Press release – 14 May 2020

Early calculation of actual COVID-19 fatality rate in France made possible by mechanistic-statistical model

By the end of March, an INRAE team had already produced the first infection fatality ratio calculation for COVID-19 outside China. They identified a fatality rate of 0.5% based on French hospital data, adjusted to 0.8% when nursing home data were included. Their results were confirmed at the end of April by work carried out at the Pasteur Institute in Paris and by a New York study. The team's analysis was published on 8 May 2020 in *MDPI Biology*.

Measuring the progression of epidemics is always a challenge because they are driven by 'latent' processes that are not immediately observable by scientists or medical practitioners, one such process being, for example, the mechanism by which a virus is transmitted. Meanwhile, those managing the health emergency urgently need an estimate of the infection fatality ratio of a virus and the probable number of those infected. At the outset of an epidemic, the observed case fatality rate (CFR), derived from the number of deaths attributable to the illness divided by the number of known cases, provides little help in evaluating the risks associated with the epidemic as this value is strongly influenced by rates of detection. For example, the case fatality rate observed in France on 10 May stood at 19% but this did not accurately reflect the facts of the epidemic on that date. This would require the calculation of the actual fatality rate, or infection fatality ratio (IFR), derived from the number of deaths attributable to COVID-19 divided by the actual number of cases, which would need to be estimated. In March, data on the infection fatality ratio for COVID-19 were available from just two populations – individuals who had been repatriated from Wuhan and the Diamond Princess cruise passengers. These two populations each had very specific characteristics and were not representative samples of the global population.

In the field of epidemiology, two methods of calculating the fatality rate are available to researchers, but each has its disadvantages:

- The statistical, data-driven approach, which is based on data collected and analyzed by scientists (number of people tested positive, number of deaths etc.) but which takes no account of the particular mechanisms of the epidemic (how the virus is spread, contacts between people etc.).
- The mechanistic approach, which focuses on these mechanisms, but whose outputs are hard to compare to the field data because the latter are strongly influenced by collection methods (sampling bias, observation errors, etc.). Given that the actual process involved is reflected only indirectly in such data, this process is described as latent (or hidden).

Scientists from INRAE's NUMM and SPE divisions have for many years been using mechanistic-statistical modelling methods that couple these two approaches in order to track plant epidemics and biological invasions. The principles they have developed for this purpose are equally applicable to epidemics in animals and humans. What is more, such

methods are particularly useful where the observation data (in this case, the number of positive tests) do not directly represent the latent process itself but are instead indirectly derived from it.

Developing the model

Mechanistic-statistical methods are particularly suited to the COVID-19 epidemic because, while we have access to a large amount of field data, these data are impacted by multiple biases. Here, the teams described the latent progress of the epidemic by using a standard epidemiological model that took the particular characteristics of this virus into account and calculated certain key parameters (including the famous R_0). A probabilistic model was then used to link observed data (number of known cases, number of tests carried out, number of deaths) with the epidemiological model. This probabilistic observation model included a further estimated parameter, namely, the relative probability of a 'healthy' person being tested vs a person infected by SARS-CoV-2 (COVID-19).

Using these methods, our scientists were able, as early as March 2020, to calculate the infection fatality ratio for the COVID-19 epidemic in France, the first such ratio to be computed outside China. They calculated a fatality rate of 0.5% based on hospital data, adjusted to 0.8% when data from nursing homes were included. They also calculated a basic reproduction number (R_0) of 3.2. These results were confirmed at the end of April by the work of the Pasteur Institute, whose own calculations also arrived at an IFR of 0.5% based on French hospital data, and by antibody tests carried out in New York which produced a ratio of 0.6%. These mechanistic-statistical methods are thus immensely valuable to epidemiology, allowing the calculation of crucial parameters such as the IFR at a very early stage in an epidemic.

References

Roques, L.; Klein, E.K.; Papaïx, J.; Sar, A.; Soubeyrand, S. *Using Early Data to Estimate the Actual Infection Fatality Ratio from COVID-19 in France*. MDPI Biology May 2020, 9, 97. https://doi.org/10.3390/biology9050097

BioSP research team blog
https://informatique-mia.inra.fr/biosp/COVID-19

Scientific contacts:
Lionel Roques – lionel.roques@inrae.fr
Samuel Soubeyrand – samuel.soubeyrand@inrae.fr
Biostatistics and Spatial Processes (BioSP) Research Unit
NUMM and SPE Departments
Provence-Alpes-Côte d'Azur INRAE Centre

Press contact:
INRAE Press Office: +33 1 42 75 91 86 – presse@inrae.fr