**Evidence of the importance of contact tracing in fighting the COVID-19**

Okyu Kwona, \*

a*National Institute for Mathematical Sciences, Daejeon 34047, South Korea*

\*Corresponding author *Email address: okw@nims.re.kr*

**Abstract**

Background: We conducted data analysis to determine whether there are distinguishing characteristics depending on the success or failure of control for COVID-19 by country in the trend of the daily number of confirmed cases and the number of tests.

Methods: We obtained the number of confirmed cases and tests per day for almost every country in the world from Our World in Data. We applied the Pearson correlation between the two time series was calculated according to the time delay to analyze the relationship between the number of tests and the number of cases with a lag.

Results: For each country, we obtained the time lag that makes the maximum correlation between the number of confirmed cases and the number of tests for COVID-19. It can be seen that countries whose time lag lies in a special section between about 15 and 20 days are generally successful in control COVID-19. It looks like a trench on the battlefield.

Conclusion: We have seen the possibility that the success in mitigating COVID-19 can be expressed as a simple indicator of the time lag of the correlation between confirmed cases and tests. This indicator can be presumed to be a sign of an effort to actively tracing infected person.

1. **Introduction**

The novel coronavirus epidemic that started in China at the end of 2019 is still a pandemic as of April 2021. The infectious disease, called COVID-19, has been progressing for a long time and has killed many people, and has caused serious economic damage due to the reduction of social and economic activities. In many parts of the world, interventions to control COVID-19 are being implemented, focusing on quarantine of confirmed cases, social distancing, prohibition of gatherings, and closure of schools. However, controlling infectious diseases is not easy because no one knows exactly who is infected, who is not infected, or who has immunity to the infection. All we know is the infection status of those who have been tested. If the infected person, especially those who are mild or asymptomatic, has not been tested, they will continue to spread the infection to others. Therefore, in order to control infectious diseases, social distancing or prohibition of gathering is inevitably applied to the unspecified majority. This incurs enormous socio-economic costs. For this reason, finding infected people quickly through tests is of paramount importance in the fight against infectious diseases[1, 2]. The Director-General of the World Health Organization also emphasized its importance by saying that “You cannot fight a fire blindfolded. And we cannot stop this pandemic if we don’t know who is infected.”[3] Indeed, it seems that Korea and Singapore have minimized both the human and economic costs of COVID-19 by identifying who has been infected through extensive screening tests in the early stages of the outbreak. Through the test, they were able to avoid the massive socio-economic turmoil experienced in Europe and North America by detecting and quarantining infected people, even mild and asymptomatic cases, and by tracking and examining contacts related to confirmed patients.

Claims and studies have been introduced that emphasize the importance of finding infected individuals through testing and tracking contacts[4, 5, 6]. Most are to show examples for specific countries or to demonstrate their importance through simulations through computational models, We obtained data on the number of confirmed cases and the number of tests in almost all countries around the world and analyzed the correlation between them according to the time delay. As a result, we found circumstantial evidence indirectly showing that testing and tracking of contacts are effective in suppressing the spread of infectious diseases. It was confirmed that the country where the number of confirmed cases peaked and then the number of tests peaked after a certain delay of 2-3 weeks, was showing successful results in controlling COVID-19. The purpose of this study is to show that test and contact tracking are effective in fighting the pandemic, not from simulations using imaginary or computational models, but from various actual data in various countries.

1. **Materials and Methods**

Our world in data is a scientific online publication focusing on large-scale global issues such as poverty, disease, hunger, climate change, war and inequality. In 2020, Our World in Data has become one of the leading organizations publishing global data and related research on the COVID-19 pandemic[7]. It creates and maintains a worldwide database of COVID-19. It is used by the United Nations, the White House, the World Health Organization, epidemiologists and researchers.

