May 12, 2020: Results of a model for the analysis of the dynamics of COVID-19

Francisco Castro Rego

The model explained in the Annex is based on the dynamics of symptomatic confirmed cases and it was applied to various countries with data provided in May 12 from: https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide. These data were used for the construction of the graphs from figures 2 to 5.

However, these data report all confirmed cases making no distinction between symptomatic and non-symptomatic cases. The total number of confirmed cases roughly corresponds, in a first phase, to the number of symptomatic confirmed cases but later it is highly dependent on the effort made in testing. Therefore, the results must be interpreted with caution. As testing efforts increase with time, the projections might be overestimated.

This analysis overcomes the problems raised with the use of data for all confirmed cases, as they are including greater proportions of non-symptomatic cases that tested positive. These results illustrate the potential use of this simple method that has the advantage of relying very much on the data, with easy daily updates, and with few assumptions that contribute to its robustness. The good fit to the data is also a very good indication of the potential use of the model to evaluate possible scenarios and compare different situations and countries.

For Portugal, we show the results in figure 1.

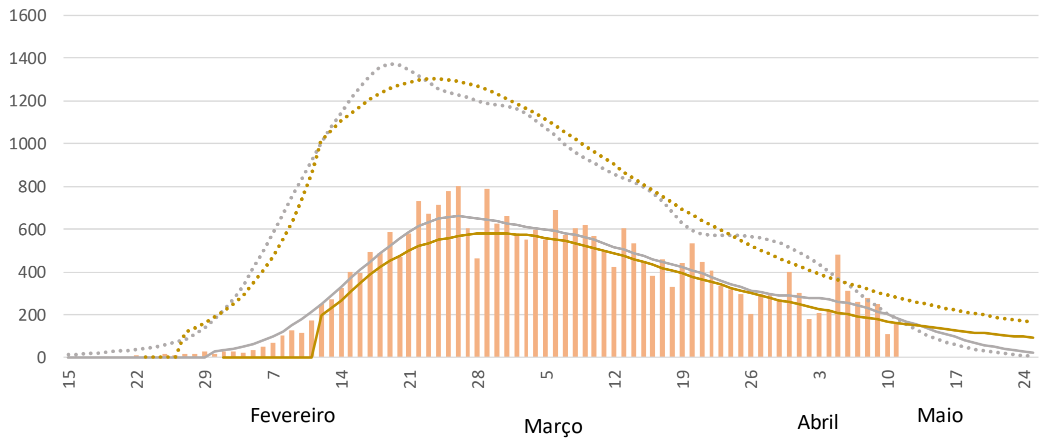


Figure 1. Analysis of the estimated dynamics of transmissions and symptomatic cases estimated by two different models for Portugal from February 15 to May 11, and the projections for the period May 12 to May 25, modelled from DGS data. The assumptions of the incubation period and the asymptomatic rates are in the model.

The analysis of that figure allows for the understanding that the number of transmissions decreased since late March. The two models now converge in projecting a gradual decrease during May.

Results for China, South Korea, Iran and Turkey show that for the first two countries the estimated values are close to the observed values, as expected, since this is a data driven method and the episodes are practically controlled, but new episodes, starting with a very few cases, might still occur. For Iran the numbers are increasing again (Figure 2).

