



Chaos, percolation and the coronavirus spread: a two-step model

Hua Zheng^{1,a} , Aldo Bonasera^{2,3}

¹ School of Physics and Information Technology, Shaanxi Normal University, Xi'an 710119, China

² Cyclotron Institute, Texas A&M University, College Station, TX 77843, USA

³ Laboratori Nazionali del Sud, INFN, 95123 Catania, Italy

Received: 3 September 2020 / Accepted: 25 September 2020 / Published online: 9 October 2020
© Società Italiana di Fisica and Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract We discuss a two-step model for the rise and decay of a new coronavirus (Severe Acute Respiratory Syndrome-CoV-2) first reported in December 2019, COVID-19. The first stage is well described by the same equation for turbulent flows, population growth and chaotic maps: a small number of infected, d_0 , grows exponentially to a saturation value, d_∞ . The typical growth time (aggressive spreading of the virus) is given by $\tau = \frac{1}{\lambda}$ where λ is the Lyapunov exponent. After a time t_{crit} determined by social distancing and/or other measures, the spread decreases exponentially analogous to nuclear decays and non-chaotic maps. Some countries, like China, S. Korea and Italy, are in this second stage while others including the USA are near the end of the growth stage. The model predicted 15,000 (± 2250) casualties for the Lombardy region (Italy) at the end of the spreading around May 10, 2020. Without the quarantine, the casualties would have been more than 50,000, one hundred days after the start of the pandemic. The data from the 50 US states are of very poor quality because of an extremely late and confused response to the pandemic, resulting unfortunately in a large number of casualties, more than 70,000 on May 6, 2020, and more than 170,000 on August 21, 2020. S. Korea, notwithstanding the high population density ($511/\text{km}^2$) and the closeness to China, responded best to the pandemic with 255 deceased as of May 6, 2020, and 301 on August 21, 2020.

1 Introduction

Chaotic models have been successfully applied to a large variety of phenomena in physics, economics, medicine and other fields [1–7]. In recent papers [8, 9], a model based on turbulent flows and chaotic maps has been applied to the spread of COVID-19 [10, 11]. The model has successfully predicted the rise and saturation of the spreading in terms of probabilities, i.e., the number of infected (or deceased) persons divided by the total number of tests performed. Also a dependence on the number of cases on the population density has been suggested [8], and the different number of fatalities recorded in different countries (or regions of the same country) was attributed to facilities overcrowding [9]. In this paper, we would like to extend the model to the second stage, i.e., the decrease in the number of events due to quarantine or other measures [12]. Different fitting parameters of the model are due to the different actions,

^a e-mail: zhengh@snnu.edu.cn (corresponding author)

social behaviors [13], population densities [8], pollution [14], etc., of each country, but there are some features in common, and it is opportune to first have a look at some data available on May 6 which we updated to the end of June and to the middle of August, 2020 when available.

In Fig. 1, we plot the number of positive (top panels) and deceased (bottom panels) as function of time in days from the beginning of the recordings. Some data have been shifted along the abscissa to demonstrate the similar behavior. Different countries are indicated in the figure insets. As we can see, all the EU countries display a very similar behavior including the UK notwithstanding Brexit. The USA data have been shifted by 38 days, which is the delay in the response to the pandemic resulting in the large number of fatalities. In contrast, S. Korea reacted promptly and was able to keep the number of positives and more importantly the death rate down. Among the EU countries, Germany shows the lowest number of deceased cases, which could be due to different ways of counting (for instance performing autopsies to check for the virus, as in Italy). In any case, the analysis in Ref. [9] shows that some regions of Italy have lower mortality rates (for instance the Veneto region which borders the Lombardy region—the highest hit) compatible to Germany. Thus, similar to [9], we can assume that different overcrowding of health facilities, retirement homes, jails, etc., might be the cause [13, 14] for the differences displayed in Fig. 1. The striking feature in Fig. 1 is that all countries seem to have reached saturation while the USA and Russia are still growing. Notice the striking similarity of the data in Fig. 1 to those of the model of population growth as originally proposed by Verhulst [15].

To counteract the pandemic, many countries have adopted very strict quarantine measures. Social distancing and other measures [9, 13, 14] decrease the probability of becoming infected. Thus we expect that countries with lower population density might have better and faster success in controlling the virus. On the other hand, if some countries adopt non-effective measures or are too late in responding, the lower population density might extend the problem for some time. Thus in order to better stress the efficacy of the quarantine, we have plotted in Fig. 2 the number of cases DIVIDED by the population density, assuming that it is much easier to perform social distancing if the population density is low. In Fig. 2, we see that S. Korea and Japan, even though their densities are rather high, $511/\text{km}^2$ and $334/\text{km}^2$, respectively, perform best. We should also consider that S. Korea and Japan are in close proximity to China, the epicenter of the infection [8, 9], while the other countries are located across a continent or an ocean giving further advantages in organizing a response. The last data points for China reflect an adjustment to the death rate in Wuhan, which probably had similar problems to those of the Lombardy region in Italy [9]: We will not be surprised to see future corrections.

There is a ‘hidden’ parameter in Figs. 1 and 2: the number of tests performed daily. Zero tests, zero cases and, according to some politicians, no problem but then the hospitals get filled with sick people and we have a pandemic. In order to have realistic information on the time development of the virus, it is better to calculate the total number of cases DIVIDED by the total number of tests. This defines the probability to be infected or the death rate probability due to the virus. We stress that such probability may be biased since often the number of tests is small and administered to people who are hospitalized or show strong signs of the virus [8, 9]. The values we will derive must be considered as upper limits, but the time evolution should be realistic.

Not all the countries provide the number of tests performed daily (China). In Fig. 3, we plot the probabilities versus time for the same countries as in Figs. 1 and 2. As we can see, some cases show a smooth behavior indicating prompt and meaningful data taking. Large fluctuations or missing data are also seen at the beginning, which means that the number