From Our World in Data, we obtained the number of cases and tests per day for almost every country in the world. Not only their daily report values, but also smoothed time series are provided. Figure 1 shows a graph of the number of cases and the number of tests for the 30 representative countries selected based on the population. Usually, the number of tests is much higher than the number of confirmed cases. Therefore, the graph was drawn by normalizing the two trends to the largest value in the period we observed to make it easier to see and compare with the eye. Table 1 shows the maximum number of confirmed cases and tests by country. Data on confirmed cases are recorded almost entirely in all countries. However, there are some countries where the test data is completely missing or partially missing. In the case of China and Egypt, there is no test data completely. In Brazil and Vietnam, test data are sparsely missing. It was judged that there was a big difference in the environment of the spread of infectious diseases between before and after vaccination was started, and the date when vaccination was first started is indicated on the graph. The start date differs from country to country, and some countries have not started vaccination during the period we have observed.

1. **Results and discussion**

In fig. 1, it can be seen that the number of confirmed cases per day and the number of tests per day are generally moving in a similar trend. It can be simply understood that the number of confirmed cases increases in proportion to the number of tests. In particular, in the case of the United States, South Africa, Japan, Colombia, Kenya, and Mexico, the trend of the number of tests and the number of confirmed cases seems to be almost the same. Another feature that can be seen with the eye is the cases in which the number of tests peaks after a period of time after the number of confirmed cases peaks. The case of South Korea seems to be a representative example. Another characteristic is that the number of tests continues to be high even after the number of confirmed cases peaks. India and the Philippines are such examples. It can be understood that if the number of tests is maintained even after the number of confirmed cases has peaked, the spread of the infection is suppressed and the number of confirmed cases gradually decreases because hidden infected persons are continuously discovered. Another characteristic is the case of Russia, where the number of tests peaked first, and then the number of confirmed cases peaked at a time difference. It can be understood that if the test is reduced at an early stage, the spread of the infection continues by undetected infected people, and the number of confirmed cases later increases. The number of dates varies from country to country, and only data up to the day before vaccination was used.

**3.1. Correlation with time lag**

In order to systematically analyze the relationship between the number of tests and the number of cases with a lag, the Pearson correlation between the two time series was calculated according to the time delay. We obtain the Pearson correlation as a function of the time delay . And the functional form can be written as

,

,

,

Where and are the number of confirmed cases and tests on the day of country , respectively. is the date length of the data used for country . After vaccination, it was thought that there was a significant change of situation in the spread of infection, so only the period when the vaccine was not inoculated was considered. It can be easily understood as calculating the correlation between the number of confirmed cases on day and the number of tests after days from day. Through this calculation, the time difference , which produces the maximum correlation, was obtained for each country.

Figure 2 shows the results of the correlation analysis, excluding 6 countries that do not have complete test data among the top 30 countries in the population. In general, it shows the shape of a mountain with one peak, and the lag value of at which the peak is located differs from country to country. Most are around , and there are seldom cases where is outside the range of ±10 days. In many countries, the number of confirmed cases and tests are showing a similar trend without a large time lag.

In the case of South Korea, which is considered one of the countries that successfully control COVID-19, the is 16 days. It reflects the trend of the time series in Figure 1, where the number of test peaks 16 days after the number of confirmed cases peaked. On the other hand, in the case of the United States, which is evaluated as unsuccessful in controlling COVID-19, the is 0. This well reflects the graph in Figure 1, where the number of confirmed patients and the number of tests show the same trend without lag.

* 1. **The possibility of success indicator of defeating COVID-19**

Focusing on the results of the above correlation analysis, we examined whether there is a possibility that the character of the time lag between the number of confirmed cases and the number of tests could be an index to evaluate the successfulness of control COVID-19. For all countries where the number of confirmed cases and the number of tests were completed, the time lag value () with the greatest correlation between them was calculated. And we consider the total cumulative number of confirmed cases per 1,000 population as the degree of success for defeating COVID-19 of the country. To see the relationship between the time lag and the success, a scatter plot between and total cases per thousand was drawn as shown in fig. 3 (a). It can be seen that many countries are clustered near where the number of confirmed cases relative to the population is widely distributed from small to many cases. On the other hand, countries whose is outside the range of -10 days and 10 days tend to have relatively few confirmed cases per population. In particular, countries where is between the 15 days and the 20 days show a remarkably low number of confirmed cases compared to the population. South Korea, Taiwan, and Japan are located in that range. To see this trend more clearly, the average of the y-axis values for all the countries included in the window of 11 days length along the x-axis in fig. 3 (a) was calculated. The result can be seen in fig. 3 (b), and it can be seen that a deep valley appears between 15 and 20 days of . The fact that the number of tests peaks 2-3 weeks after the number of confirmed cases peaks can be understood as being actively searching for targets to be tested through contact tracing. As the number of confirmed cases increases, the number of related contacts increases, and tests for them are followed. Therefore, the number of tests does not decrease immediately, but increases and then decreases with a lag.

