

What statistics should we observe for the COVID-19 disease?

We receive a lot of information about the Covid-19 disease diffusion and we can access a lot of data world-wide. However, most of the readily available statistics can be misleading and have two negative consequences:

- increase the sense of insecurity among the population causing panic and irrational behaviours;
- lead the decision makers to emotional driven decisions under the sentiment of frustration.

Before talking about the statistics let me make the point about what I consider now common knowledge about this disease.

- The diffusion of the contagion is extremely rapid, most likely due to fact that the virus is totally unknown for the population.
- The real negative impact of this disease concerns essentially the elders and people already affected by other chronic diseases.
- The combination of the above two is an exponential growth of very severe cases requiring intensive care. For this reason it is likely that the intensive care departments of the national health system risk to saturate rapidly implying a collapse of the whole system.

It is clear from the above that the precise decision problem decision makers are facing is how to avoid the collapse of the health system and more precisely saturation of the intensive care departments. The reason is both ethical (do not abandon the more vulnerable: elders and sick) and practical (a collapse of the health system would have a tremendous impact).

We can now turn our attention to the statistics.

1. The first and more widely used statistic is the daily rate of positive tests usually at national/regional/local level. Although the figure is immediate it is clearly misleading.
 - Different countries use different policies as far as testing the population is concerned: this means that comparing such figures among different countries is meaningless.
 - Unless we test 100% of the population at the same time and at regular time

intervals we really ignore the real shape of the contamination spread and at any given time we ignore which portion of the population is effectively contaminated or not.

- For the same reason any increase of the number of positive tests tells us very little about the contamination speed and acceleration (first and second derivate of the curve). It could easily be that the increase of positive tests is a result of doing more tests: if at day t we have k more positive tests with respect to day $t - 1$ we now we discovered k more contaminated not that in day k we actually had k new contaminated.

- Finally, since we do not know the true extension of the disease we cannot have a real idea of how severe it is. Indeed we know that most of the population is asymptomatic or that develop very mild symptoms, but we have no data showing how many these are.

2. A second, widely used statistic concerns the “deaths”. It is very emotional (obviously) and for this reason even more misleading.

- Once again different countries count differently the number of “deaths” (a typical difference being deaths due to covid-19 vs deaths of patients being positive to covid-19) which makes any comparison among countries meaningless.

- For this reason any ratio between deaths rate and positives rate is even more meaningless because combines statistics which presently cannot be used.

- Finally, it should be noted that the statistics about deaths observes the end (unfortunate) of each individual process. Unless we can declare the end of the pandemic and count and compare the different types of processes of the different types of patients any statistic is meaningless.

What should we observe instead and why? My claim is that the only useful and meaningful statistic in this moment is the daily occupation at the “Intensive Care Departments”.

- This statistic is complete: we know everyday exactly the 100% of the patients, positive tested, who join the ICDs.
- It is unambiguous worldwide, the protocol for declaring a patient necessary to recover in ICD being practically the same worldwide.
- It remains a safe proxy of how the contagion is progressing. Although we do not know the real curve and the law between the contagion and ending

at the ICD, we know there is a positive correlation between the two. If the occupation at the ICDs increases or decreases and if this happens rapidly or slowly this tells us that the contagion itself is behaving at the same way.

- It gives a precise idea about the capacity of our health system to resist to the crisis and to respond to the “real problem” mentioned at the beginning.

Aiding to make appropriate decisions implies being able to identify key information which can lead us to undertake reasonably good action. Choosing the appropriate statistic to observe, while within such a major crisis is a critical decision to make.

Alexis Tsoukiàs, LAMSADE
Paris, 23/03/2020