Coronavirus Pandemic

Quantitative analysis and mathematic modeling of the global outbreak of COVID-19

Yanyan Jiang[1#], Xuefeng Jiang[1#], Wenjun Tong[1], Jingming Zhou[2]

[1]*Jiangsu Collaborative Innovation Center of Biomedical Functional Materials, Jiangsu Key Laboratory of Biomedical Materials, School of Chemistry and Materials Science, Nanjing Normal University, Nanjing, China*
[2] *Jingling High School Hexi Campus, Nanjing, China*

[#] Authors contributed equally to this work.

Abstract
The coronavirus pandemic is the biggest in the past 100 years, affected over 200 countries and killed over 300 thousand people. To better understand the epidemics in different areas, the progress percentage was generated in this study by dividing everyday total confirmed case number by the up-to-date total case number, so data obtained from different countries and territories can be put together and compared directly regardless of the large difference in the magnitude of numbers. The global outbreak data were analyzed and categorized into 4 groups based on different epidemic curve stages. The grouping pattern suggests that the geographical position may not play a critical role in the progress of COVID-19 epidemic. In this report, we also used a mathematic model to predict the progress of COVID-19 outbreak in UK, USA and Canada in Group 3, providing valuable information for assessing the risk in these countries and the timing of reopening business.

Key words: COVID-19; epidemic; outbreak; progress percentage; projection.

Global glance of pandemic

The coronavirus pandemic has infected and killed more people than any epidemic disease globally in the past 100 years. The World Health Organization (WHO) declared the COVID-19 outbreak a global public health emergency in late January and defined it a pandemic in March 2020. The disease has spread to 215 countries, territories and international conveyances around the world, infected 5.73 million people, and caused 361 thousand death by 26 May 2020. Centers for Disease Control and Prevention (CDCs), emergency agencies, and research scientists have swarmed to studies on epidemic data analysis[1,2], virus property characterization [3], disease pathology [4], trend projection [5], comparisons with previous pandemics [6] and the current pandemic in different coutries [7] as well as necessary policies that help to stop the current pandemic and to prevent future outbreaks [6,8,9].

This study collected total confirmed case number by 26 May 2020 from countries that currently rank Top 20 or ever ranked Top 10 in the past and plotted against time. First, the total confirmed cases in individual countries were plotted (Figure 1A). USA has the largest total confirmed case number and accounts for almost one third of the global number. Since its total case number is much bigger than the other countries' and dwarfs their curves overwhelmingly, it is difficult to make a direct comparison between all selected countries. Thus, the selected countries excluding USA were plotted (Figure 1B) for better comparison.

Different stages of epidemic

To better compare the global epidemic curves, the progress percentage of each country or territory was generated in this study by dividing everyday total confirmed case number by the up-to-date total case number (shown in Figure S1-S4), so data from different countries can be compared directly regardless of the large difference in the magnitude of numbers (shown in Figure 2).

If we look at an epidemic curve, an epidemic in an area can be divided into 4 stages based on the increase of total confirmed cases: Stage 1, sporadic cases or limited local transmission, creating undulating curve; Stage 2, substantial local transmission, featuring a takeoff of the curve; Stage 3, emerging of large number of new cases, causing an exponential increase; Stage 4, fewer new cases reported than peak, leading to

flattening of the curve till the end of the epidemic. By using the progress in percentage plot, it is easy to compare the stages of different countries and categorized them into 4 groups based on different stages. Group1, including China, S. Korea, Italy, France, Germany, Switzerland, Austria, and Norway, are in the latter half of Stage 4 (Figure 3A). Group 2, including Spain, Turkey, Belgium, Netherlands, Japan, and Denmark, have successfully passed the peak and entered Stage 4 (Figure 3B). Group 3, including USA, UK, Iran, Canada, Ecuador, Sweden, and Singapore, are developing new cases steadily. Although some of them have primitive evidence showing that the peak has been reached, further observation is needed to confirm the transition (Figure 3C). It is worth noticing that Iran was in Group 2 before 23 April 2020, but recently moved to Group 3 due to the second outbreak, about 2 weeks after gradual reopening of business since 11 April 2020. The timing of the second outbreak strongly suggests a causal relation rather than simply a coincidence. Group 4, including Brazil, Russia, India, Peru, Chile, Saudi Arabia, Mexico, Pakistan, Qatar, and South Africa, are still in Stage 3 and haven't shown any evidence of approaching the peak yet. (Figure 3D).