In fig. 3 (a), the features of the country’s location can be seen divided by continent. In most European and North American countries, the time lag is near zero or has a negative value. And there are many countries with a large number of infected people. It can be assumed that this is because European and North American countries that value personal freedom have difficulty in tracking and inspecting contacts due to difficulties in handling personal information. On the other hand, in many Asian countries, especially South Korea, Taiwan, and Japan, the time lag is a positive value and the number of infections is low. In Korea, after the spread of MERS in 2017, it has become legally possible to handle personal information in an infectious disease outbreak, and accordingly, it is possible to track contacts of confirmed cases using information such as mobile phones and credit cards. Asian countries in such a situation, as mentioned earlier, exhibit the characteristic of generating positive values of as a result of active contact tracing. Most countries in the African show a time lag near zero, as in Europe and North America. However, the number of infected people remains fairly low. Human migration between continents or regions is relatively low compared to countries on other continents, so the spread of infectious diseases may not be significant.

1. **Conclusion**

Through the analysis of the time-lagged correlation between the number of confirmed cases and the number of tests, we found evidence that can infer that the finding of infected cases and the finding of additional infected cases through contact tracing with the confirmed cases is an important factor in the fighting against COVID-19. From the empirical data, it is confirmed that the nations showing the number of tests continuously increased for 2-3 weeks and then peaked after the number of confirmed cases peaked were effectively controlling the spread of infectious diseases. It can be assumed that the reason the test lasts for 2-3 weeks is that the test is actively performed to confirm the spread of the infection by tracing contacts from confirmed cases. These efforts are believed to be of great help in controlling the spread of infectious diseases.

Obviously, some countries have a very low rate of confirmed cases even in countries where is outside the 15-20 day range. Contact tracing is not the only method for COVID-19 pandemic response. Through containment and strong social distancing, contact between people can be reduced. In addition, through thorough hygiene management such as wearing a mask or washing hands, the likelihood of infection can be reduced during contact. There may be countries in which infection is suppressed regardless of the value of by these various mitigation strategies.

The fact that countries with a lag of 15-20 days, or 2-3 weeks, are successful in fighting against COVID-19 may be a special result reflecting the principle of contact tracking and the incubation period characteristics of infectious diseases. The value may contain the secret of an excellent strategy for effective contact tracing and for carrying out inspections accordingly. It will be a meaningful study in the future to find out why this particular time lag occurs through an infectious disease spread simulation model that contains the mechanism of contact tracing.

**Acknowledgment**

This project was supported by the National Institute for Mathematical Sciences founded by the Ministry of Science, ICT & Future Planning (B21910000)

References

1. Morel, Chantal M., Olof Lindahl, and Volkan Ozenci. “Lessons from COVID-19 on the role of the state and the market in providing early testing”. Journal of Global Health 10.2 (2020).
2. Subbaraman, Nidhi (2020-03-23). “Coronavirus tests: researchers chase new diagnostics to fight the pandemic”. Nature. doi:10.1038/d41586-020-00827-6
3. WHO Director-General’s opening remarks at the media briefing on COVID-19 - 16 March 2020
4. Salathe, Marcel, et al. “COVID-19 epidemic in Switzerland: on the importance of testing, contact tracing and isolation.” Swiss medical weekly 150.1112 (2020).
5. Jo, Y. S. A guide to Singapore’s Covid-19 contact-tracing system, Multimedia. The Straits Times https://www.straitstimes.com/multimedia/a-guide-to-singapores-covid-19-contact-tracing-system (2020).
6. Bradshaw, William J., et al. “Bidirectional contact tracing could dramatically improve COVID-19 control.” Nature communications 12.1 (2021): 1-9.
7. <https://ourworldindata.org/coronavirus>

