

Covidalculation and Covidmunication

Some methodological insights

Jacques Martin (ESOE)

Claudio Bacarani (University of Verona)

Federico Brunetti (University of Verona)

Almost everywhere in the world people have been bombarded with figures about the coronavirus pandemic since March 2020. We tried to point to a number of instances of erratic, unclear and contradictory information in a short paper last April (Martin J., Bacarani C., April 2020). The movement has amplified since then and is reaching a new apex with the so-called second wave of the pandemic, particularly in Europe, which is the most affected part of the world, bar the USA and Brazil.

The figures communicated by the media and governments, for the latter to give a 'scientific' ground to the political decisions made to combat the virus and its spread, raise a number of questions. The problem is then twofold. One, what is the validity of the figures used as a basis to make decisions? As we shall see, some methodological flaws and the big error margin of the results obtained cast doubts about this validity, hence about the decisions made. Two, is the way the figures are communicated appropriate? Here too, as shown below, in so far as in several ways the communication has got effects which are the opposite of the ones intended, it can be said that the communication is largely inappropriate.

First of all, these figures are never properly explained to the public, certainly not understood by the media people who repeat them like parrots, and most probably not grasped by a majority of government members, not to mention 'ordinary' politicians. The result is that the general public is utterly confused and has little or no trust in the word of governing authorities. If the first periods of confinement were rather easily accepted in the Spring of 2020, during the present second wave, there is significant resistance to the measures taken by European governments in such varied countries as Spain, France, England, Germany or Italy. In some cases, this opposition can be expressed in a violent way, as we have witnessed in Italy, especially in the South of the country, and sporadically in France. Compared to the situation in the Spring of 2020, the present situation is harder to bear for people not only because the communication of this avalanche of figures is confusing, not understood and not understandable, but also because the economic and psychological state of populations is quite different. In the Spring of 2020, individuals and businesses were in a position to be more resilient, which was largely supported by governments pouring money into economies to try and smoothen the shock caused by the pandemic. Remember President Macron of France's phrase eight months ago: "whatever it may cost". Today, populations are much less resilient. They are weary, disillusioned, depressive in many cases, in dire economic situations for many of them even if most governments, and the European Union, still follow the 'whatever it may cost' policy, and have lost almost all trust in their government. They have a strong feeling of having been cheated being back to square one.

We would like to present some tools which are regularly used to give the public a picture of the pandemic and take France and Italy as two instances, but the situation is pretty similar in other European countries, to illustrate how these tools are used and how the information they provide is communicated.

Starting from a global outlook, we must be aware that the organization of the health system is quite different in France and Italy. In France the system is centralized and monitored by the French Ministry of Health, although we can get data at the regional or a lower level. At first sight, this should be a guarantee for the consistency of the data and the communication. Unfortunately it is not actually the case. In Italy, the responsibility for health issues lies with the regions. This implies that we can have variations in the data provided and the way they are processed and of course in the opportune and effective character of the decisions made and the solutions implemented.

First, what we can notice when we compare the social environment during the 'first wave' in the Spring of 2020 and the 'second wave' in the Autumn of 2020, is that, in both countries, the attitude of the people has significantly changed due to the length of the pandemic. A sort of denial of the 'official' communication and a distrust toward the management of the crisis has spread. The health personnel who were seen as heroes during the first wave is now largely ignored and violent actions can take place. For example ambulances have been attacked in Italy, and demonstrations have taken place both in France and Italy against governmental measures.

Second, we can wonder about the validity of the scientific tools, not in themselves (they are long-established ones in epidemiology and have proved their relevance but in **controlled conditions**), but in the way they are used by governments to communicate with the general public. There are, what we can call, methodological flaws in the use of the tools and the communication of the results so that people have got a sort of intuitive understanding that they are deceived. Moreover, we see almost every day panels of experts spending their time contradicting one another, as if they were more interested in asserting their own 'superior' knowledge than informing the public. In Italy for example a researcher of the National Council of Research disavowed the President of the High Council of Health about the exponentiality of the contagion curve – rightly as will be seen below.

We would like to examine the most common tools used and the way they are communicated.

We will have a look at four tools used day after day to show the people the state of the pandemic, and, on the basis of the results obtained, to justify the political decisions made. So, we will consider the 'rate of positivity, the 'rate of incidence', the 'R0' and the function of exponentiality.

