The peak value is defined as the cumulative number of confirmed cases. After this peak point, the temporary number of cases gradually decrease and the epidemic subsides **Error! Reference source not found.**. The epidemic trend of 118 countries are first simulated all over the world, and the peak number of confirmed cases and its arrival time in each country are predicted (See Appendix 2 for details). According to different peak value intervals, these countries are divided into six different categories. Among these categories, U.S. and Italy, whose peak values are higher than 100,000 based on simulation. 21 countries including German, Spain, Turkey and France have peak values higher than 10,000. And the peak values of most of these contries come after April, 11, 2020. 34 countries have peak values ranging from 1000 to 10,000, and all of their peak values arrive after April 18, 2020. Romania and other 33 countries have peak values ranging from 100 and 1,000. 27 countries have peak values between 10 and 100. And 9 countries have peak values under 10.

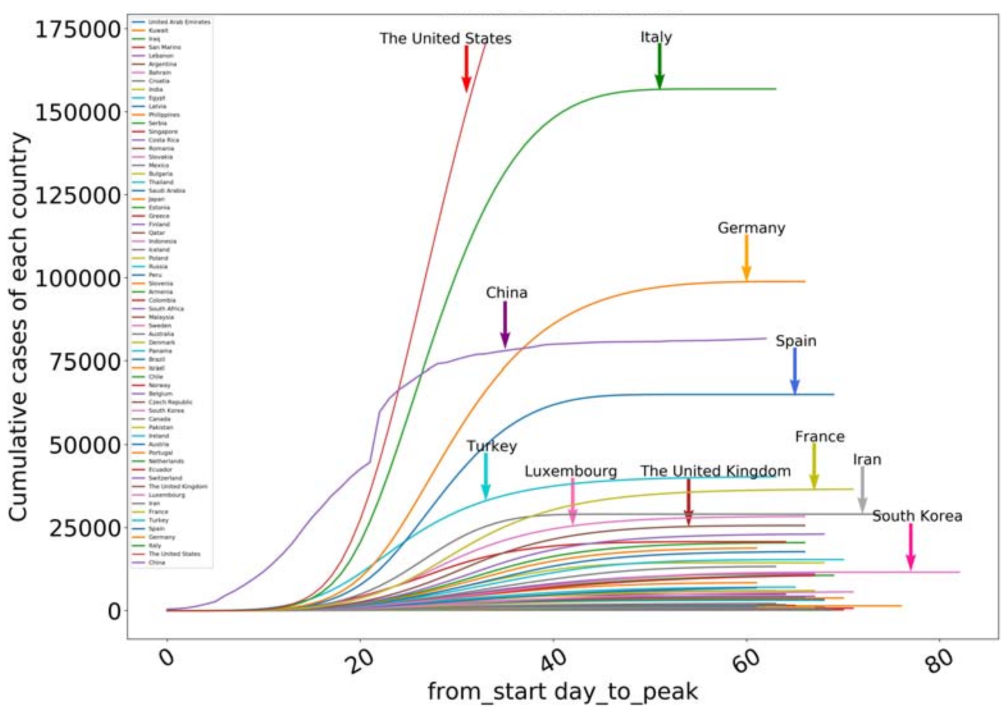

Figure 3. Comparisons between the estimated trends of cumulative number of confirmed cases in 64 countries or regions, which has the number of confirmed cases exceeding 100, as of March 21, 2020. X-axis is days since the outbreak, and y-axis is the number of confirmed cases.

U.S. is the country with the highest exposure to the risk of the coronavirus.

According to our estimation, the peak number of real-time confirmed cases of U.S. is predicted to reach 267,324, and its arrival time is on July, 21, 2020. It is the only country in our prediction where the number of confirmed cases exceeds 200,000. It should be suggested that the U.S. should take effective measures to overcome the epidemic as soon as possible.

Germany, Italy and Spain monopolize one category, out of all 118 countries, due to its high risk of the coronavirus. The peak number of real-time confirmed cases of Italy is predicted to reach 78,237 according to our estimation. As one of the worst-hit countries, Italy has already attracted the world's attention. The president of Italy issued a national blockade on March 10, however, the daily number of new confirmed cases continued to rise. The other 18 countries whose peak values are estimated to reach over 10,000, such as the France, Iran, United Kingdom and South Korea, are all in severe situation at present. Poland, Brazil and Australia and several other countries are the next worst-hit areas with their peak value over 1000.

