COVID-19 spreading: a recursive model

Sergey O. Ilyin

A.V. Topchiev Institute of Petrochemical Synthesis, Russian Academy of Sciences,
29 Leninsky prospect, 119991 Moscow, Russia; s.o.ilyin@gmail.com

Abstract: This communication describes a recursive mathematical model of the spreading COVID-19 infection, which allows estimating the effectiveness of quarantine measures. This model takes into account the contagiousness of infected people during the incubation period of the disease and the conditionally non-contagiousness of sick people due to their isolation. The model was used to analyze the situation in eight countries and to find the viral transmissibility, which made it possible to give a brief prediction of the COVID-19 spreading.

The first mathematical models to predict the development of infectious diseases appeared in the early XX century.[1,2] In 1927, Kermack and McKendrick proposed the use of differential equations for calculations, dividing the human population into people susceptible to disease (S) and those who had already recovered (R).[3] The susceptible persons became infected (I) at some rate of transmission and then recovered at a different rate. Their model became known by the acronym SIR, which means that the model simultaneously calculates the number of susceptible, infected and recovered persons. This model served as a basis for the development of subsequent models by modifying the equations and adding to the calculation of other persons not belonging to the three specified basic categories, which allowed taking into account the features of particular diseases. Models have been created that consider the possibility of re-infection (SIS model)[4] and death (SIRD),[5] the existence of an incubation period (SEIR),[6] temporary immunity of infants (MSIR)[7] etc.

When a new infection appears, neither the set of population categories to be considered in the model nor the rate of transition of people from one category to another is known. Current information about the features of the novel coronavirus infection COVID-19 and the manner in which people perceive it and act should serve as a basis for building a model to describe the spread of this virus. These features are as follows: firstly, it is the presence of a long incubation period, during which the infected persons is contagious to others, and secondly, it is the isolation of discovered infected persons, which as a result become conditionally non-contagious. The combination of these two factors makes the new coronavirus infection unique. Usually the opposite is true: infected people are not dangerous to others during the incubation period and become contagious after its expiry. For this reason, a new model that considers these circumstances is needed to predict the COVID-19 spreading. Unfortunately, the duration of the immunity produced after recovery from novel coronavirus is currently unknown. In addition, there is also too little information to calculate correctly the rate of recovery of patients: a small

percentage of the population recovers just a week after infection, while the majority of people are sick for a long time. Therefore, the proposed model cannot be final, but it is necessary for forecasting and management decisions.

The COVID-19 spreading model is based on a set of parameters whose values are unique for each country due to differences in population density and humans' behavior, date of virus penetration and government actions. The set includes the following parameters:

d_0 is the date of the initiation of the epidemic; it is not the date of detection of the first infected person but the date of appearance of the first undetected (or detected too late) one;

d_1, d_2, d_3 are dates of change in the behavior of the population, e.g., due to the awareness of the reality of what is happening, the introduction of quarantine and its tightening;

t_D – is the average time from infecting to isolating the infected person, which is equal to the incubation period that I assume to be six days (from 5.2 to 6.4 days according to different sources);[8,9] theoretically, this parameter can be reduced by total testing of the entire population, but it is feasible only for small communities;

R_0, R_1, R_2, R_3 – are the viral transmissibilities that are equal to the average number of people who will be infected by one person before its isolation and depend on the behavior of the population at different stages of the epidemic; when R is less than 1.0 the epidemic fades, and vice versa;

r_0, r_1, r_2, r_3 – are the reduced transmissibilities that are equal to the average number of people who will be infected by one person per day: $r = R/t_D$; to suppress the COVID-19 spreading, r should be less than 0.167.

The evaluation of the virus spreading is based on the calculation of the following data:

$N_D(d_i)$ is the number of detected infected persons on d_i date, which equals the total number of infected persons six days earlier:

$$N_D(d_i) = N_T(d_i - t_D);$$

$N_T(d_i)$ is the total number of infected persons on d_i date, which is the sum of the total number of infected persons the day before and the number of new infected persons that, in turns, is equal to the product of the reduced transmissibility and the number of active infected persons the day before (taking into account that those who have been previously infected cannot re-infected):

$$N_T(d_i) = N_T(d_i-1) + r_0 \times N_A(d_i-1) \times (1 - N_T(d_i-1)/N_P),$$

where N_P is the total population;

$N_A(d_i)$ is the total number of active (undetected) infected persons on d_i date, which equals the difference between the total number of infected persons and the number of detected infected persons on the same day:

$$N_A(d_i) = N_T(d_i) - N_D(d_i).$$

At the start of the epidemic (d_0 date), $N_A(d_0)=1$, $N_T(d_0)=1$, and $N_D(d_0)=0$.

Thus, in order to calculate the virus spreading dynamics it is necessary to know the values of only two parameters – d_0 and r_0. In case of changing the behavior of the population from the date d_1, parameter r_0 changes its value from this date to become r_1. If the behavior changes again, a pair of d_2 and r_2 will appear, etc.