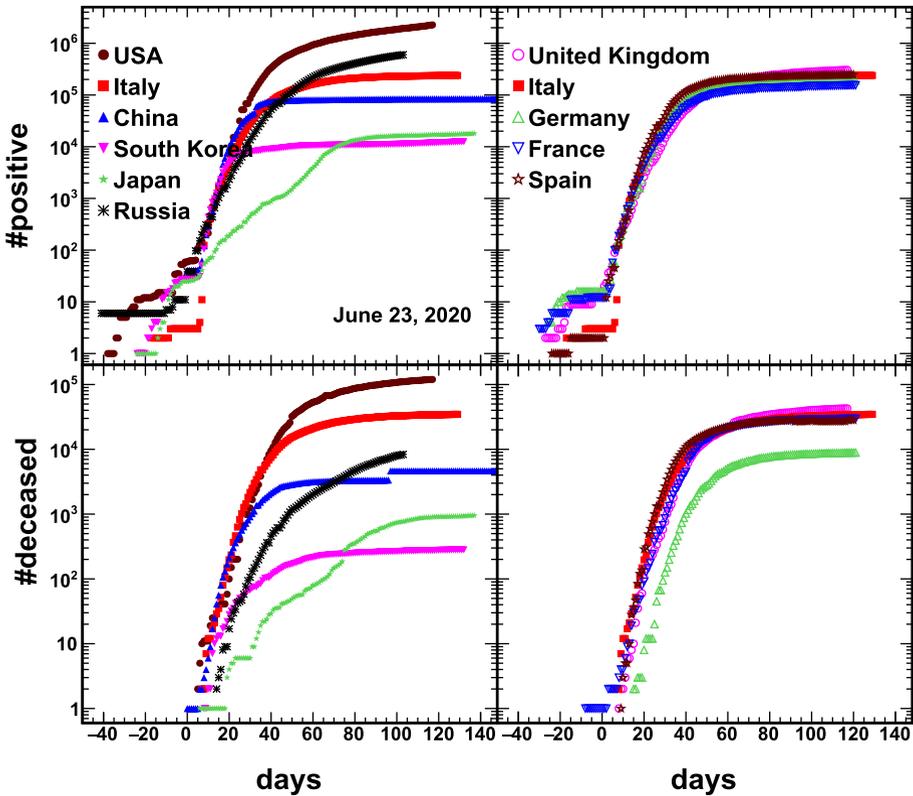


Fig. 1 Number of positive (top panels) and deceased (bottom panels) as function of time for different countries are indicated in the inset. Time $t = 0$ was suitably chosen to match the exponential growth for the number of positive, and it was kept the same for all the other plots, this figure and Figs. 2 and 3

of early daily tests was very small. All countries show a decreasing behavior at long times both for the positive and deceased cases suggesting that the mitigation procedures adopted can be effective, but with different rates. S. Korea and Japan display a similar behavior but with much lower values. Germany ‘performs best’ among the EU countries analyzed here, most importantly regarding the death rate. The USA, which was showing an increasing trend in Figs. 1 and 2, displays a decreasing probability but at a lower rate and the increase in the previous figures may be attributed to the increase in the test number. A similar behavior can be noticed for Russia but with lower probabilities for both cases. In particular, we may notice that Russia is ‘performing best’ among all the countries even better than Germany. This is in part due to the fact that, as for other northern and colder countries, they had more time to organize a strong response to the pandemic building upon experience gained in the Italian case as well. To this, we have to add the clearly good health structures and their longstanding research on virus propagation and control. It is no surprise that they were the first to announce a vaccine for COVID-19, <https://www.sciencemag.org/news/2020/08/russia-s-approval-covid-19-vaccine-less-meets-press-release>. If these premises are maintained and ‘scientifically’ corroborated, a vaccine may be available before the month of November 2020 but only for the few countries who have enough resources to afford it like the USA and Russia of course. If this will happen, we may expect a similar behavior to the one in Fig. 3

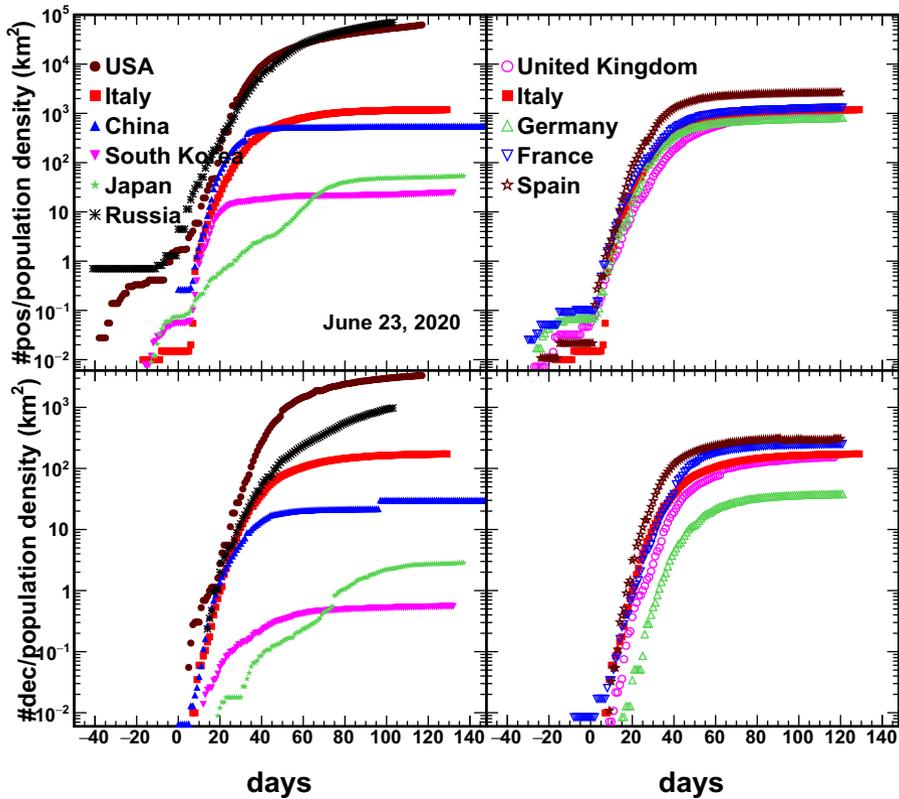


Fig. 2 Number of cases divided by the population density of each country versus time. Compare to Fig. 1

for Russia, a large number of positive due to the vaccine and a relatively small number of deceased primarily controlled by either preexisting conditions or the impossibility for a reason or another to get the vaccine in time.

2 The model

We have discussed and applied the first stage of the model in Refs. [8,9]. We briefly recall it and write the number of people (or the probability) positive to the virus (or deceased for the same reason) as [1,2,4–9,15]:

$$\Pi(t) = \frac{d_0 d_\infty}{d_0 + d_\infty e^{-\lambda t}} \tag{1}$$

In the equation, t gives the time, in days, from the starting of the pandemic, or the time from the beginning of the tests for the virus. At time $t = 0$, $\Pi(0) = d_0$ which is the very small value (or group of people) from which the infection started. In the opposite limit $t \rightarrow \infty$, $\Pi(\infty) = d_\infty$, the final number of people affected by the virus. Equation (1) has the same form observed in Figs. 1 and 2, but in reality it should be applied not to the number of positives (or deceased) but to their probabilities, i.e., the number of cases divided by the total number of tests. The main reason for this definition is to avoid the spurious time dependence due to