Furthermore, 36 of all 118 countries will not reach the peak value of 100, and 58.3% of 118 countries will have the peak value under 1000. However, 18.6% of the countries are suffering a lot that they will have a peak value over 10,000. We are supposed to help each other to overcome the hardship. We compared peak arrival time in countries whose peak value are higher than 500. As we can see, most of the countries will reach their peak value in May.

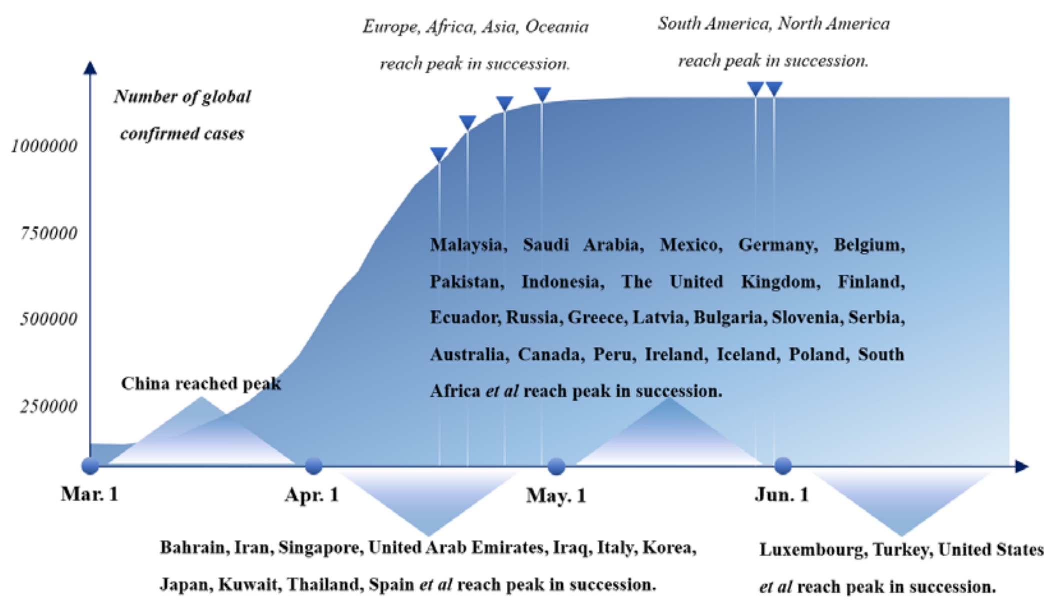

Figure 4. The peak time of the number of confirmed cases in majar countries and 6 continents, as well as the global trend chart of the cumulative number of confirmed cases. X-axis is date, and y-axis is the global number of confirmed cases.

4.3. Estimate of the Peak of Confirmed Cases in Continent-Level

It can be observed in Table 2 that Africa has the the minimum peak value and the second earliest peak arrival time. There are only two countries in Oceania for our study, the sparse population makes it earlier to reach the peak value quickly. It is estimated that Asia's peak value would reach 153,174 (95% CI: 135,640 – 223,879). As of March 22, 2020, the daily increment of confirmed cases only in the U.S. is 4763. This also led to an outbreak trend in North America, which is estimated to reach a peak of 283,128 (95% CI: 189,394 – 632,844) on May. 30 (95% CI: May. 5 - July. 15) under the current policy intervention. Europe has the highest peak value indicating the severe situation in Europe. South America also has the second latest peak-arrival time but a much lower peak value than North America.