Figure 1. Total confirmed cases in selected countries.
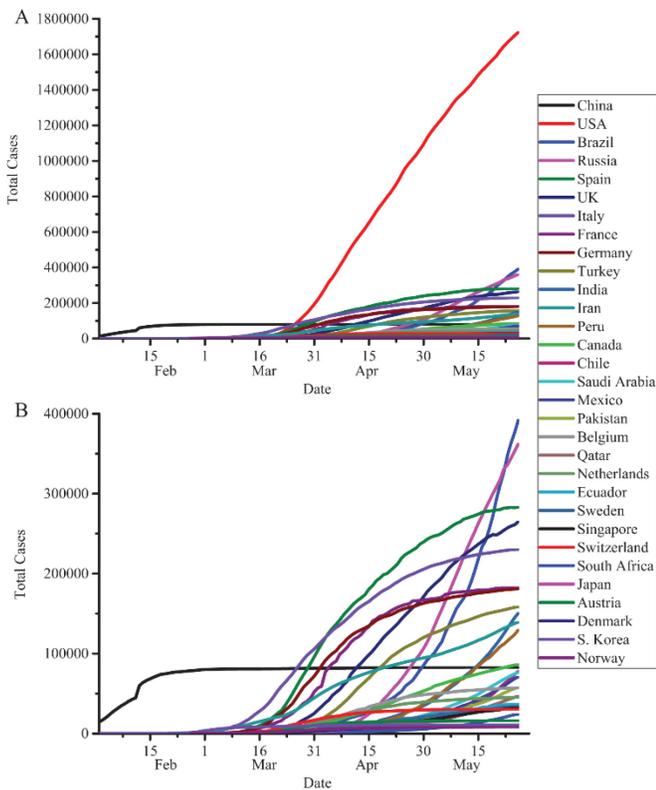
(A) with USA and (B) without USA.

Figure 2. The progress percentage in selected countries.
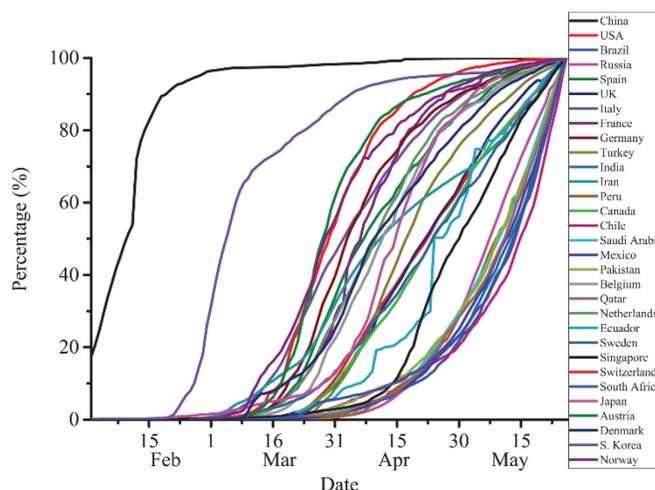
The grouping has revealed some interesting facts. In east Asia, Japan (in Group 2) has shown totally different epidemic progression despite the close geographic distance to China and S. Korea in Group 1. European countries are mostly included in Group 2, however, UK and Sweden are in Group 3. In Group 3, UK, USA, and Canada have shown almost identical trends despite that UK is geographically separated from the other two by Atlantic Ocean. All BRIC (Brazil, Russia, India and China) countries except China are in Group 4 and show similar epidemic trends. These facts suggest that the geographical position may not play a critical role in the epidemic progression in an area. The curves of UK, USA and Canada are almost identical at the beginning, possibly due to frequent travels. But after these countries have tightened their borders, their curves still resemble each other (Figure 3C), suggesting that similar cultural background, including the public response to social distance policy, as well as the position in global economy, or even the form of government may also contribute a lot to the progression of the epidemic in an area.

It is safe to say that the outbreak is under well control in Group 1 countries, and wouldn't post a big threat to Group 2 countries provided that protective measures are taken. For countries in Group 4, it is too early to make any predictions at this moment. But some Group 3 countries, such as USA, UK and Canada, are approaching the intersection of accelerating again or slowing down and passing the peak. Naturally, many people would be eager to know the projections in these countries, due to their critical importance in the global economy and personal attachment.