Table 1 The maximum number of confirmed cases per day and the maximum number of daily tests for the period from January 1, 2021 to March 27, 2021 in the top 30 countries by population size.

|  |  |  |  |
| --- | --- | --- | --- |
| Continent | Country | Maximum cases | Maximum tests |
| Asia | China | 4,607 | - |
| Asia | India | 93,199 | 1,196,972 |
| North America | United States | 250,744 | 1,864,725 |
| Asia | Indonesia | 12,865 | 49,183 |
| Asia | Pakistan | 6,533 | 42,212 |
| South America | Brazil | 77,129 | 98,318 |
| Africa | Nigeria | 1,666 | 19,768 |
| Asia | Bangladesh | 3,810 | 23,648 |
| Europe | Russia | 28,501 | 565,313 |
| North America | Mexico | 17,559 | 37,414 |
| Asia | Japan | 6,446 | 74,090 |
| Africa | Ethiopia | 1,879 | 21,880 |
| Asia | Philippines | 8,055 | 47,614 |
| Africa | Egypt | 1,575 | - |
| Asia | Vietnam | 56 | 24,857 |
| Africa | Democratic Republic of Congo | 249 | 869 |
| Asia | Turkey | 33,307 | 203,635 |
| Asia | Iran | 13,630 | 63,195 |
| Europe | Germany | 25,757 | 238,862 |
| Asia | Thailand | 943 | 30,718 |
| Europe | France | 56,225 | 515,517 |
| Europe | United Kingdom | 59,829 | 1,409,016 |
| Europe | Italy | 35,073 | 329,462 |
| Africa | Tanzania | 29 | - |
| Africa | South Africa | 19,042 | 66,585 |
| Asia | Myanmar | 1,505 | 26,928 |
| Africa | Kenya | 1,327 | 8,890 |
| Asia | South Korea | 1,047 | 54,529 |
| South America | Colombia | 17,857 | 73,144 |
| Europe | Spain | 37,011 | 257,295 |