Table 2. Sum of peak values and average peak arrival time of different continents

Continents	Peak value	Peak time (in 2020 year)
Europe	555756 (95% CI: 438675 - 982349)	Apr. 15 (95% CI: Apr. 11 - June. 8)
Asia	153174 (95% CI: 135640 - 223879)	Apr. 22 (95% CI: Apr. 7 - May. 6)
North America	283128 (95% CI: 189394 - 632844)	May. 30 (95% CI: May. 5 - July. 15)
South America	54821 (95% CI: 38924 - 93248)	May. 28 (95% CI: May. 3 - June. 8)
Oceania	6060 (95% CI: 3924 - 13299)	Apr. 28 (95% CI: Apr. 26 - May. 3)
Africa	5210 (95% CI: 3294 - 9234)	Apr. 19 (95% CI: Apr. 16 - May. 14)

4.4. Geographical Distribution Analysis of COVID-19

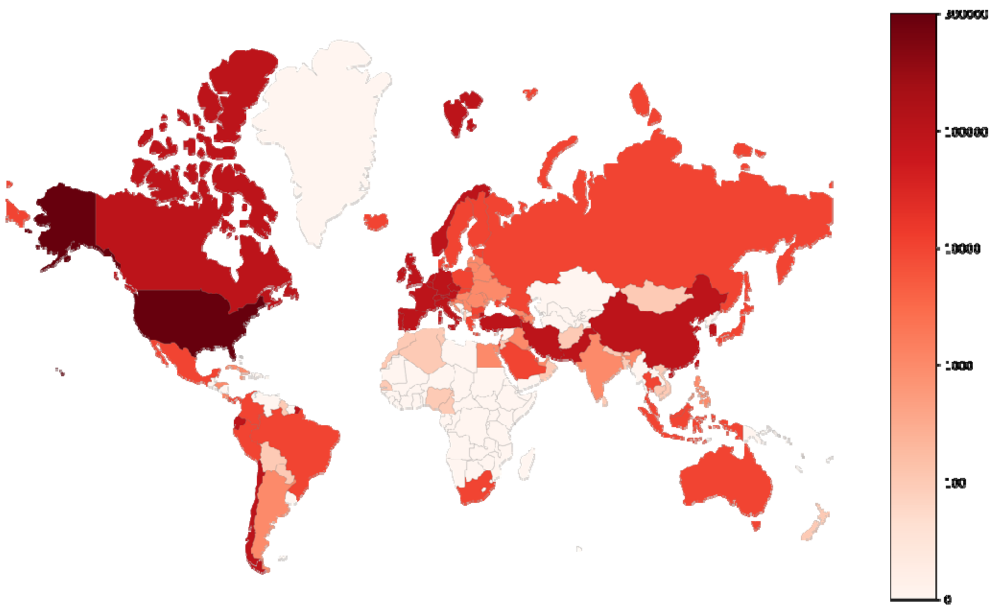

Figure 5. Peak value of confirmed cases in 118 countries over the world. In this world map, the same set of color coding is used for different actual geographical locations. Red means the higher the diagnosis cases peak, while white means the lower the diagnosis cases peak.

The predicted final epidemic scale is visualized on the world map to show the geographical characteristics of the predicted epidemic peak of confirmed cases. According to [21], geographic proximity transmission is very prominent in epidemic transmission. That is to say the phenomenon of regional transmission is popular. This opinion can also be confirmed on the map shown in Figure 5. The outbreak shows multiple clusters, and has a more obvious near aggregation effect. But we can still find that North Korea is geographically close to South Korea, Japan and China, which has the worst situation. The neighboring countries of North Korea are suffering from severe epidemics, but North Korea has no confirmed cases until now. This is because at the time of the outbreak in China in January 2020, North Korea adopted strict control and management. The North Korean government blocked all entry borders, not allowing any people to travel home and abroad.

The choropleth map indicates that the spread of COVID-19 will form four major regional clusters. The first cluster ranges from Southeast Asia to Oceania, including China, South Korea, Japan, Indonesia, the Philippines, Australia, etc. South Korea, Japan, Indonesia, Malaysia, and the Philippines will all have peak values of more than 1,000. Water transportation in this area is well developed, with frequent commercial exchanges and more opportunities for people to gather, which may accelerate the spread of COVID-19. Among these cases, the Diamond Princess incident is the most typical one. The effect of epidemic spread caused by gathering behaviors such as cruises is likely to be much higher than the R_0 level in general sense [21]. As Mar. 22, 2020, 712 people have been diagnosed with the disease on the diamond princess. In the case of unknown diseases, patients will spread the virus to more susceptible groups.