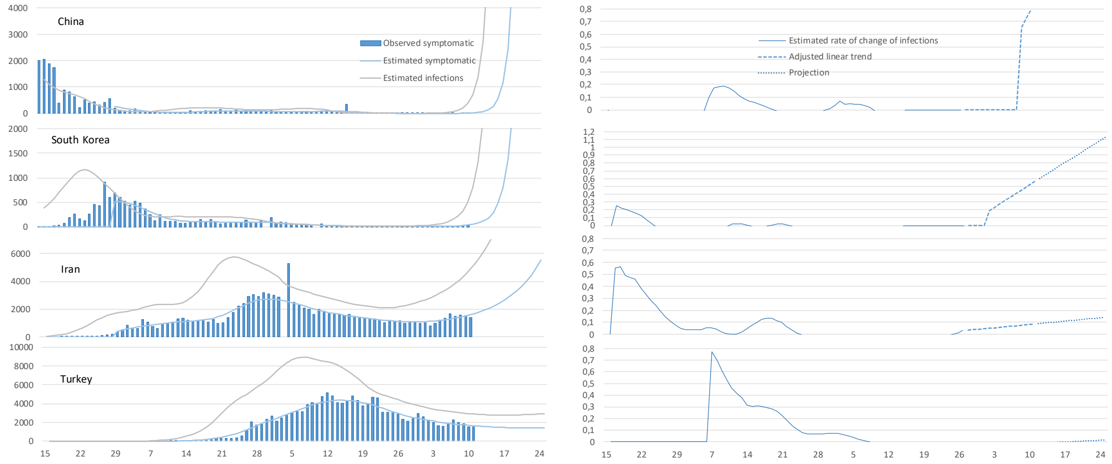
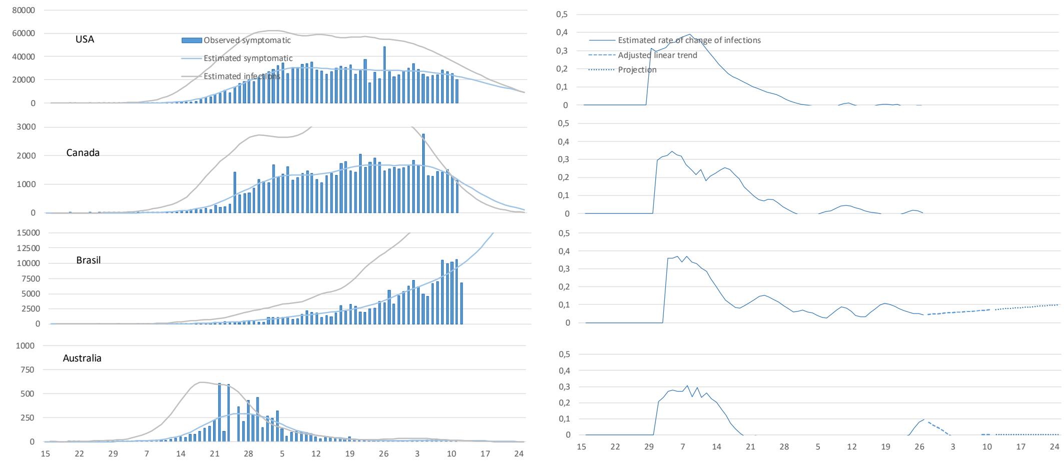


Figure 2. Observed and estimated symptomatic cases and number of infections with the corresponding rates of changes of transmissions in China, South Korea, Iran and Turkey.

The same analysis in other continents with more recent episodes suggest delayed responses, but the episode seemed already finished in Australia. The numbers are still not decreasing very steadily in the USA and Canada and increasing in Brasil (Figure 3).



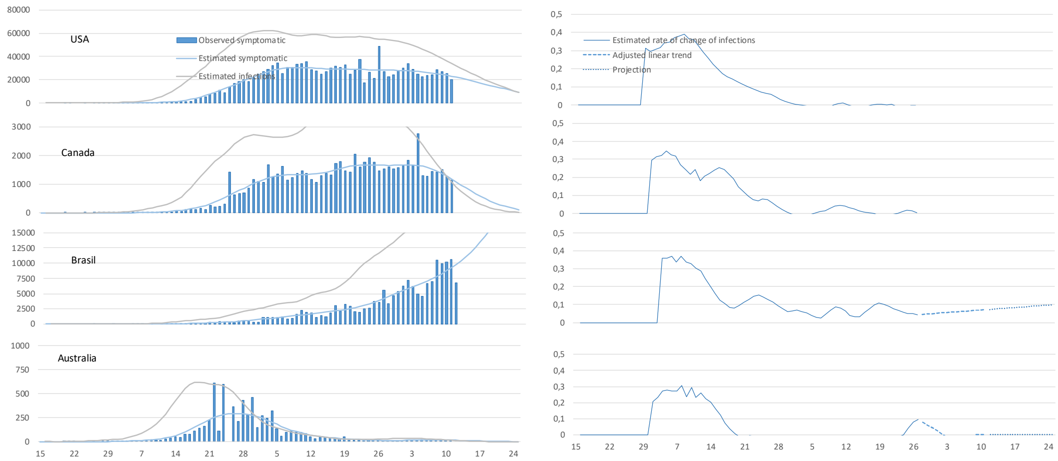


Figure 3. Observed and estimated symptomatic cases and number of infections in the USA, Canada, Brasil and Australia.

In Western Europe, the indication of the model is that the number of infections and symptomatic cases was decreasing in most countries but at a variable speed but in the last period new increases are possible, as in Belgium. In Sweden the numbers show a tendency to increase (Figure 4).