Projections of the disease in USA, UK and Canada

This study used a mathematical model to simulate the epidemic progression in USA, UK and Canada. We took China's curve as a reference (Figure 4). On 12 February 2020, China CDC published a new version of diagnostic standard, which caused a dramatic increase in the daily new case number, and affected the numbers of the following few days. It is obvious that except for this sharp peak, the decline of daily new cases after the peak which had been reached around 4 Feburary 2020 corresponds to a nearly straight line with relatively big negative slope, $y = 3874.6 - 164.56x$ ($R^2 = 0.9239$) where x is the number of days after 3 February 2020. The decline of daily new cases would oscillate along this line until a bottom is reached, which indicates the end of the epidemic. After the bottom was reached in March, there were occasional oscillation mainly caused by imported new cases with sporadic domestic new cases.

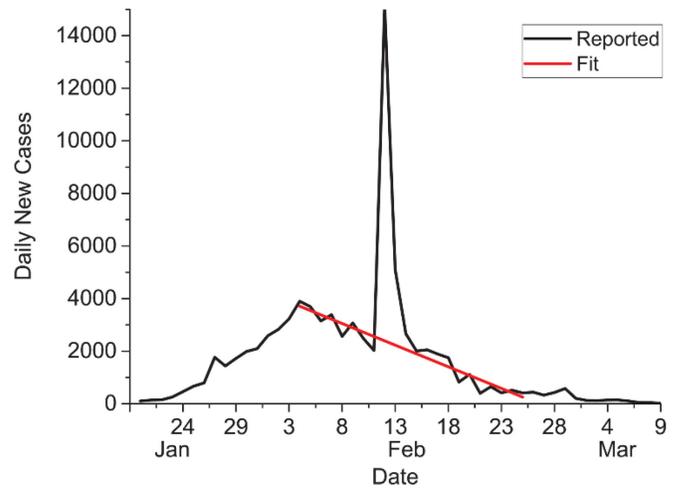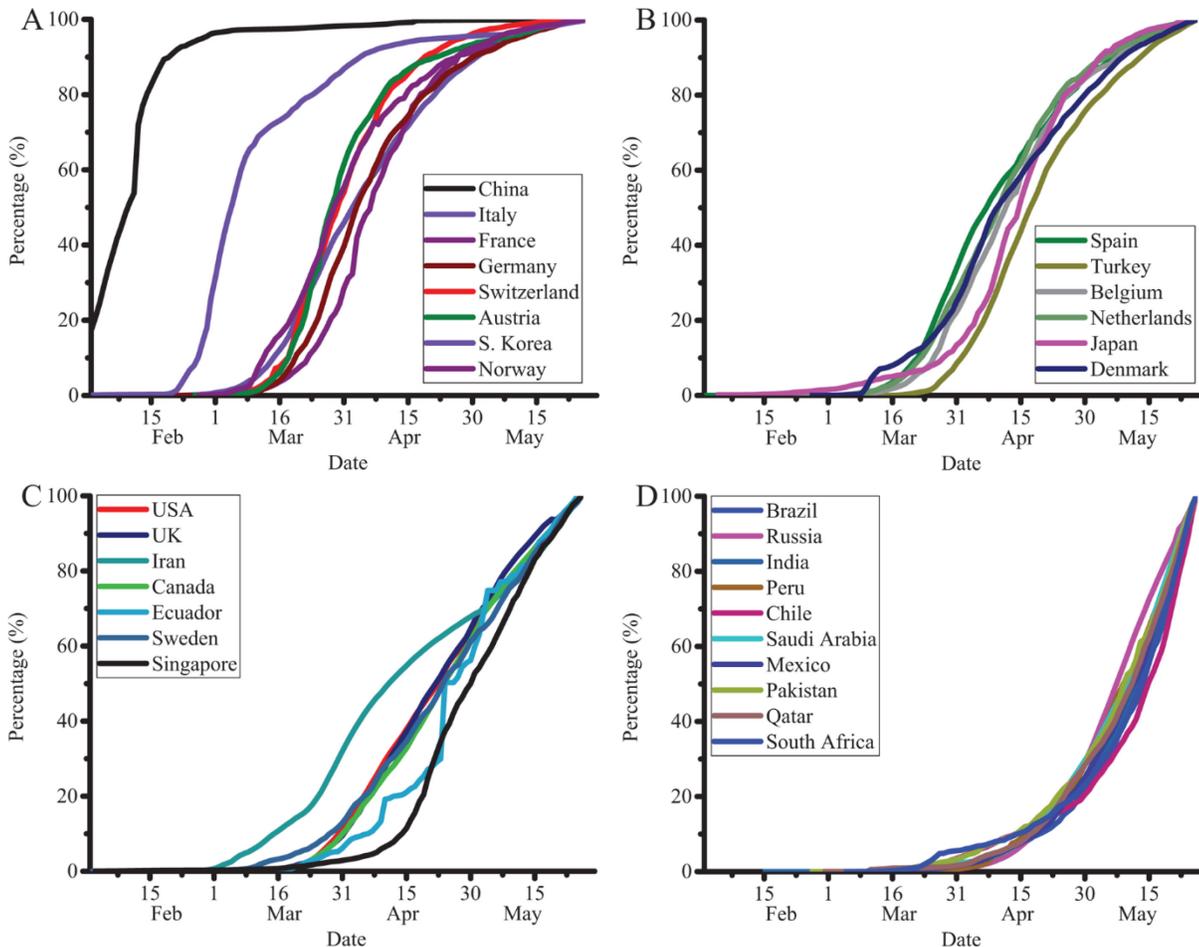**Figure 4.** The trend of pandemic in China.