Number of 'new cases' and positivity rate:

Ministers and journalists constantly use the word 'indicator', but they do not seem to understand the proper meaning of the word. We hear phrases such as "the indicators are bad, the indicators are not good, the indicators are worsening, the indicators are improving". What is an indicator? "something that shows what a situation is like" (*Cambridge Dictionary*). So, an indicator is a tool, not a figure or result. There is a confusion between 'what shows' and 'what is shown' (between the thermometer and the temperature). Therefore "if the indicators are bad, there must be changed!" (Candau P., 2005). Then we have the 'number of positive cases', labelled 'new cases', which is a raw figure, which is meaningless as it must be related to the number of tests performed, which is not systematically given. Consequently, people tend to think that it is an absolute number. This figure is also quite meaningless because it is not representative in any way of the population. Who gets tested? People who have got Covid-19 symptoms; so, automatically the probability of being positive is higher. People who have been 'in contact' with positive people, as far as they can know (!), which also increases the probability of positiveness. People who are anxious; so, the key parameter is not the disease, but becomes the level of anxiety. In France, people can go for a test after seeing their GP but they can also decide by themselves to be tested, which introduces a big bias. In Italy GPs carry out an evaluation of the health condition of the patient first. But even upon medical recommendation, some people may shun the test because they are afraid to go to hospital or to be handicapped in their jobs. Thereupon,

a positivity rate, which is $(\text{nb of positive cases}/\text{nb of tests}) \times 100$, is calculated and thrown at the faces, or ears rather, of people without being explained. Such a figure has in fact no statistical relevance nor significance. Some other figures are also given in a sporadic way: the number of deaths (normally reliable, but not quite), the number of people going into intensive care (Italy) or in intensive care (France and even in Italy), which is not the same, the number of people cured (Italy, never in France), the level of saturation of hospital structures, with sometimes strange ways of calculating it; we have heard for example in France that the rate of occupation of intensive care beds was 140%! Logically this means that there is more than one person per bed.

Rate of incidence:

We can turn now to the rate of incidence. First of all, we are sure that the phrase itself does not mean anything to anyone who is not an epidemiologist. If we look, for example, at the definition given by the French Ministry of Health, this is what we find: "The rate of incidence corresponds to the number of positive tests for 100,000 inhabitants. It is calculated in this way, according to the formula given by the French Ministry of Health $(100,000 \times \text{number of positive cases})/\text{population}$." We must confess that this is not quite limpid. Fortunately, we can find something more explicit, such as: "The ratio between the number of new cases of disease during a specified time interval and the average population during the time interval" (cdc.gov).

How is this 'rate' given by the media? In this way: 140 for 100,000 inhabitants. Like for the positivity rate, it is not stated that the first figure is the number of new cases observed, the time interval is absent (although the French Ministry of Health, for example, calculates it on a daily and weekly basis; plus the fact that this interval can be fixed or moving), and the word 'average' is also absent though it has a negligible impact on calculation on a daily or weekly basis. In Italy this indicator is not used in the case of the Covid pandemic, although it is used for other diseases. And of course differences between regions would be large and probably amplified by the decentralization of the management of health. It is sometimes used by experts and also expressed in $x/100,000$, which may be meaningful for experts but not for the general population. So, it could be said that it is wiser not to communicate with this 'rate'.

The first remark we can make is that this presentation is not a rate. A rate is normally expressed as a percentage (%) - except in some specific cases -, which, as the word indicates, is out of a hundred, something people are used to, whereas we do not think they can 'visualize' a result out of 100,000, even if it were expressed as $\frac{\quad}{100,000}$. So, let us use our example and express it as a (percentage) rate. What we get is $(140/100,000) \times 100 = 0,14\%$. We can anticipate that the perception of the population would be different if it was expressed in this way. The risk might of course be that they would believe that there is no problem at all.

The second remark we can make is derived from what is said above about the positivity rate. As the number of new cases is not reliable from a statistical point of view and hence does not reflect the real situation, this rate of incidence has got a big margin of error, giving it little meaning.

R0:

What is the R0? We are pretty sure that very few (no?) journalist or minister would be able to give a full explanation of it. Let us have look at a definition. «The R0 (initial reproduction rate of the virus) is calculated from a population which is entirely susceptible of being infected (i.e. not vaccinated nor immunized). It corresponds to the product of three factors: $R0 = \beta c D$. (French Ministry of Health)

- β = the risk of contracting the virus during a contact
- c = the number of contacts during a time interval
- D = the number of days during which a person is contagious (up to 14 days in the case of the coronavirus)