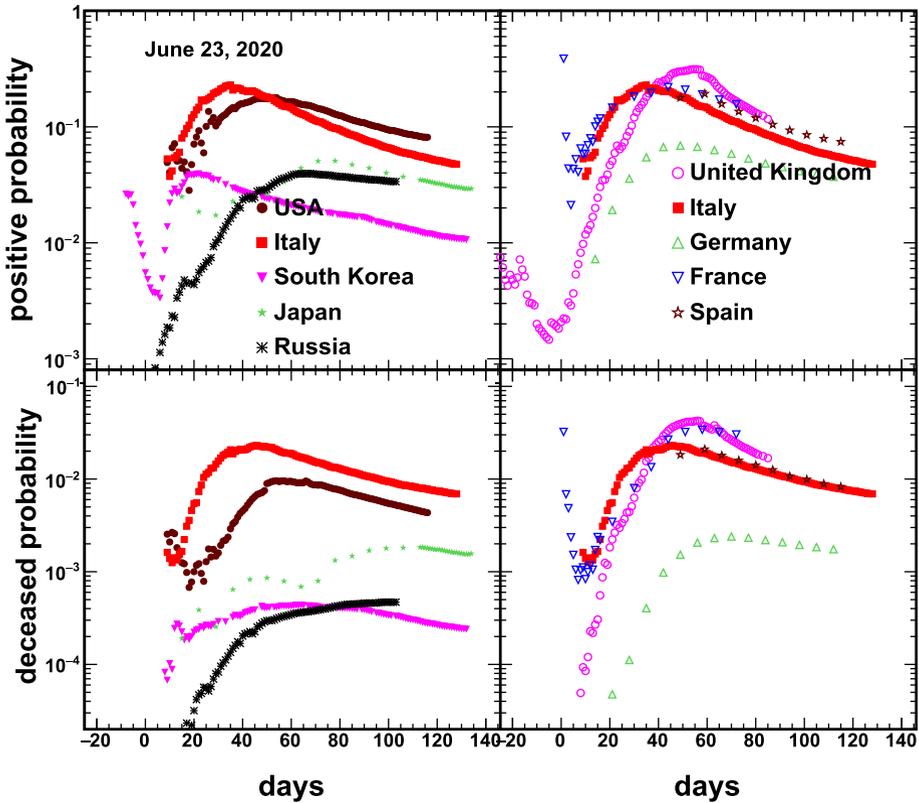


Fig. 3 Probabilities versus time for the countries indicated in the inset. Some countries stopped providing the number of tests performed daily (France on May 5th and the UK on May 22nd), other countries are providing this information periodically (Spain, Germany)

the total number of tests, which varies on a daily basis and very often not in a smooth way [8,9]. In Fig. 3, we have plotted the probabilities for different countries since the data are available. It is important to stress that the information on the total number of daily tests is crucial and should be provided also to avoid questions on data handling. If we treat Eq. (1) as a probability, then we expect to saturate at d_∞ at time t_{crit} . At later times, if social distancing is having an effect, we expect the probability to decrease and eventually tend to zero. In Fig. 4, we see exactly such a behavior for the cases of two Italian regions: Lombardy and Sardinia [9], <https://github.com/pcm-dpc/COVID-19>. For times larger than t_{crit} , the decrease is exponential and can be described as for nuclear decays and non-chaotic maps [1,12]:

$$\Pi(t) = d_\infty e^{-\alpha(t-t_{crit})}. \tag{2}$$

α and t_{crit} are fitting parameters. Values for the Lombardy region are $\alpha = 0.0268(0.025)d^{-1}$ and $t_{crit} = 39(42)d$ for the positives (deceased). We can infer the decay time as $\tau_d = 1/\alpha = 37(40)d$ suggesting that roughly τ_d after the maximum the pandemic should be over, i.e., $t_{max} \approx t_{crit} + \tau_d = 76(82)$ days from February 24, 2020. To demonstrate the predictive power of the model, we updated the data for the Lombardy case in “Appendix”, Fig. 19. A similar agreement to other Italian regions is found [9].

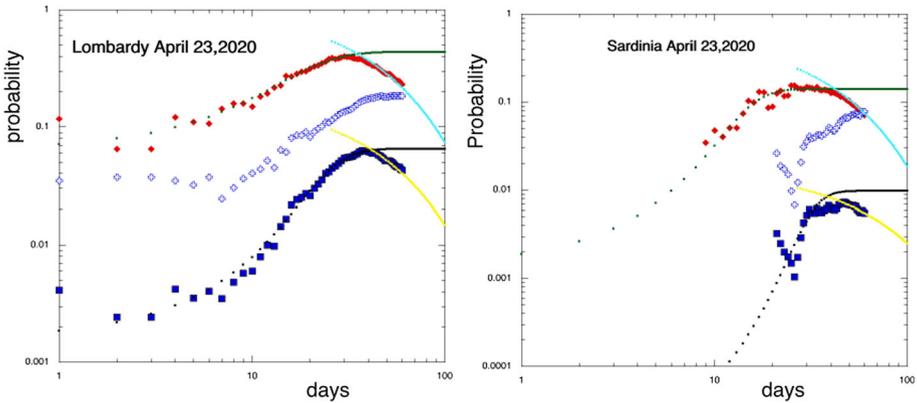


Fig. 4 Probability for positive (rhomb symbols) or deceased (square symbols) versus time in days for the Lombardy (left) and Sardinia (right) regions. The open crosses give the ratio positive/deceased and reach almost 20% for Lombardy [9]. The continuous points are obtained from Eq. (1) and the exponential decay from Eq. (2). Updated data are given in Fig. 19

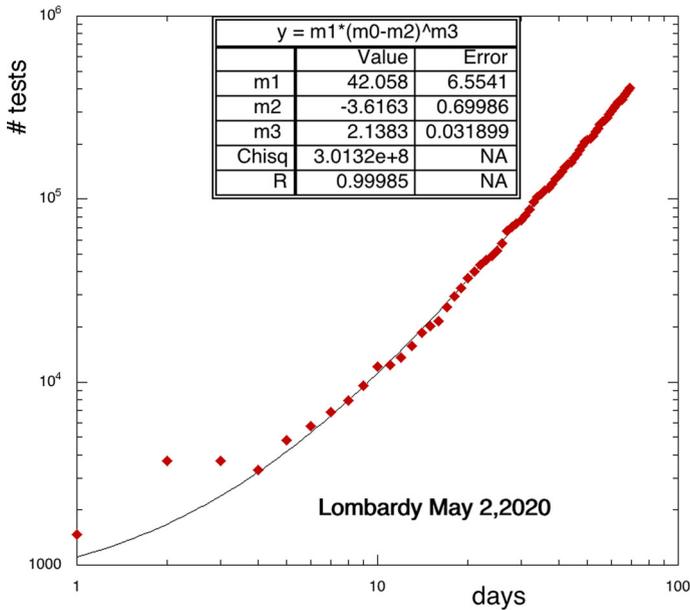


Fig. 5 Total number of tests as function of time for Lombardy. The fitting function and its values are displayed in the insets