The second outbreak cluster occurs in Western Europe, including Spain, Italy, France, Poland, Germany, and many other neighboring countries. Germany, Italy and Spain are estimated to have 98,893, 156,857 and 65,082 confirmed cases, respectively, given that the current policies are not changed. France, Luxemburg, Switzerland, Netherlands are all expected to have peak values of over 20,000. Considering the number and density of European countries, there would be a serious cross-border phenomenon. As the first outbreak center, China has a population of about 1.4 billion, and the number of patients in China has basically reached the relative saturation level, because the increment of confirmed cases in China is less than 100 more than one week. As of March 22, 2020, the cumulative number of patients is 81,649 in China. Under the current transmission environment and government measures, the estimated number of patients is likely to be 559,356, with 740 million population in Europe. Based on the above data, the per capita prevalence rate in Western Europe will be at least 11.96 times higher than that in China.

The third outbreak cluster is in the Middle East. According to the current trend, the peak number of confirmed cases in Iran will reach 29051. According to the current trend, the epidemic situation in Iran has gradually slowed down, and the number of confirmed cases is expected to reach peak around April 7, 2020. In addition, the peak values of Qatar and Saudi Arabia are both estimated to be over 1,000. Estimated by the current situation, the peak of diagnosis in the Middle East will come in early May this year.

The last outbreak luster occurred in North America. Up to now, although there's no direct evidence to prove that wearing masks can inhibit the spread of the epidemic, the phenomenon of crowd gathering in airports and other places in the U.S. still exists, and many people have not taken any isolation and protection measures. Given that the epidemic has spread in the U.S. (35,209 cases in total, as of March 22, 2020), if the current situation is maintained, the peak number of cases in the U.S. will reach 267,324, which may have a huge impact on American people's livelihood and social economy. If the U.S. follows the existing policy, the peak arrival time of COVID-19 in the U.S. will be much later than that of global epidemic.In addition, Canada is expected to have a peak value of 13,336, making it the second most affected country in the region.

It can be seen clearly that the most countries or regions where the outbreak is expected to occur has relatively developed trade system and their population flows are relatively large. Timely isolation measures are necessary and the epidemic situation is urgent. The cumulative number of global patients will finally be 1,225,059 (95% CI: 486,934 – 4,533,392).

5. CONCLUSION

We performed simulation analysis of COVID-19 based on SEIR model. From analysis, we draw the following conclusions:

(1) We collected data before March 22, 2020 to simulate the natural transmission of the SARS-CoV-2 in 118 countries. The average basic reproductive number R_0 of these countries was estimated to be 2.48 (95% CI: 1.60 – 4.33).

(2) 23 countries will have a peak value over 10000, including France, Luxemburg, The United Kingdom, Switzerland *et al*. And 50 of the 118 countries will have peak values lower than 100. 20 of 118 countries or regions, including Bahrain, Iran, Singapore *et al*, will reach peak value of confirmed cases in April, 2020. 39 countruies or regions will peak in May, 2020. Sweden, Luxembourg, Turkey, Brazil and U.S. will be saturated after Jun. 1, 2020.

(3) Among the six continents, Europe has the highest peak value of 555,756 (95% CI: 438,675 – 982,349), and the average peak-arrival date of Europe is Apr. 15 (95% CI: Apr. 11 - June. 8). The lowest peak of 5,210 (95% CI: 3,294 – 9,234) appears in Africa, which will arrive on Apr. 19 (95% CI: Apr. 16 - May. 14). Asia has a peak value of 153,174 (95% CI: 135,640 – 223,879), and China accounts for about 53% this value. (see Table 2 for more details.)

(4) Western Europe, Southeast Asia to Oceania, the Middle East, and North America will be the four epicenter with the most severe situation. The epidemic situation in North America and Western Europe are estimated to far exceed that in China.

6. DISCUSSION

In this paper, we mainly estimated and analyzed the transmission trend of COVID-19 based on the daily updated data of confirmed cases and recovered cases.

In fact, the trend of COVID-19 is related to multiple factors. China's prior data were used in the third phase of the epidemic (see Figure 2), however, the testing policies and standards in each country and region are different, which is highly related to the governance system and safety consciousness. In addition, in the second phase of epidemic transmission, human-to-human transmission plays an important role. The intensity of people flow directly affects the infection degree of the virus. Some factors, such as subway traffic, catering facilities, customs or habits of each region remains to be researched.