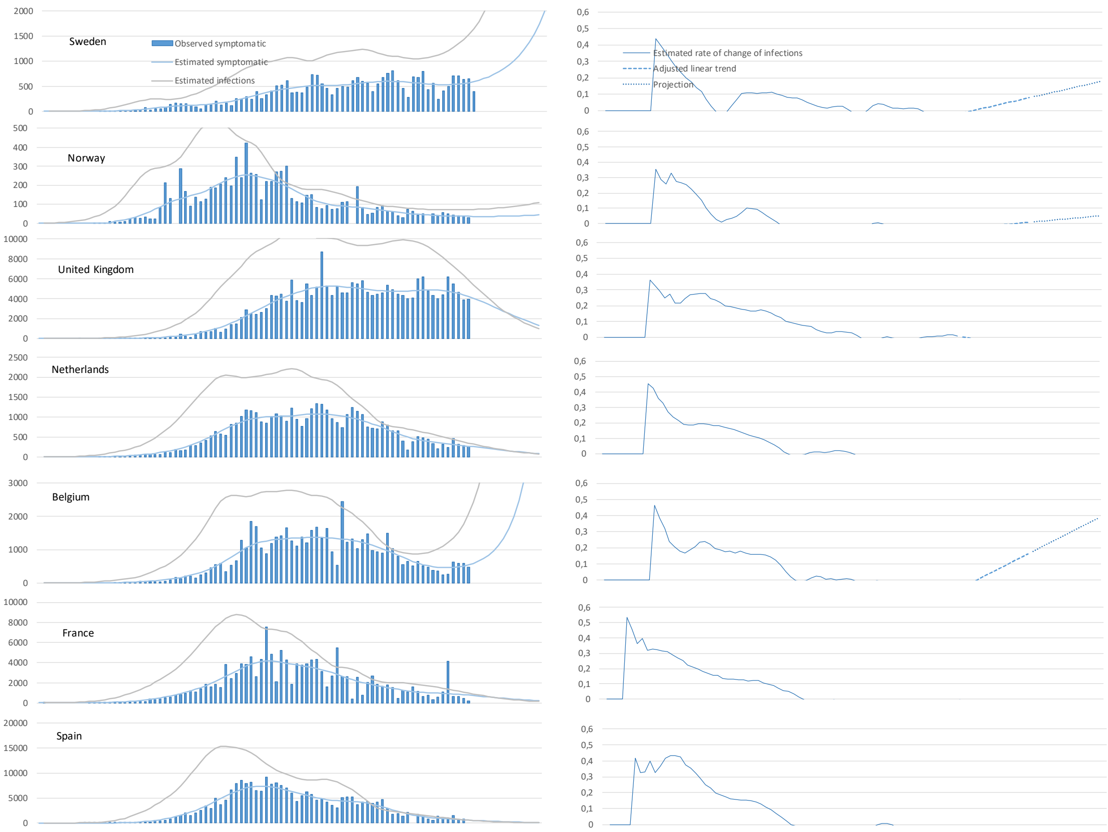


Figure 4. Observed and estimated symptomatic cases and number of infections in Western Europe.

Finally, the analysis of the situations in Italy, Switzerland, Germany and Austria (Figure 8) using this model seems to indicate that the number of transmissions and the number of symptomatic cases already decreased significantly in all cases. However new episodes may happen (Figure 5).