It is more difficult to model human losses correctly. Two more parameters appear:

L is the apparent lethality rate that is equal to the ratio of the number of deaths to the sum of those who died and recovered;

t_L is the average time from infection to death.

These two parameters depend on the efficacy of treatment and may vary as physicians gain experience and as hospitals overflow. Number of deaths on d_i date equals total number of people infected t_L days earlier multiplied by the lethality rate:

$$N_L(d_i) = N_T(d_i\text{-}t_L)\times L.$$

Due to the presence in the equation of two parameters (t_L and L) that have the same effect on the resulting value, the precision of their evaluation is lower than that of transmissibility. It should be understood that the less asymptomatic and mild cases of the disease have been detected the more the lethality rate is overestimated. The average time from infection to death was found to be about 8 days and this duration will be used to make calculations for all the countries.

The situation with predicting the number of recovered persons is even worse due to the appearance of an even greater number of independent parameters:

$$N_R(d_i) = N_T(d_i\text{-}t_M)\times k_M + N_T(d_i\text{-}t_S)\times k_S$$

where k_M, k_S, t_M, and t_S are the shares of mild and serious ill patients as well as the times from their infection to healing respectively; $k_M+k_S+L=1$.

The model equations are presented in the discrete form (instead of differential one) thanks to what for carrying out of calculations the given model is easy for reproduction in any spreadsheet editor. At first glance, it seems that the model does not take into account the existence of asymptomatic carriers of infection, but this is not true: since the share of asymptomatic carriers in the population does not change over time, their presence is taken into account implicitly by the value of the transmissibility. This model can be denoted by the SILRD abbreviation, which means that it takes into account Susceptible, Infected, isoLated, Recovered, and Dead persons.

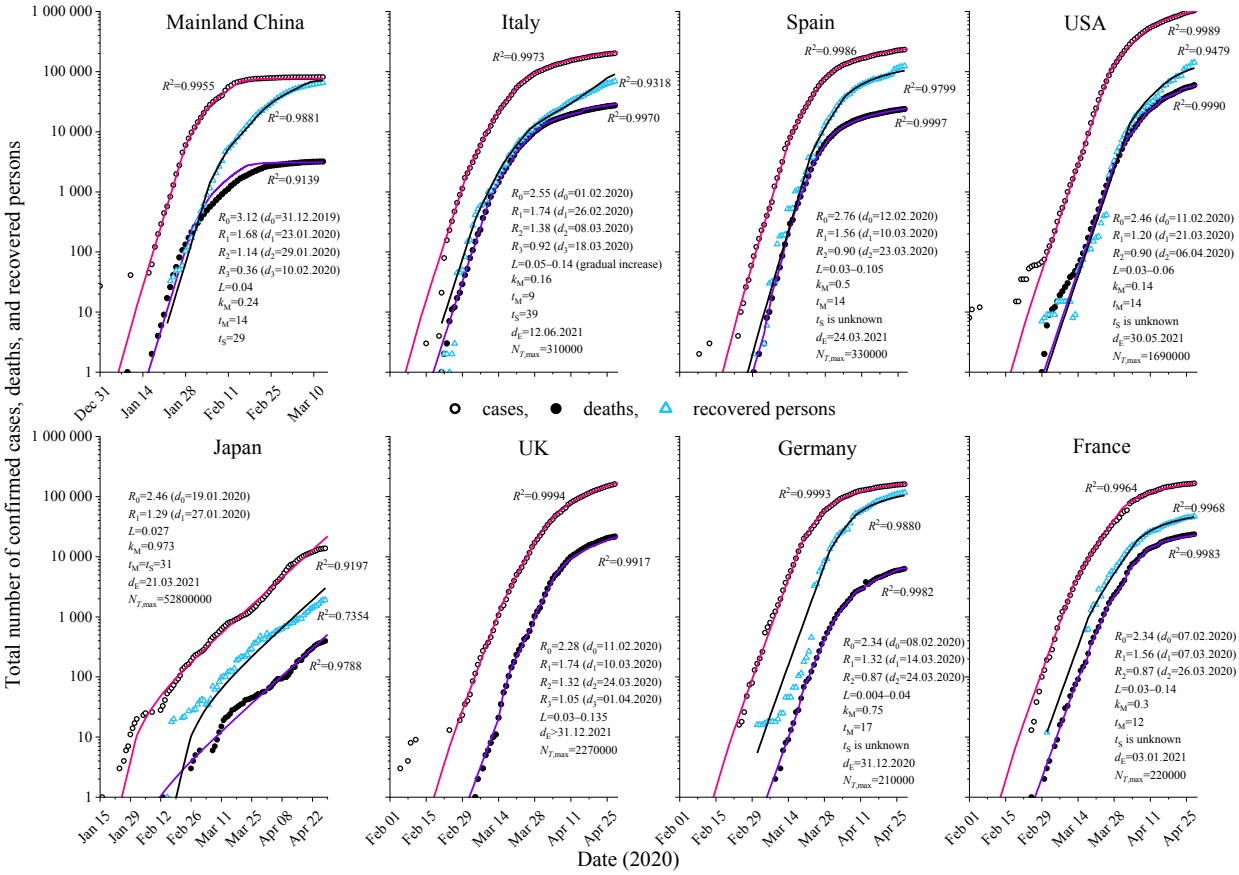

Figure. Time dependences of the total cases, deaths, and recovered from COVID-19. Dots show real data, while lines represent the result of calculations using the model.