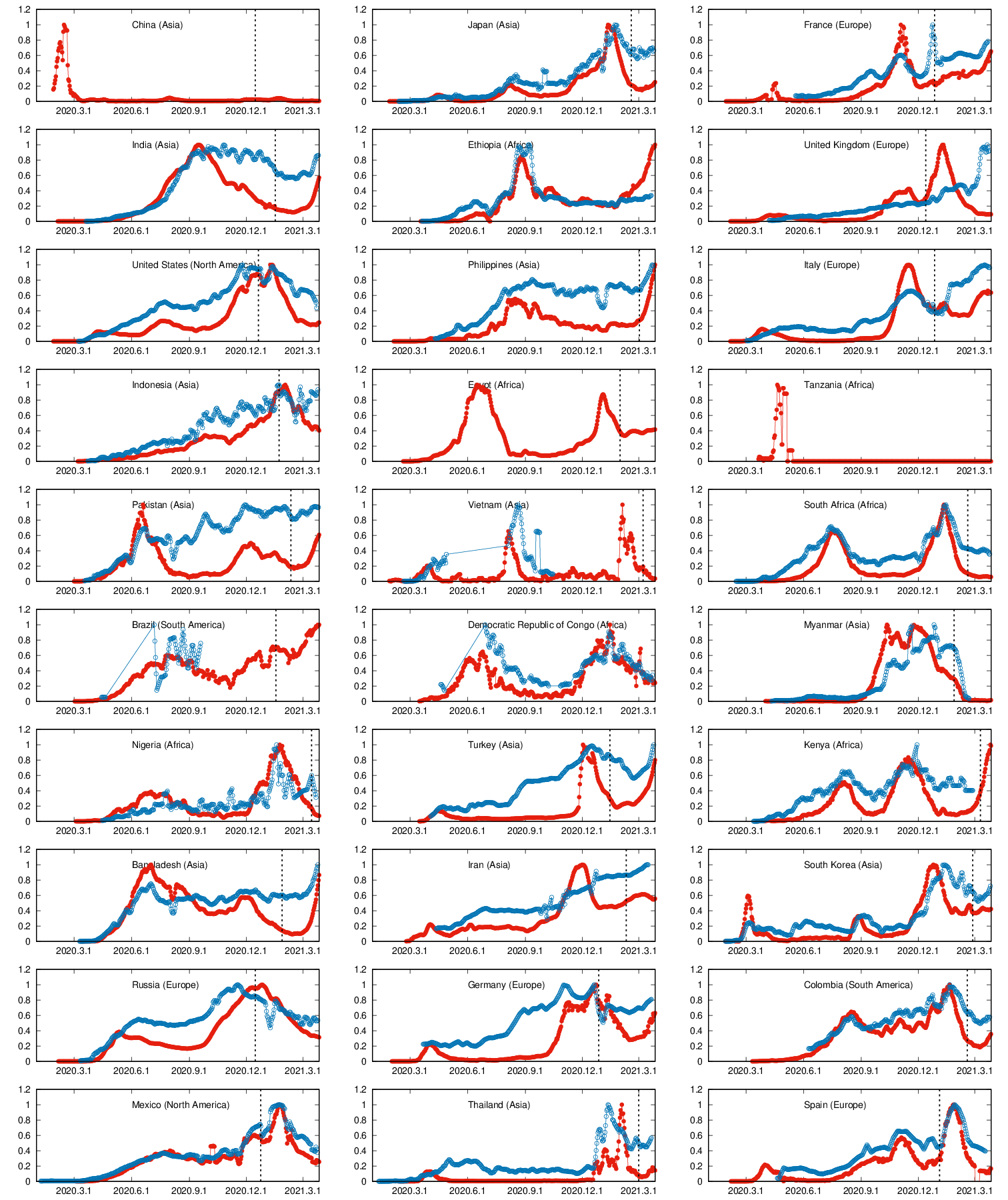


Figure 1 The smoothed daily number of new confirmed cases (red) and test (blue) from January 1, 2021 to March 27, 2021. The vertical dashed line separates the unvaccinated and vaccinated periods. No test data exist for China, Egypt and Tanzania. Brazil, Vietnam, and Democratic Republic of Congo do not have complete data about test.