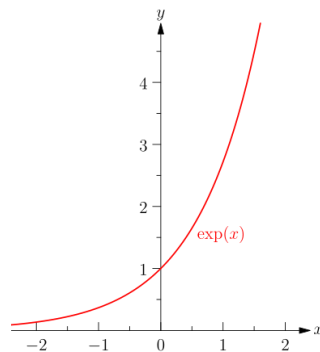
We can make the following remarks. β is a parameter for which there is sufficient scientific evidence to have an acceptable value. c , this is where there is the rub. If this parameter is fine from a theoretical point of view. In practice it is impossible to have a reliable figure for it in the social conditions in which the virus can spread. So, we can but have an approximation, which can be the cause of a big bias and render the result non-significant. D is also a parameter for which there is sufficient scientific evidence to have a reliable value. Therefore, this indicator, which is fine in theory and operational in controlled conditions, can only give a broad estimate with a big margin of error. Moreover there are three ways of calculating the R_0 : the first is to use the number of confirmed cases, the second is to use the number of people admitted in a hospital emergency ward and the third is to use the number of people hospitalized for Covid-19. The three ways of course give different results.

From the point of view of the communication, on the one hand the time interval is rarely, if ever, given, and on the other hand the way (1, 2 or 3 above) it is calculated is never given.

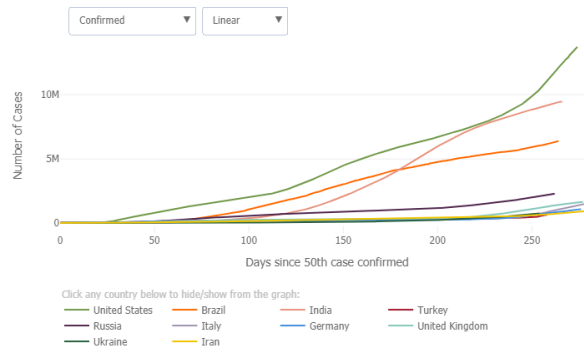
It can be that giving that sort of information can increase the confusion among the public instead of making the situation clearer. On the one hand because the public is not told how this indicator is calculated, and on the other hand because, even if it were, the majority of people would not understand it. So, it adds to anxiety rather anything else.

Exponential function:

We hear again and again that the spread of the virus grows exponentially. Strangely enough, when figures go down, we hardly hear anything. The problem is that the media and politicians use this expression without knowing what an exponential function is. An exponential function is expressed in this way: $\exp(x) = e^x$, a function which is its own derivative. Its graphic representation is like this:



Let's compare, for example, the exponential function with the number of cases in different countries (*Johns Hopkins University*)



Whatever way we look at the data about the pandemic, we can never see an exponential function in its real mathematical sense. Even if we may witness an exponentiality over a short period, not more than three or four weeks, the curve rapidly flattens and tends to become of a 'Gaussian type', or in the case of successive waves, a succession of Gaussian type curves but with different shapes of the 'bell'. This is a good sign in fact as it shows that the measures taken to control the pandemic are effective. 'Communicators' just want to say that sometimes, and only sometimes, values increase faster. It would be so much simpler to say so, and not to impress the people and deceive them. It would be wiser to speak of 'acceleration', and then 'slow down' or 'decrease' than using mathematical tools that are incorrectly applied and not understood by media people, most of politicians and of course the general public. The result is that such communication creates fear and confusion and open the way to all sorts of hypotheses of a political and behavioral nature that nourish various theories of conspiracy.

These four examples show that politicians, relayed by the media, use data and tools to 'represent' the coronavirus pandemic, which are not inappropriate in themselves, but which are biased for different reasons, mainly methodological ones, inducing, implicitly, certain behaviors only creating mistrust and anxiety with an effect which is often the opposite of what is intended. And worse, could we say, they justify their decisions on the basis of such communication. As 'popular wisdom' says: "By dint of taking people for imbeciles, you prove that you are an imbecile yourself". Figures about such a dramatic situation with huge psychological and sociological influence should be consistent, meaningful and always based on the same criteria. In short, simplicity and clarity in communication are essential.

"What is well conceived, is expressed clearly" (N. Boileau, *L'Art poétique*, 1674)

References:

- Candau P. (2005): Qu'est-ce qu'un processus ?, IAE Aix en Provence, France
Centers for Disease Control and Prevention (2020): *Coronavirus Disease 2019*, cdc.gov
Johns Hopkins University (2020) : *Coronavirus resource center*, coronavirus.jhu.edu
Martin J., Baccarani C. (2020): Some thoughts on Covid-19 communication transparency, EISIC.net
Ministère des Solidarités et de la Santé (2020): *Covid-19*, santepubliquefrance.fr