Based on historical data on disease development in eight countries (People's Republic of China, Italy, Spain, USA, UK, Japan, France and Germany), the model was tested and most of its parameters were found (Figure). For all the countries at the start of the epidemic, the viral transmissibility was between 2.28 and 3.12. The greatest transmissibility was found in China, where one person infected three, probably because of higher population density. The introduction and progressive strengthening of quarantine measures resulted in a decrease in the viral transmissibility, which was noticeable six days later in the decline in the rate of new cases. All the countries have introduced quarantine measures gradually. The initial restrictions reduced the viral transmissibility to 1.20–1.74, which was not enough (it was necessary to achieve transmissibility of less than 1.0) and the virus continued to spread with acceleration. As a result, all the countries, with the exception of Japan, went for tighter measures, thus reducing the transmissibility to 0.87–1.14. Japan has focused on the timely detection and isolation of the infected persons. It can be concluded that this strategy does not work: the curve of the total number of cases for Japan alternately slows down and then accelerates again. This is the result of the fact that Japan has been successfully isolating most of the infected persons, but a few people remain non-isolated and they start the epidemic again. By contrast, the PRC further strengthened the measures, which resulted in a reduction of the transmissibility to 0.36 and a quick win over the epidemic (in 6 weeks according to the model).

To date, all the European countries (except the United Kingdom) have only managed to reduce the transmissibility slightly below 1.0. From a practical point of view, this means that the number of people falling ill on a daily basis in these countries is gradually decreasing, but so slowly that the end of the epidemic in them will not be sooner, at best, than the end of this year. At the same time, it should be understood that any alleviation of quarantine measures would lead to increased transmissibility and a resumption of accelerated virus spreading. To prevent this from happening after the quarantine restrictions have been removed, the viral transmissibility must remain below 1.0. By way of example, the original transmissibility was 2.55 in the case of Italy, therefore it is necessary that 61% of the Italian population be either infected and then recovered (provided the immunity produced is durable and strong) or vaccinated against coronavirus so that when quarantine is abolished, the transmissibility remains less than 1.0. At present, 0.33% of the Italian population has been infected according to official statistics. Statistics may not take into account asymptomatic and mild cases of the disease, number of which may be 4-50 times more (for the time being only by rumor) than that of the officially recorded cases. Even if this is true, the percentage of infected and recovered persons is still significantly lower than necessary: removal of quarantine will inevitably lead to a return of the growth rate of the number of infected people to almost their original level. In other words, rapid vaccine development is vital to defeating the disease in the near future.

Thus, the model allow making a forecast of the situation development and the conclusion about the effectiveness of quarantine measures. By way of example, determine the current number of active infected persons ($N_A(d_i)$), the approximate date of isolation of the last infected person (d_E), and the number of people that could eventually be infected under the current quarantine ($N_{T,max}$). According to the calculations, the efforts made by many European countries, the USA and Japan to stop the spread of the COVID-19 infection are not as effective as those implemented previously in the People's Republic of China. Most countries were able to achieve a daily reduction in the number of infected people, but even in these cases, the viral transmissibility remains high enough, which does not allow to defeat the epidemic within a reasonable time. At the same time, suppressing the epidemic, albeit slowly, allows time for vaccine development.

References

1. Brauer F, Castillo-Chávez C. Mathematical Models in Population Biology and Epidemiology. NY: Springer, 2001.

2. Daley DJ, Gani J. Epidemic Modeling: An Introduction. NY: Cambridge University Press, 2005.

3. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. P Roy Soc A Math Phys 1927; **115**: 700–721.

4. Nåsell I. The quasi-stationary distribution of the closed endemic SIS model. Adv Appl Probab 1996; **28**: 895–932.

5. Wang P, Jia J. Stationary distribution of a stochastic SIRD epidemic model of Ebola with double saturated incidence rates and vaccination. Adv Differ Equ 2019; **2019**: 1–6.

6. Li MY, Muldowney JS. Global stability for the SEIR model in epidemiology. Math Biosci 1995; **125**: pp.155–164.

7. Bichara D, Iggidr A, Sallet G. Global analysis of multi-strains SIS, SIR and MSIR epidemic models. J Appl Math Comput 2014; **44**: 273–292.

8. Backer JA, Klinkenberg D, Wallinga J. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. Euro Surveill 2020; **25**: 2000062.

9. Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. N Engl J Med 2020; **382**: 1199–207.