Figure 3. The progress percentage in selected countries in groups.

Group selected countries based on stages: (A) Group 1, Close to the end, (B) Group 2, Over the peak, (C) Group 3, Steady growth, (D) Group 4, Accelerating.

Due to its simplicity and accuracy, we applied this linear fit model to the data obtained from USA, UK and Canada, and we found that after 15 April 2020, the trends of these countries started to show that the peak has been reached. We proposed that without further outbreak the decline of daily new cases in these countries would oscillate along a straight line and eventurally reaches the bottom like the trend observed months ago in China. The linear trends are $y = 31877 - 271.94x$ ($R^2 = 0.4326$) for USA, $y = 5349.5 - 69.212x$ ($R^2 = 0.571$) for UK, and $y = 1867.2 - 21.241x$ ($R^2 = 0.7287$) for Canada, where x is the number of days after 15 April 2020 (red lines in Figure 5).

Without further outbreak, these trends predict that the epidemic will end on 10 August 2020 in USA, on 1 July 2020 in UK and on 12 July 2020 in Canada, which are 117, 77, and 88 days after 15 April 2020, respectively. The total cases will be 2.50 million for USA, 303 thousand for UK and 110 thousand for Canada, which suggest 780, 37, and 23 thousand more confirmed cases from the total cases on 26 May 2020, respectively. As mentioned earlier, Iran provides a bad example, showing that premature reopening may trigger a second outbreak. Hence, policy makers need to take extra caution when assessing the risk of outbreak again in the country and the timing of reopening business. If reopening business is inevitable, protective measurements such as mask wearing, hand washing, frequent ventilation and disinfection, as well as universal test are essential to prevent any further outbreak.

Summary

This study is the first to use progress percentage to plot epidemic curve and enables direct comparison of data regardless of the large difference in the magnitude of numbers. Side-by-side comparison of global epidemic curves make it possible to group the countries at similar stage and compare the impact of control measurements on the epidemic progression. This study also used a simple but accurate mathematical model to project the epidemic progression in USA, UK and Canada, providing valuable information for assessing the risk of outbreak and the timing of reopening business.

After the original manuscipt was submitted, due to the unexpected incidents and the premature reoponing business, there was a second outbreak starting from mid-June in USA. Detailed discussion of this outbreak is included in the supporting information.(Figure S5) Some other countries developed second outbreak as well but the same model can be applied.