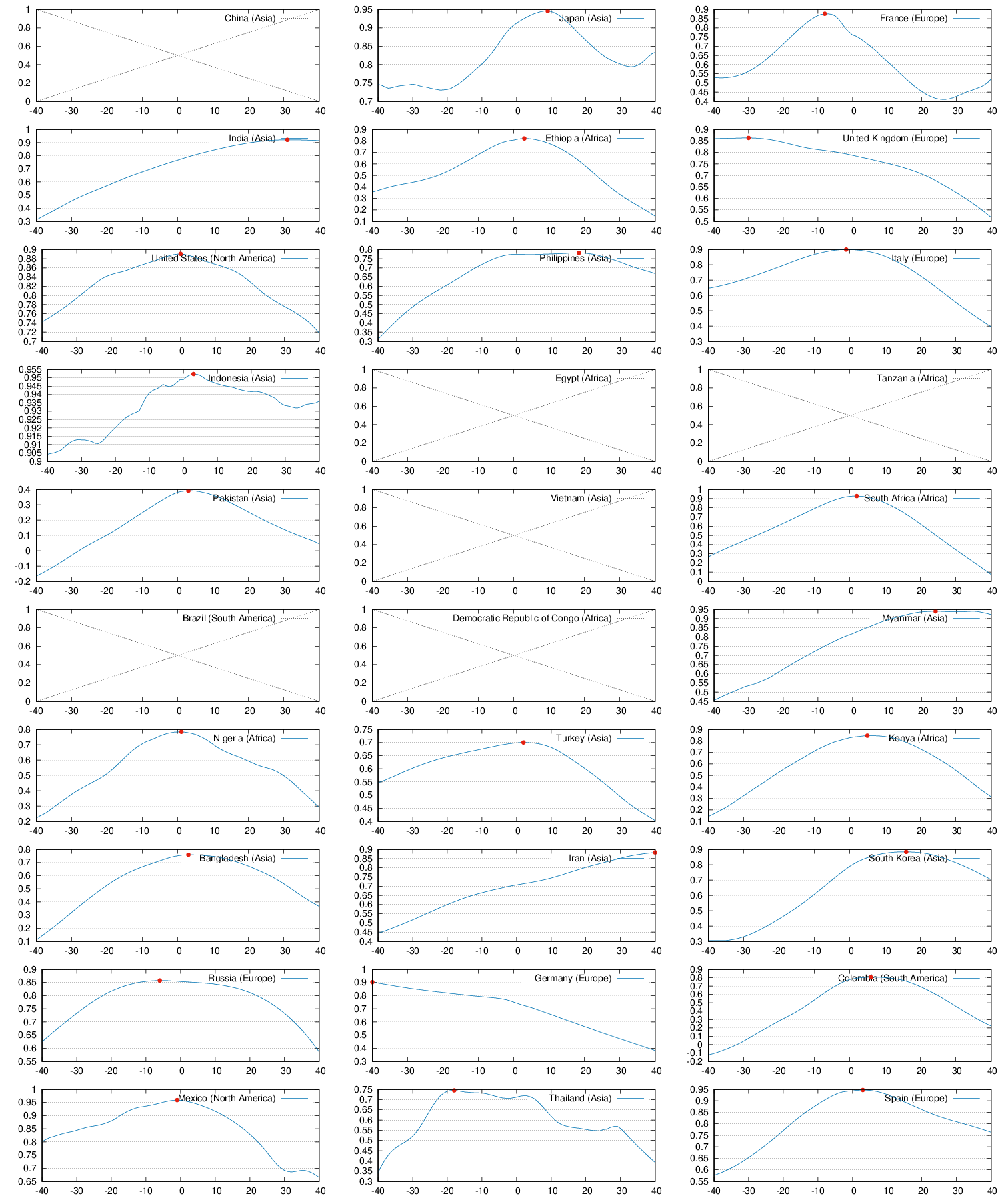


Figure 2 The correlation coefficient according to the time lag between the number of confirmed cases and the number of tests. The x-axis is the time difference and has a unit of day. The y-axis is the correlation coefficient value. The location with the largest correlation coefficient value is indicated by a red dot.

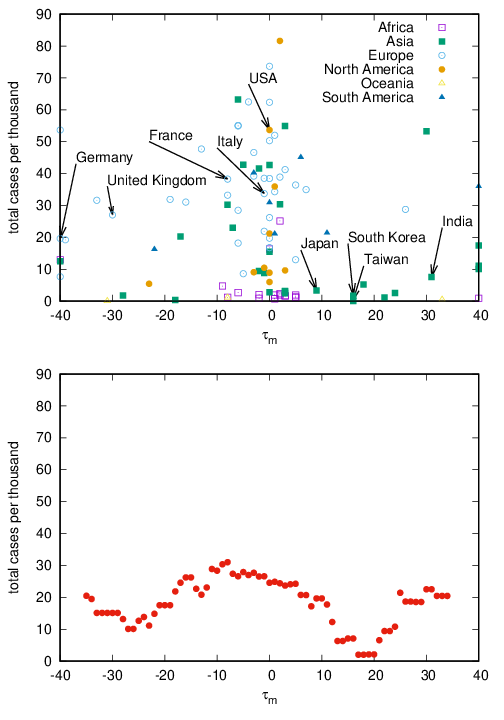


Figure 3 Each point in (a) represents a country, the x-axis is the time lag value at which the correlation of the country is maximized, and the y-axis is the total number of confirmed cases per thousand population. Each country is marked by different colors and shapes for each continent. There are 106 points. Each point is marked by different colors and shapes for each continent to which each country belongs. (b) shows the average of the total number of confirmed cases per thousand population for all countries included in the 11-day window of the x-axis.