Global Health Security:
COVID-19 & Its Impacts

Pandemic Tracking:
The Unfortunate Use of Statistics

By Jikon Lai

SYNOPSIS

There are two problematic issues with the use of statistics in the context of COVID-19: the use of inappropriate indicators and inapt use of indicators. While statistics might offer us an insight into the disease's development over time, we should not focus too much on the actual numbers that are reported, particularly when comparing situations between countries.

COMMENTARY

IT IS now cliché to suggest that COVID-19 has lifted the veil on shortcomings in almost all aspects of our various dimensions of public life. One that has not received a lot of attention however is our use of statistics.

Day after day, one media report after another continues to cite statistics related to COVID-19 in ways that are problematic at best, misleading at worst. Issues with the use of these statistics in recent months fall into two main areas: the use of inappropriate indicators and the inapt use of indicators.

Use of Inappropriate Indicators

Almost every news report and analysis on COVID-19 cite and use the total number of cases. In the context of discussing the situation within a single country, this is understandable. The total number of cases convey information about the severity of

the situation. It is also something that can be used to compare changes from one day to another.

The issue arises however when this indicator – the total number of cases – is used to compare situations in different countries, something which observers do every single day. Many use the total number of cases reported to imply, infer or convey the severity of the situation in one country relative to another.

For instance, it is not uncommon to read statements that are similar to the following: 'with X number of cases, country A is now the worst-hit country in the world'. The unstated assumption or implied observation here is that country A has a higher number of cases than others.

The issue with this logic lies in the use of the raw number of cases to compare different countries. For example, if China had a total of 50,000 cases and Singapore had a total of 40,000 cases, would it be correct to conclude that the situation is worse in China than it is in Singapore…? With those numbers, many would probably argue that the situation is worse in Singapore than in China because China has a population that is almost 250 times larger than Singapore's.

The Right Way to Use Numbers

As this example suggests, when comparing the situation in different countries, the relevant indicator should be the total number of cases normalised against the respective countries' populations rather than the raw total number of cases. At the time of writing, many were looking in horror at the situation in Brazil because with over 330,000 cases, it had one of the top three highest number of cases in the world and was deemed to be the epicentre of the pandemic in Latin America.

However, if we look at the number of cases normalised against the size of population, Brazil recorded approximately 1,565 cases per one million population. In contrast, Ecuador reported approximately 2,034 cases per 1 million population, Chile 3,239, and Peru 3,393.

I am not suggesting that the situation in Brazil is/was not dire but it is problematic to suggest, as many observers and media reports do, that the situation in Brazil is somehow the worst in Latin America simply based on the raw total number of cases. Even on a count of the number of deaths, it is problematic to suggest that the situation was the worst in Brazil among Latin American countries. Although Brazil does have the highest total number of deaths among Latin American countries, at the time of writing both Peru and Ecuador had higher number of deaths than Brazil when those statistics were normalised against the size of their respective populations.

[A related but different issue here is whether it makes analytical sense to compare two very different countries such as, for example, Singapore vs China.]

Inapt Use of Indicators

From numerous media reports and government statements, we know that there are a range of issues with the reported statistics on COVID-19.

There are debates over what is or is not counted as a case of COVID-19 because there are different ways of detecting someone who has contracted the virus. Countries also adopt different decisions on whether to officially count deaths that are suspected, but that have not been confirmed through a test, to have been caused by COVID-19 (see for example this article by [The Economist](#)).

Above and beyond definitional issues, there is also scepticism over the reliability of the measurement of cases. There have been reported shortages of, as well as faulty, test-kits. Policy-makers set different conditions under which testing is conducted, thus making tests widely available in some countries, but not others. Some governments even chose not to test!

The above suggests that the reported statistics reflect different things in different countries. This in itself is not necessarily an issue. First, however a country chooses to define the statistic, and however it is measured, the statistic might still have meaning within its national context.

Second, since the possibility of getting a true and accurate measure of COVID-19 is an open question, deviations from 'the truth' is not necessarily an issue provided that the deviation is not materially substantial and that the 'error' in measurement is consistently applied over time.

This leads us to the following: (1) our focus should be on the trend-line and not the numerical level of the statistics being reported; and (2) numerical comparisons of the statistics reported by different countries is problematic at best. Neither of these issues is unique to COVID-19, but are instead common in the construction of most statistical measures.

Caveat Emptor

Numbers are seductive. People like to use it for the certainty that they seemingly offer. However, caveat emptor: While statistics on COVID-19 might offer us an insight into the disease's development over time, we should not focus too much on the actual numbers that are reported.

There is a good reason that statisticians use the term 'estimates' when reporting statistical measures. I am not suggesting that we do away with statistics altogether but we should exercise considerable care in our use and reliance on them, particularly in our attempts to compare situations between countries.

Jikon Lai is Assistant Professor in the Centre for Multilateralism Studies (CMS), S. Rajaratnam School of International Studies (RSIS), Nanyang Technological University (NTU), Singapore. This is part of an RSIS Series.