References

1. Li Q, Guan X, Wu P, Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KSM, Lau EHY, Wong JY, Xing X, Xiang N, Wu Y, Li C, Chen Q, Li D, Liu T, Zhao J, Li M, Tu W, Chen C, Jin L, Yang R, Wang Q, Zhou S, Wang R, Liu H, Luo Y, Liu Y, Shao G, Li H, Tao Z, Yang Y, Deng Z, Liu B, Ma Z, Zhang Y, Shi G, Lam TTY, Wu JTK, Gao GF, Cowling BJ, Yang B, Leung GM, Feng Z (2020) Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. N Engl J Med. 382: 1199-1207.
2. Huang R, Liu M, Ding Y (2020) Spatial-temporal distribution of COVID-19 in China and its prediction:A data-driven modeling analysis. J Infect Dev Ctries. 14: 246-253.
3. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, Graham BS, McLellan JS (2020) Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science 367: 1260-1263.
4. Xu Z, Shi L, Wang Y, Zhang J, Huang L, Zhang C, Liu S, Zhao P, Liu H, Zhu L, Tai Y, Bai C, Gao T, Song J, Xia P, Dong J, Zhao J, Wang FS (2020) Pathological findings of COVID-19 associated with acute respiratory distress syndrome. Lancet Respir Med 8: 420-422.

Figure 5. Projections of the disease in (A) USA, (B) UK and (C) Canada.
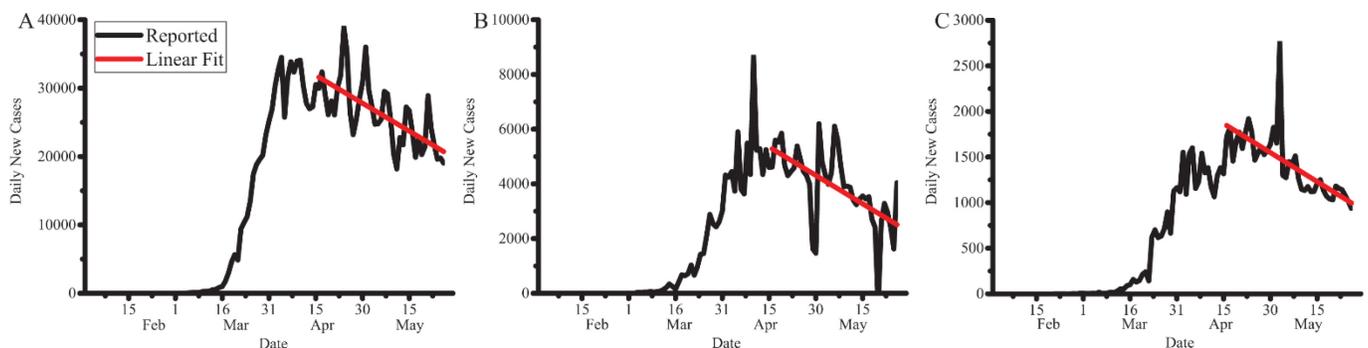
5. Tang Y, Wang S (2020) Mathematic modeling of COVID-19 in the United States. Emerg Microbes Infect 9: 827-829.
6. Contini C, Nuzzo MD, Barp N, Bonazza A, De Giorgio R, Tognon M, Rubino S (2020) The novel zoonotic COVID-19 pandemic: An expected global health concern. J Infect Dev Ctries 14:254-264. doi: 10.3855/jidc.12671.
7. .Porcheddu R, Serra C, Kelvin D, Kelvin N, Rubino S (2020) Similarity in Case Fatality Rates (CFR) of COVID-19/SARS-COV-2 in Italy and China. J Infect Dev Ctries 14:125-128. doi: 10.3855/jidc.12600.
8. Kelvin DJ, Rubino S (2020) Fear of the novel coronavirus. J Infect Dev Ctries 14: 1-2. doi: 10.3855/jidc.12496.
9. Musa TH, Ahmad T, Khan M, Haroon, Wei P (2020) Global outbreak of COVID-19: a new challenge? J Infect Dev Ctries 14: 244-245. doi: 10.3855/jidc.12530.

Corresponding author
Professor Wenjun Tong, PhD
School of Chemistry and Materials Science, Nanjing Normal University, Wenyuan Road 1, 210023 Nanjing, China
Phone: +86 (25) 8589-1319
Email: wjtong@njnu.edu.cn

Jingming Zhou, PhD
Jingling High School Hexi Campus, Mengdu Road 60, 210019, Nanjing, China
Phone: +86 (25) 8647-6007
Email: jingmingzhou@gmail.com

Conflict of interests: No conflict of interests is declared.