It should be noted that because the analysis above is based on that the official statistics of confirmed cases and of recovered cases are true in a unified environment. As a matter of fact, there may be deviations in the norms of statistical data in various countries, which will also lead to errors between our conclusions and the real situations.

Abbreviations

COVID-19: Coronavirus Disease 2019; SARS-CoV-2: Severe Acute Respiratory Syndrome Coronavirus 2; CI: Confidence Interval; SEIR: Susceptible-Exposed-Infectious-Recovered.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The datasets used and analyzed during the current study is available from open resources.

Competing Interests

The authors declare that they have no conflict of interest.

Funding

This work was not supported by any funding.

Authors' Contributions

Conceived and designed the experiments: Qinghe Liu, Junkai Zhu, Junyan Yang, Qiao Wang.

Performed the mathematical modelling: Qinghe Liu, Junkai Zhu, Qiao Wang.

Analyzed the data: Qinghe Liu, Junkai Zhu, Zhicheng Liu, Yuhao Zhu.
Collect the data: Zefei Gao, Deqiang Li.
Wrote the paper: Junkai Zhu, Qinghe Liu, Deqiang Li, Liuling Zhou, Zefei Gao, Zhicheng Liu, Yuhao Zhu.
All authors read and approved the final manuscript.

Acknowledgments
Not applicable.

Reference

[1] 'Definition of suspected cases of unexplained pneumonia', the National Health Commission of the People's Republic of China (in Chinese) /[http://www.nhc.gov.cn/]

[2] Li Q, Guan X, Wu P, *et al.* Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia[J]. The New England Journal of Medicine, NEJM.

[3] 'Update on COVID-19 as of 24:00 on March 14' (in Chinese)/[http://www.nhc.gov.cn/]

[4] 'WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020'/[https://www.who.int/]

[5] WHO, Novel coronavirus (2019-nCoV) situation report 54. World Health Organization, 2020. /[https://www.who.int/]

[6] 'WHO Director-General's opening remarks at the media briefing on COVID-19 - 13 March 2020'/[https://www.who.int/]

[7] WHO, Novel coronavirus(2019-nCoV) situation report. World Health Organization, 2020. /[https://www.who.int/]

[8] Li Y, Liang M, Yin X, *et al.* COVID-19 Epidemic Outside China: 34 Founders and Exponential Growth. medRxiv 2020:2020.03.01.20029819.

[9] Zhuang Z, Zhao S, Lin Q, *et al*. Preliminary estimating the reproduction number of the coronavirus disease (COVID-19) outbreak in Republic of Korea and Italy by 5 March 2020. medRxiv 2020:2020.03.02.20030312.

[10] Zhan C, Tse CK, Lai Z, *et al*. Prediction of COVID-19 Spreading Profiles in South Korea, Italy and Iran by Data-Driven Coding. medRxiv 2020:2020.03.08.20032847.

[11] Zheng Z, Wu K, Yao Z, *et al*. The Prediction for Development of COVID-19 in Global Major Epidemic Areas Through Empirical Trends in China by Utilizing State Transition Matrix Model. medRxiv 2020:2020.03.10.20033670.

[12] Li L, Yang Z, Dang Z, et al. Propagation analysis and prediction of the COVID-19. medRxiv 2020:2020.03.14.20036202.

[13] Bandoy DJDR, Weimer BC. Pandemic dynamics of COVID-19 using epidemic stage, instantaneous reproductive number and pathogen genome identity (GENI) score: modeling molecular epidemiology. medRxiv 2020:2020.03.17.20037481.

[14] Qinghe Liu, Zhicheng Liu, Deqiang Li, Zefei Gao, Junkai Zhu, Junyan Yang, Qiao Wang, Assessing the Tendency of 2019-nCoV (COVID-19) Outbreak in China, medRxiv, 2020.02.09.20021444.

[15] Aron JL, Schwartz IB. Seasonality and period-doubling bifurcations in an epidemic model[J].Journal of theoretical biology, 1984, 110(4): 665-679.

[16] Li MY, Muldowney JS. Global stability for the SEIR model in epidemiology[J].