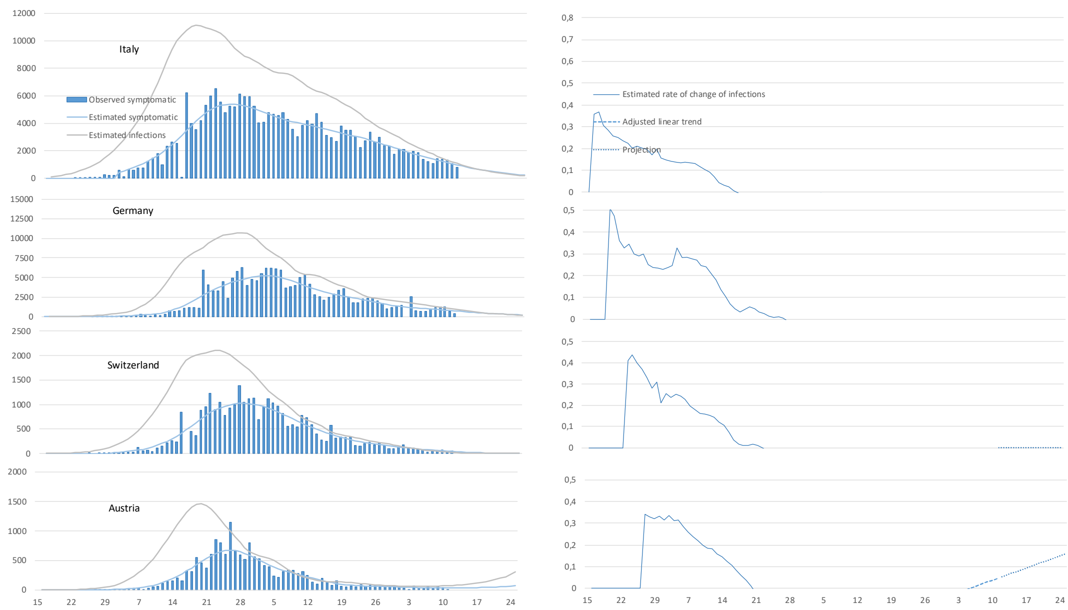


Figure 5. Observed and estimated symptomatic cases and number of infections in Italy, Switzerland, Germany and Austria, showing that the number of infections is already decreasing significantly in all countries. However, new episodes are possible.

In summary, the model of analysis seems to describe well the observations, to be easily interpretable and to allow for some projections with the assumption that the rate of change of the transmissions adjusted for the last 14 days continues constant in the following 14 days. Obviously, this assumption is never completely true but examples from episodes in China and South Korea indicate that this can be a reasonable approximation.

The fact that this system of analysis is data driven is simultaneously a strength and a weakness. The quality of the data is also an issue, with symptomatic cases depending on number of tests done and various ways of reporting. Also, the results are very much influenced by the phase of the episode. Very recent infections have obviously the tendency to indicate exponential growth without any detectable trend to decreased acceleration. However, this model is not depending on many assumptions and provides a very objective way to assess current trends and consequent projections.

Finally, the results of this model of analysis are simple to obtain and projections are estimated daily with updated data.

Annex: The models of analysis

The analysis of the dynamics of COVID-19 infections in the various countries has to use the information from the countries where it first spread, China and South Korea. These analyses should also take into account the information from Japan from the cruise ship Diamond Princess.

First, an estimate of the proportion of the infected people that show symptoms is important. As in an experiment, Mizumoto et al. (2020) indicated that of the 634 confirmed cases in the Diamond Princess only 306 were symptomatic. The asymptomatic percentage of coronavirus disease was a little more than 50%. Chen et al. (2020) also used this value in their mathematical model, suggesting that this may be a good reference value for other situations.

Secondly, it is necessary to have information on the time between infection and the detection of symptoms, the incubation period. The work by Baker et al. (2020) on the incubation period estimated from infections among travelers from Wuhan, China, provides such information. These authors found that the incubation period could be described by a Weibull distribution with a parameter of scale of 6.4 (95% confidence interval 5.6-7.7) and a parameter of shape 2.3 (95% confidence interval 1.7-3.7). Figure 1 shows a representation of the probabilities of infection in the days before the detection of symptoms based on the Weibull distribution.

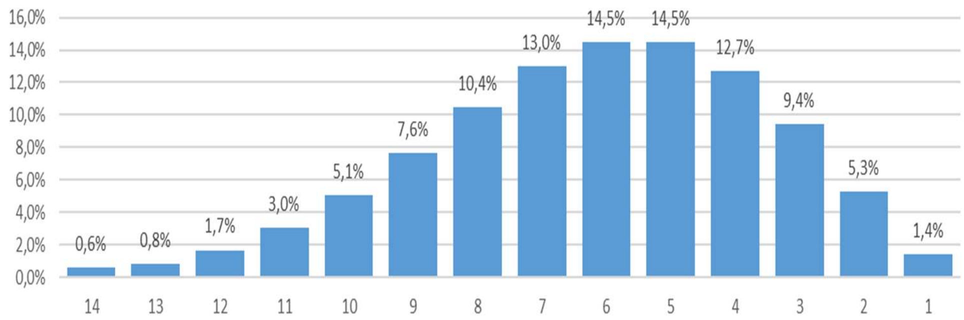


Figure 1. Percentage of symptomatic cases (y axis) distributed by the number of days of the estimated infection prior to detection of symptoms (x axis), following the Weibull distribution of scale and shape parameters 6.4 and 2.3 as in Baker et al. (2020).