From Figs. 4 and 19, it is quite easy to derive the value of t_{crit} given by the maximum. This value differs slightly for the positive and the deceased as well as for the different regions. Thus it is important to have enough data to perform best fits using Eqs. (1) and (2). The value of t_{crit} depends on many factors including the population density, the weather temperature, humidity, etc., and especially social distancing or any other measures used to mitigate the pandemic. If no measures are adopted (herd immunization or natural selection approach), such as occurred in some countries like Sweden (and the UK at first), then we expect the

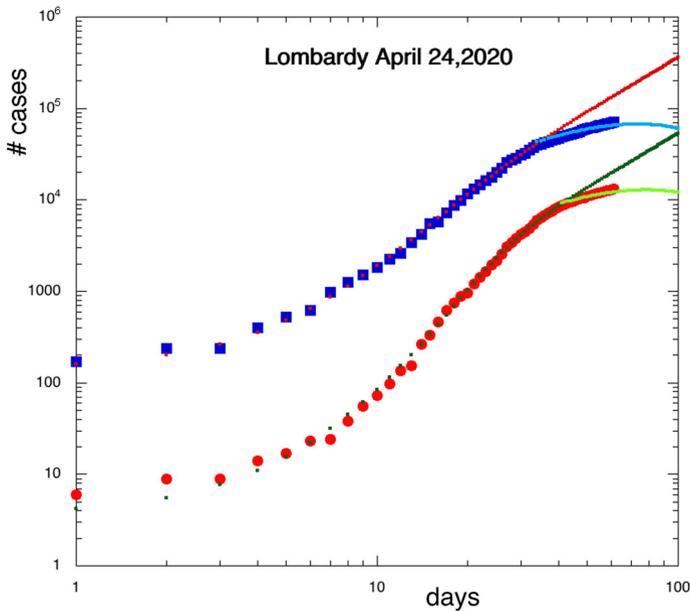


Fig. 6 Predicted cases with and without quarantine as function of time, see text. Data for positive and deceased are given by the square and circle symbols respectively

plateaus in Figs. 3 and 4 to last longer but eventually the process will be described by Eqs. (1) and (2). However, in such cases we may also expect to be flooded by positives and deceased persons jeopardizing the health structures and harming the sanitary personnel [9]. A country like Sweden with excellent sanitary structures and low population density ($25/\text{km}^2$) may succeed in this task, but the same attempt in the UK ($279/\text{km}^2$) was a disaster and quickly abandoned as can be seen from Figs. 1, 2 and 3. In particular in Fig. 3, we see that the UK has the largest probabilities, <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>. Of course the predictions only have validity if the conditions are not changed, for instance relaxing the quarantine too soon. If these conditions are modified, then we may have a rapid increase in the cases again and return to the original curve given by Eq. (1); such a behavior might be noted in Fig. 3 for Japan. At the time the Olympics 2020 were under discussion, Japan interrupted the COVID-19 testing as can be seen from the plateaus in Figs. 1 and 2 lasting approximately 15 days. Thus it is important to understand when to relax the measures, and for this reason, we have plotted in Fig. 4 two cases. Lombardy is the worst case in Italy with more than 14,000 deceased as of May 6, 2020. In the figures, we can see that the probabilities are much lower for Sardinia, which could be regarded in some sense as the future of what should eventually happen in Lombardy because of the quarantine. The population density of Sardinia is relatively low, $69/\text{km}^2$, and it is an island away from the mainland. This situation is in many respects very similar to S. Korea with lower population density. Thus, the measures could be relaxed in Sardinia following the example of S. Korea after careful instructions to the population and random every day testing to search for positive cases and isolate them. This will provide crucial information on the social behavior and on the virus spread. We will show below that the model predicts a small number of positive and deceased for Lombardy around or after May 10, 2020; thus, shelter at home might be extended up to that day. It would be important to send some signals to

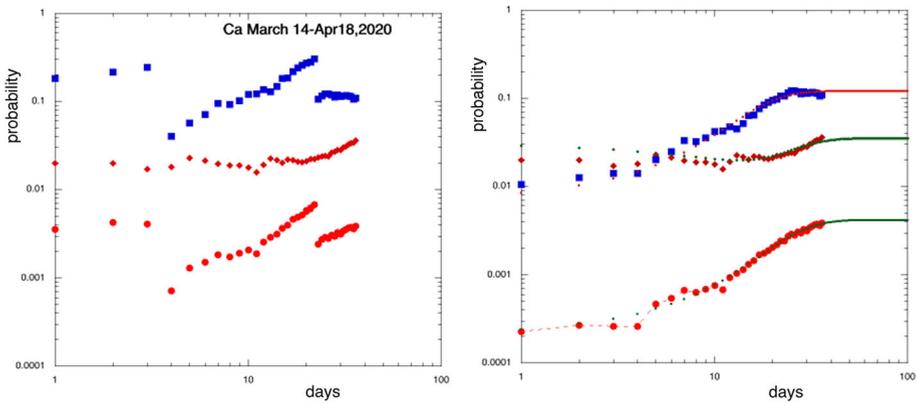


Fig. 7 Positive (square symbols) and deceased (circle symbols) probabilities versus time in days. The rhomb symbols represent the ratio deceased/positive independent on the number of tests. The right panel is obtained after renormalization, see text

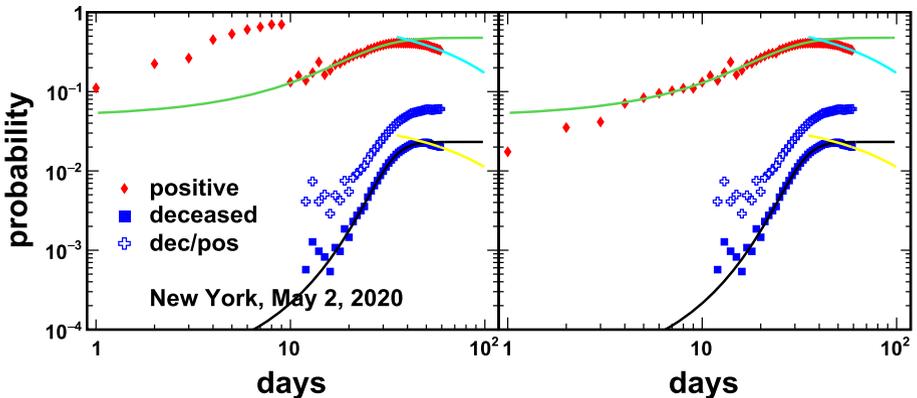


Fig. 8 Same as Fig. 7 for the state of New York. See Fig. 20 for an update

the population of return to normality after months of sheltering by organizing for example sportive events in Sardinia. The Italian national sport, “Serie A”, might organize 2–3 games per day in different Sardinian towns, with empty stadiums and broadcasted live. Other limited activities but strongly controlled could be allowed in less affected regions such as Calabria, Abruzzo and other southern Italian regions discussed in Ref. [9]. Releasing all measures for the entire country at the same time might be not too wise. Looking at other countries experiences, we would suggest that quarantine should not be released before the probability for positives is less than 4% (the maximum of S. Korea, Fig. 3). Below such a value, the other countries may follow the S. Korean approach, but if they are not organized to do that, reopening too soon may be dangerous.