Mathematical biosciences, 1995, 125(2): 155-164.

[17] Li MY, Smith HL, Wang L. Global dynamics of an SEIR epidemic model with vertical transmission[J]. SIAM Journal on Applied Mathematics,2001,62(1):58-69.

[18] Hethcote H W . The Mathematics of Infectious Diseases[J]. SIAM Review, 2000, 42(4):599-653.

[19] Dietz, K. The estimation of the basic reproduction number for infectious diseases[J]. Statistical Methods in Medical Research, 1993, 2(1):23-41.

[20] G. Chowell, P.W. Fenimore, M.A. Castillo-Garsow, C. Castillo-Chavez. SARS outbreaks in Ontario, Hong Kong and Singapore: the role of diagnosis and isolation as a control mechanism [J]. Journal of Theoretical Biology, 2003, 224(1): 1–8.

[21] Riley S , Fraser C , Ca. D , *et al*. Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions[J]. Science, 2003, 300(5627):p. 1961-1966.

Appendix 1 Data Source (All data used in paper is public.)

Reported Cases Data:
https://www.who.int/
https://bnonews.com/index.php/2020/02/the-latest-coronavirus-cases/
http://www.nhc.gov.cn/

National-Level Migration Data:
https://bluehub.jrc.ec.europa.eu/migration/app/

Appendix 2 Simulation Results in 118 Main Countries or Regions

Data used as of March 22, 2020. All countries or regions are in alphabetical order. 'none' indicates that the number of confirmed cases in the country or region as of March 22, 2020 does not exceed 100, so detailed analysis will not be conducted temporarily.

Country	Peak value	Peak time	Inflection point	R0
Afghanistan	[20, 100]	none	none	none
Albania	[20, 100]	none	none	none
Algeria	[20, 100]	none	none	none
Andorra	[20, 100]	none	none	none
Argentina	343	2020-04-29	2020-03-23	1.923618
Armenia	4075	2020-05-25	2020-04-05	2.777125
Australia	6049	2020-05-13	2020-03-31	2.431222
Austria	17827	2020-05-16	2020-04-03	2.636313
Azerbaijan	[20, 100]	none	none	none
Bahrain	356	2020-04-06	2020-03-13	2.11718
Bangladesh	<20	none	none	none
Belarus	[20, 100]	none	none	none
Belgium	11055	2020-05-05	2020-03-29	2.574309
Bhutan	<20	none	none	none
Bolivia	<20	none	none	none
Bosnia and Herzegovina	[20, 100]	none	none	none
Brazil	7378	2020-06-14	2020-04-03	2.496003
Brunei Darussalam	[20, 100]	none	none	none
Bulgaria	1104	2020-05-10	2020-03-30	1.831309
Cambodia	[20, 100]	none	none	none
Cameroon	[20, 100]	none	none	none
Canada	13336	2020-05-13	2020-04-04	2.448616
Chile	10682	2020-05-23	2020-04-06	2.660947
Colombia	4236	2020-05-21	2020-04-05	2.78129
Costa Rica	855	2020-05-20	2020-03-29	2.168627
Croatia	383	2020-04-23	2020-03-24	2.071571