More than half of the cases had an incubation period between 4 and 7 days. The proportion of cases infected 14 days (or more) before the development of symptoms was very small (0.6%) which has been used as the quarantine period.

The above information can be used to estimate the daily number of infections associated to a number of detected symptomatic cases. If 1000 new symptomatic cases were detected in day 15, it can be estimated that 6 of them were infected in day 1, 8 cases in day 2, and so on. If these simple calculations are made for a sufficient large period, it is possible to estimate the number of infectious transmissions with a time lag of 14 days.

We can illustrate the proposed process of modelling the dynamics of the virus in the diagram shown in Figure 2 where the situation of Portugal is used as an example.

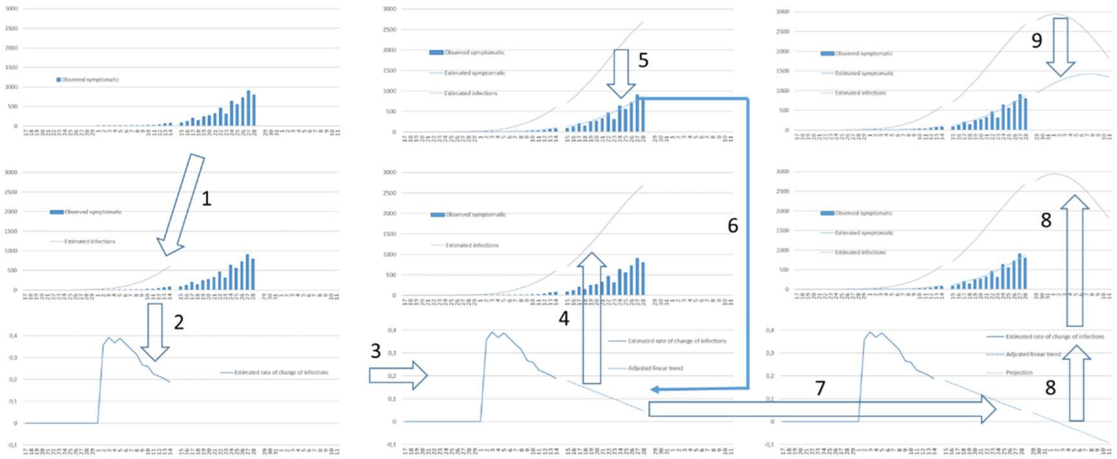


Figure 2. Steps of the model of analysis of the dynamics of COVID-19 proposed:

(1) From the daily number of symptomatic cases it is possible to estimate the daily number of infections 14 days before , using the Weibull distribution suggested by Baker et al. (2020) and the asymptomatic rate of 50% indicated by Mizumoto et al. (2020) and Chen et al. (2020).
(2) From the estimated number of infections the rate of change of infections is calculated.
(3) The last value of the rate of change from the previous period is used as the beginning of a linear trend of the rate of change for the current 14 day period with an unknown slope.
(4) A tentative slope is tried originating a corresponding curve for the estimated number of infections.
(5) From the estimated curve of infections, an estimated curve for symptomatic cases is calculated using in reverse direction the indicated Weibull distribution and the asymptomatic rate.
(6) A sum of squares of the differences between the observed number of symptomatic cases and the corresponding estimated numbers is computed and the system runs until the sum of squares is minimized, defining the slope of the rate of change of infections for the current 14 day period.
(7) Under the assumption that the trend is constant, the same slope is used for the projection period, the next 14 days.
(8) With the slope of the infection determined the curve of the estimated number of infections for the projection period is computed.
(9) Finally the curve of estimated symptomatic cases is computed from the number of infections as before.

The second model is based on the fitting of a global double exponential equation to the full time series data of the estimated number of new infections.

Literature cited:

Backer JA , Klinkenberg D, Wallinga J. 2020. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. Euro Surveill. 2020;25(5):pii=2000062. https://doi.org/10.2807/1560-7917.ES.2020.25.5.2000062

Mizumoto K, Kagaya K, Zarebski A, Chowell G. 2020. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. Euro Surveill. 2020;25(10):pii=2000180. https://doi.org/10.2807/1560-7917.ES.2020.25.10.2000180

Chen, T., Rui, J., Wang, Q. Zaho, Z., Cui, J., Yin, L. 2020. A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. Infect Dis Poverty 9, 24. https://doi.org/10.1186/s40249-020-00640-3.