The model describes very well the data and might be used for the everyday control on the resurgence of the pandemic. It offers another great advantage: We have described a way to eliminate misleading inputs due to the number of everyday tests. We can proceed in the inverse direction in order to predict the total number of deceased and positive cases. The task that we have now is much easier, and it is the prediction of the daily tests for each

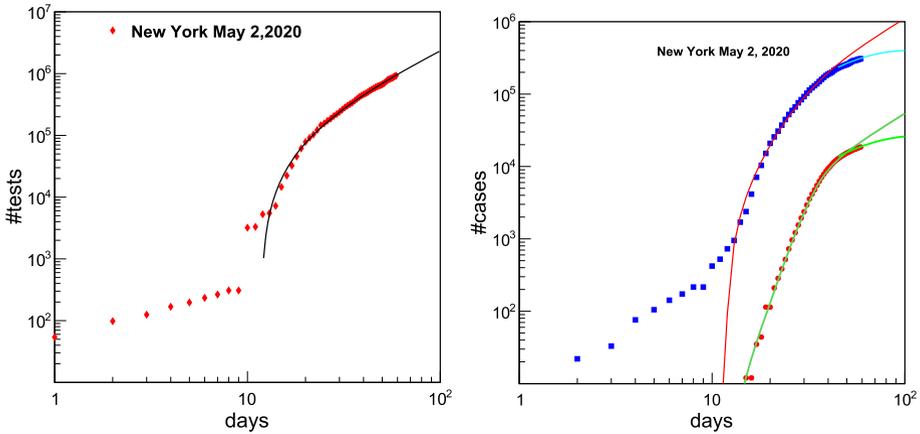


Fig. 9 Same as Figs. 5 and 6 for the state of New York

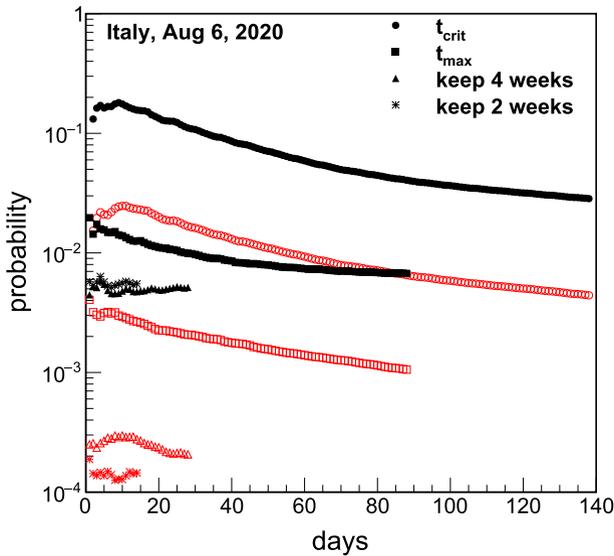


Fig. 10 Rescaled positive (full symbols) and deceased (open symbols) probabilities as function of time for Italy. The different scaling times are displayed in the figure. Notice that the values when keeping 2–4 weeks are constant, about 0.5% for positive, or keep decreasing for deceased. This is an indication that the pandemic is under control but still present. Measures such as wearing a mask, hygiene, social distancing, etc., must still be enforced until these probabilities are zero

case. As we have seen from Figs. 1, 2 and 3, there were some wrong decisions taken by the different countries at the beginning of the pandemic (apart S. Korea and Japan) resulting in a very small number of tests. After 1–2 weeks, the number of tests per day was increased and eventually become constant. It is this behavior we have to predict in order to extend our model to the total number of cases. In Fig. 5, we plot the total number of tests versus time from the beginning of the recordings for Lombardy (February 24, 2020). We have fitted the data with a power law function as indicated in the figure, but any other suitable function

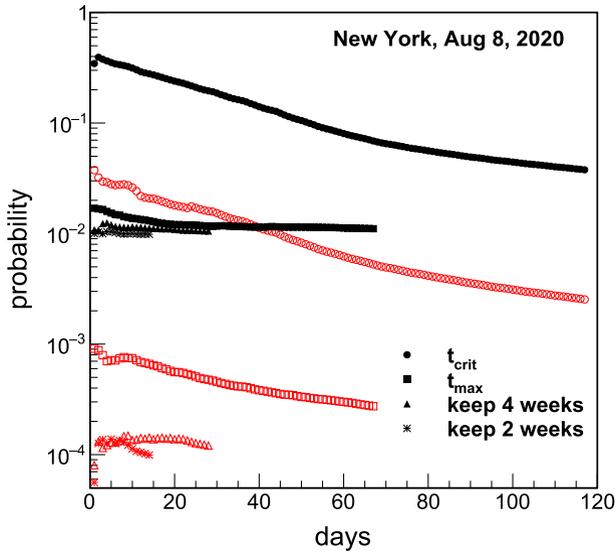


Fig. 11 Same as Fig. 10 for NY

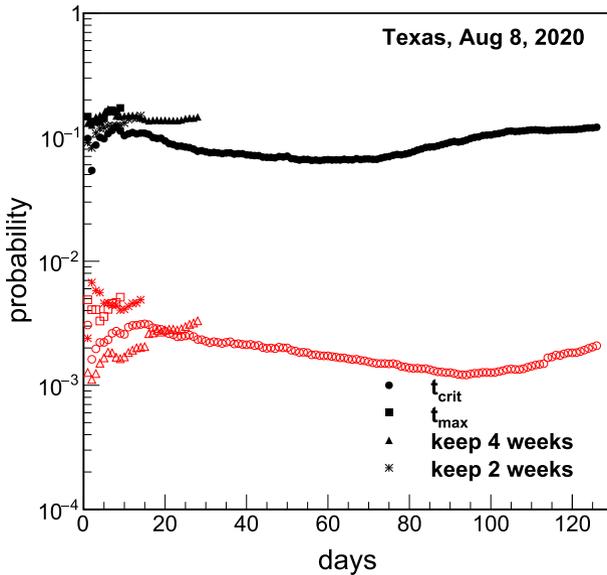


Fig. 12 Same as Figs. 10 and 11 for Texas. Notice the complete different behavior and higher non-decreasing probabilities when keeping the last 2–4 weeks respect to Italy and NY. This is essentially ‘herd immunization’ as discussed in the text. It is equivalent to the strategy adopted in some countries like Sweden or what happened in the 1918 pandemic (Spanish flu). Recall that the number of deceased in the USA because of the Spanish flu was more than 650,000. Scaling by the current USA population (3 times the population in 1918) may result in about 2 millions casualties if no vaccine is found