Cuba	160	none	none	none
Cyprus	[20, 100]	none	none	none
Czech Republic	11510	2020-05-20	2020-04-04	2.47646
Denmark	6082	2020-05-20	2020-03-26	2.493841
Dominican Republic	720	none	none	none
Ecuador	20763	2020-05-09	2020-04-08	2.450982
Egypt	568	2020-04-22	2020-03-22	1.438881
Estonia	1465	2020-05-18	2020-03-26	2.293195
Faroe Islands	[20, 100]	none	none	none
Finland	1586	2020-05-08	2020-03-25	2.313709
France	36636	2020-05-21	2020-03-23	3.278523
French Guiana	<20	none	none	none
French Polynesia	<20	none	none	none
Georgia	[20, 100]	none	none	none
Germany	98893	2020-05-05	2020-03-27	4.622566
Gibraltar	<20	none	none	none
Greece	1578	2020-05-09	2020-03-23	2.309488
Guernsey	<20	none	none	none
Guyana	<20	none	none	none
Holy See	<20	none	none	none
Honduras	240	none	none	none
Hungary	[20, 100]	none	none	none
Iceland	2183	2020-05-14	2020-03-29	1.92386
India	491	2020-04-29	2020-03-22	1.684372
Indonesia	2030	2020-05-06	2020-03-28	2.361168
Iran	29051	2020-04-07	2020-03-18	3.602484
Iraq	224	2020-04-11	2020-03-14	1.65795
Ireland	15380	2020-05-14	2020-04-07	2.595479
Israel	8472	2020-05-20	2020-04-02	2.501636
Italy	156857	2020-04-14	2020-03-14	4.185207
Jamaica	160	none	none	none
Japan	1455	2020-04-18	2020-03-14	1.819315
Jersey	<20	none	none	none
Jordan	[20, 100]	none	none	none
Korea	11610	2020-04-18	2020-03-06	2.334726
Kuwait	217	2020-04-18	2020-03-13	1.532748
Latvia	591	2020-05-10	2020-03-27	2.150965
Lebanon	286	2020-04-25	2020-03-17	1.995054
Liechtenstein	[20, 100]	none	none	none
Lithuania	[20, 100]	none	none	none
Luxembourg	28508	2020-06-08	2020-04-06	2.734804
Malaysia	4955	2020-05-01	2020-03-28	2.480225

Maldives	<20	none	none	none
Malta	[20, 100]	none	none	none
Martinique	[20, 100]	none	none	none
Mexico	1082	2020-05-04	2020-03-29	2.219194
Moldova	[20, 100]	none	none	none
Monaco	<20	none	none	none
Mongolia	<20	none	none	none
Morocco	[20, 100]	none	none	none
Nepal	<20	none	none	none
Netherlands	20528	2020-05-16	2020-04-01	2.76451
New Zealand	[20, 100]	none	none	none
Nigeria	<20	none	none	none
North Macedonia	[20, 100]	none	none	none
Norway	10726	2020-05-21	2020-03-29	2.552265
occupied Palestinian territory	[20, 100]	none	none	none
Oman	[20, 100]	none	none	none
Pakistan	14450	2020-05-05	2020-04-07	2.660701
Panama	7142	2020-05-26	2020-04-07	2.131763
Paraguay	<20	none	none	none
Peru	3393	2020-05-13	2020-04-02	2.336978
Philippines	657	2020-04-27	2020-03-22	2.181275
Poland	3102	2020-05-15	2020-03-30	2.405632
Portugal	18932	2020-05-16	2020-04-07	2.791349
Qatar	1771	2020-04-30	2020-03-24	2.418145
Réunion	[20, 100]	none	none	none
Romania	936	2020-04-20	2020-03-24	2.270333
Russia	3148	2020-05-09	2020-04-01	2.414848
Saint Barthélemy	<20	none	none	none
Saint Martin	[20, 100]	none	none	none
Saint Vincent and the Grenadines	<20	none	none	none
San Marino	234	2020-04-20	2020-03-16	1.818849
Saudi Arabia	1333	2020-05-03	2020-03-27	2.298559
Senegal	[20, 100]	none	none	none
Serbia	686	2020-05-11	2020-03-26	2.096712
Singapore	739	2020-04-08	2020-03-14	2.005895
Slovakia	1011	2020-05-17	2020-03-28	1.929672
Slovenia	3800	2020-05-11	2020-03-31	2.375646
South Africa	4348	2020-05-15	2020-04-04	2.389561
Spain	65082	2020-04-20	2020-03-23	4.121961
Sri Lanka	[20, 100]	none	none	none

Sweden	5851	2020-06-08	2020-03-25	2.052677
Switzerland	23097	2020-05-16	2020-03-29	2.860396
Thailand	1114	2020-04-20	2020-03-25	2.174744
The United Kingdom	25646	2020-05-07	2020-03-30	2.970045
Tunisia	[20, 100]	none	none	none
Turkey	40434	2020-06-09	2020-04-04	2.848875
Ukraine	[20, 100]	none	none	none
United Arab Emirates	156	2020-04-11	2020-03-12	1.705958
United States	267324	2020-07-21	2020-03-30	6.252778
Viet Nam	[20, 100]	none	none	none