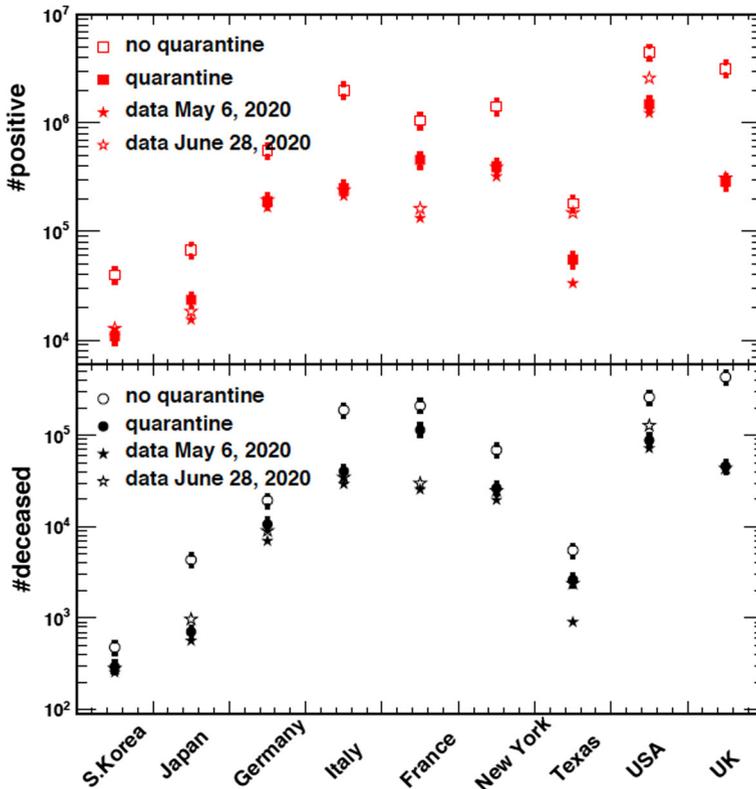


Fig. 13 Number of positives (top panel) and deceased (bottom panel) for the countries indicated on the abscissa. Predictions without quarantine measures refer to June 28, 2020. The numerical data are reported in Table 1

might do as well. As we see from the figure, the Italian data are fitted very well with a small error on the fitting. Fits performed to other countries give a power exponent ranging from 0.73 (S. Korea) to 4.1 (UK). This is also an indication of how the response to the pandemic has varied. In the ideal case, we expect the power to be about 1, the value for the UK suggests some change in strategy (from, i.e., herd immunization to quarantine), and because of such a high value we were not able to make predictions on the total number of tests say 50 days after May 6, 2020. Using more recent data for the UK—see Fig. 3, allows us to make predictions as discussed below.

Multiplying Eq. (1) or (2) by the predicted number of tests from Fig. 5 gives the total number of predicted cases which are compared to the data in Fig. 6. We assume a conservative 15% error in our estimates due to the different fit functions. Most important, an error is coming from laboratory testing with current methods [16–21]. Without social distancing, using Eq. (1) gives 360,000 ($\pm 54,000$) for the positives and 53,000 (± 7950) for the deceased 100 days after the beginning of the pandemic in Lombardy. If the exponential decay given by Eq. (2) is taken into account (due to the quarantine), the values decrease to 80,000 ($\pm 12,000$) and 15,000 (± 2250), respectively; thus, about 38,000 saved lives in Lombardy alone! There is an important difference between the two stages: If the first stage alone would be at play, the pandemic may continue after the 100 days and eventually slow down at longer times, see the

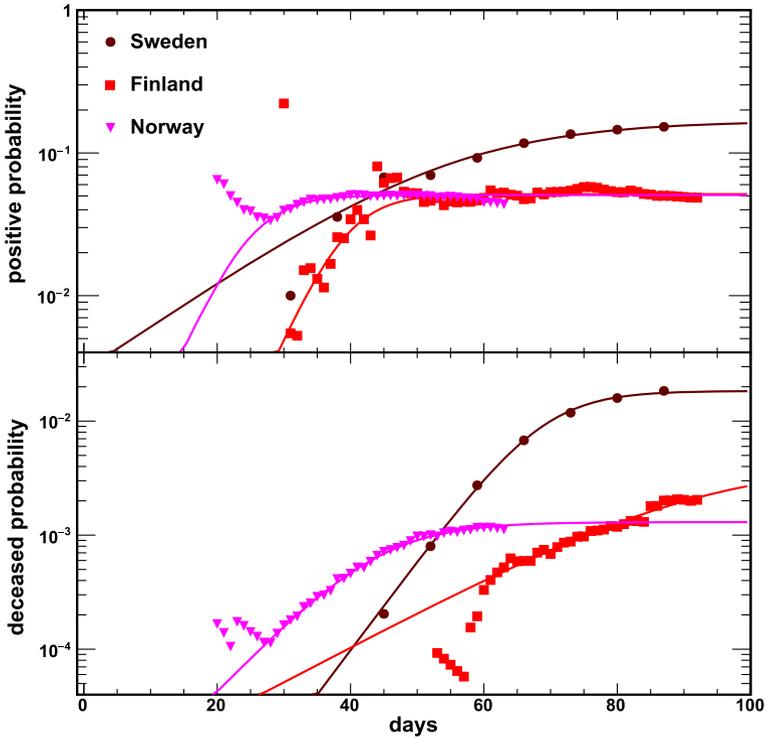


Fig. 14 Probabilities as function of time for the countries indicated in the inset. Sweden is adopting the natural selection option or herd immunization resulting in higher probabilities compared to nearby countries. Different starting data depend on which day the complete information needed for the plot was released. Updates are given in Fig. 23

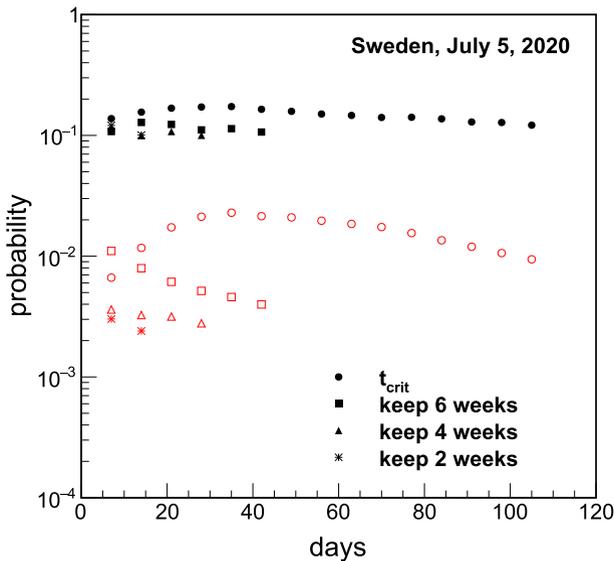


Fig. 15 Same as Figs. 10, 11 and 12 for Sweden. Notice the similar behavior to the Texas cases

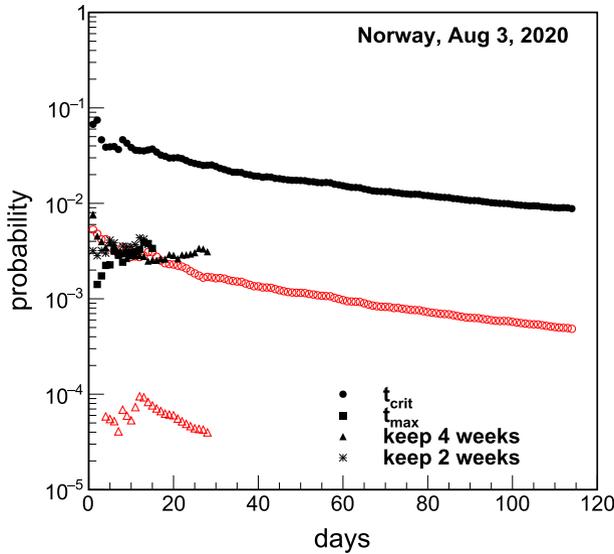


Fig. 16 Same as Fig. 15 for Norway. Notice that no deceased cases are recorded two weeks before August 3, 2020! This is in striking contrast to the Sweden case

Sweden case discussed below. Recall that the Spanish flu started in 1918 and lasted almost 36 months with an enormous death toll, <https://www.washingtonpost.com/graphics/2020/local/retropolis/coronavirus-deadliest-pandemics/>. Because of the second stage, now the predicted values are given by the maxima in Fig. 6, these occur 76 and 82 days, respectively, after the start of the pandemic recording, i.e., May 10 and 16, 2020, respectively. These values are close to the sum of t_{crit} and τ_d reported above.

If we assume a power law to reproduce the available data for the number of tests, Fig. 5, then we can write the total number of cases in the second stage as:

$$\#cases = m_1(t - m_2)^{m_3} d_\infty e^{-\alpha(t - t_{crit})}. \tag{3}$$

The fitting parameters m_{1-3} are reported in Fig. 5 for Lombardy. To find the maximum of Eq. (3), we simply equate its derivative to zero:

$$t_{max} = \frac{m_3}{\alpha} + m_2. \tag{4}$$

Using the empirical relation above connecting t_{max} and t_{crit} , we get:

$$t_{max} \approx t_{crit} + \frac{1}{\alpha} \rightarrow t_{crit} = \frac{m_3 - 1}{\alpha} + m_2. \tag{5}$$

This relation is very useful especially when the data do not show the exponential decrease since it reduces the number of free parameters entering Eq. (2). Similar relations can be derived for different parameterizations for the total number of tests.

3 The US states

Figures 1, 2 and 3 show that the USA was hit hard by the COVID-19 resulting in different responses from the different states. In this section, we will analyze some of these

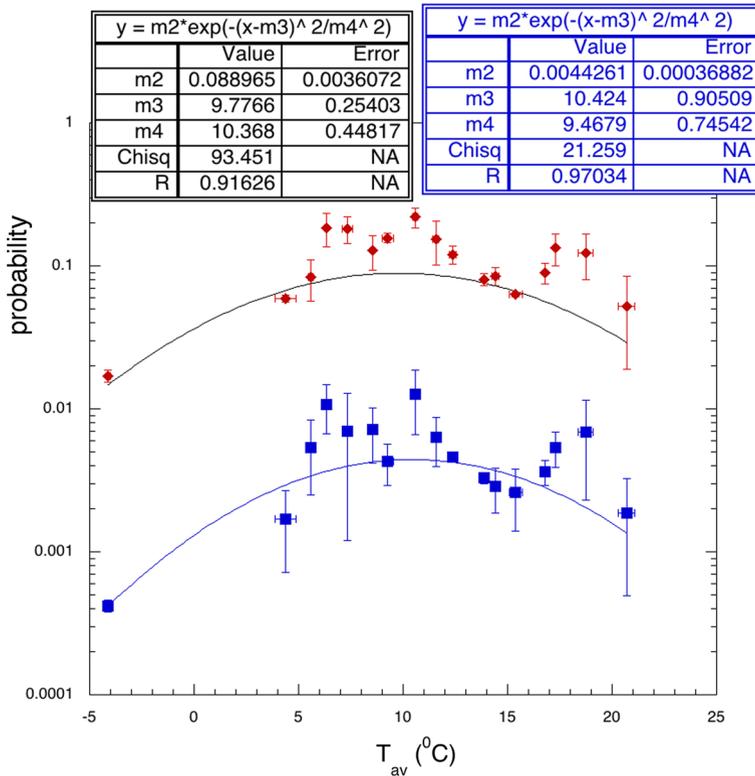


Fig. 17 Positive (rhomb symbols) and deceased (square symbols) probabilities versus average temperature of the US states in the spring season. The data to calculate the probabilities were collected on May 3, 2020. To get better statistics, averages were performed over states differing about 1 °C. Enforced Gaussian fits are also included and the fit parameters are given in the inset

states, and more analyses can be found in the appendices or available from the authors. In Fig. 7, we plot the probabilities for the state of California (Ca) for the period indicated in the inset (compare to Figs. 3 and 4). The discontinuities are due to the change in the number of tests performed daily. Notice that March 14, 2020, coincides with the quarantine declaration in Italy; thus, it was not a surprise that the virus spread quickly. Fortunately, the San Francisco mayor and the California governor placed strict restriction as early as March 6 without waiting for better testing, <https://www.sfdph.org/dph/alerts/files/HealthOfficerLocalEmergencyDeclaration-03062020.pdf>. This action saved a large number of lives and kept the ratio deceased/positives very low, compare to Fig. 4. We can correct in some cases for the low number of tests. Large data taking has a better statistical value; thus, we renormalize the data where the jumps occur to the value at later times. In the right panel, we display the result of the renormalization together with the fit using Eq. (1).

The hardest hit state was New York. In Fig. 8, we display the probabilities together with the fits using Eqs. (1) and (2), compare to Figs. 4 and 7. The ratio deceased/positives seems smaller than the Lombardy one; however, particular attention should be paid to the counting methods, and some confusion might arise if the data refer to the state of New York (NY), <https://coronavirus.jhu.edu/map.html>, or to New York city (NYC), <https://covidtracking.com/data/>

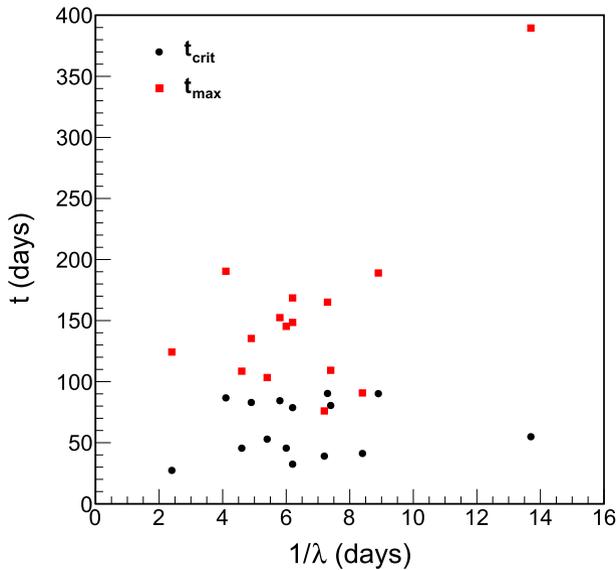


Fig. 18 t_{max} & t_{crit} versus $1/\lambda$ (the Lyapunov time) for all cases analyzed in this paper. The largest values of t_{max} & $1/\lambda$ refer to Sweden while the largest value of t_{crit} refers to the USA. All the numerical values of this figure are reported in Table 3

[state/new-york#historical](#), the difference being roughly 5000 deaths since most cases are in NYC. The bending down of the curve is evident, and we can make a prediction using Eq. (2). The resulting fit is displayed in Fig. 8; it follows well the available points but further confirmation will be given by future data. In the “Appendix”, Fig. 20, we compare the data available on June 23, 2020, to the model confirming its validity.

Using the predicted number of tests for NY given in Fig. 9, left panel, and the probability fits from Eqs. (1) and (2) displayed in Fig. 8, we can predict the total number of cases as for Lombardy. The results are plotted in Fig. 9, right panel, for the first 100 days from the start of the recordings.

4 Rescaling

In the previous sections, we have shown how the pandemic spread can be easily modeled in terms of growth and decay, and in particular we have defined two typical times, t_{crit} and t_{max} , Eqs. (2)–(5). We have also seen that low-quality data is a serious problem to make predictions. To overcome this problem, we have renormalized low statistics data to higher one with some success, Figs. 7 and 8. Day by day data are rather fluctuating; thus, it is important to average over some period, and we have chosen to use the cumulative data from a suitable defined initial time. In order to overcome the low statistics problem and, most importantly, see if quarantine actions are effective, we can rescale the results to later times, i.e., ignore the available data say to a suitable chosen time t' . In particular, we can define a

new starting time, number of cases and number of tests as:

$$\begin{aligned}t_n &= t - t'; \\ \#positive_n(t_n) &= \#positive_n(t) - \#positive_n(t'); \\ \#cases_n(t_n) &= \#cases_n(t) - \#cases_n(t').\end{aligned}\quad (6)$$

Similarly for the #deceased. The shifting of the data to later times will give important information avoiding fluctuations due to data taking or to different quarantine measures. From these quantities, we can easily calculate probabilities as before. The model has two important parameters, t_{crit} and t_{max} ; thus, obviously these times are candidates for t' in Eq. (6). In Fig. 10, we plot rescaled data for the Italian case. As expected, considering the data after t_{crit} we get the exponential decay indicating that the quarantine is effective. In the figure, we see exactly this with the usual small delay in the deceased case decay. If we shift the data to t_{max} , we expect to see the tail of the exponential decay, i.e., a rather small probability, around 1% and 0.1%, respectively, for the positive and deceased. These values are similar to the S. Korea case indicating that the pandemic is somehow under control but effectively not completely eliminated; thus, no time to relax yet. We can further improve our analysis by keeping the last few weeks to the time of writing. In Fig. 10, we show that the number of positives saturates around 0.5% 2–4 weeks before August 6, 2020, while the number of deceased is approaching zero.

We can confirm this behavior looking at the NY case, Fig. 11, where a strict quarantine has been imposed and the typical times t_{crit} and t_{max} can be defined. Similarly to Italy, there is an exponential decay after t_{crit} and the tail of the exponent after t_{max} . For the last 2 and 4 weeks before August 8, 2020, the data saturates around 1% for the positive and 0.015% for the deceased, which decreases to 0.01% at later times.

The state of Texas has not followed the strict quarantine as for the NY case, and the results are plotted in Fig. 12. As we can see after a small decrease, which allowed us to define the two typical times, activities were reopened before the number of positives was below 4%. The probability increased again and started fluctuating around 20%. This is more clearly seen in the rescaled data after t_{max} and keeping the last 2 and 4 weeks. More dramatically, we observe an increase in the death probability up to 0.6%. These are clear indications that a good occasion to bring the pandemic under control was lost and that the situation may only get worse with the arrival of the fall season and the reopening of many activities such as schools, universities and sportive events. Fortunately, the lower population density in TX has kept the actual number of deceased away from Biblical proportions, but this may not be guaranteed with the arrival of the flu season. Other cases are reported in Figs. 21 and 22.

5 Conclusions and outlook

We have proposed a two-step model for the rise and decay of the pandemic due to the COVID-19. The model needs some input parameters to predict the time evolution up to the saturation of the probability as in Eq. (1). Once the plateau is reached, given by the d_∞ parameter, the probability remains constant for some time depending on the quarantine measures or other environmental factors [13, 14]. For the Italian case, the first test was published on February 24 and Eq. (1) was fitted on March 10 before the quarantine was announced, i.e., March 14 [8]. The plateau was reached around March 24 as predicted by the model. These dates suggest that the quarantine was not effective in reducing the maximum probability and the time when this was reached. The quarantine became effective roughly 10 days after saturation. Thus, we can estimate that it takes more than 3 weeks before the quarantine gives an effect and the

probabilities start decreasing; this is the value of t_{crit} entering Eq. (2). We can assume that if the quarantine was announced say 14 days earlier, then the exponential decrease, Eq. (2), would have intercepted the rise, Eq. (1), earlier resulting in smaller probabilities. This is what happened to S. Korea and Japan, and it would explain the differences among countries: The later and the more feeble the quarantine, the higher the probabilities and the longer the time it takes to return to (quasi) normality. From these considerations, we can estimate the time it takes for other countries also if the probability decrease is not seen yet. After reaching the top of the probability, see Fig. 3, it took roughly 10 days for Italy to see the decrease. In Fig. 13, we plot the predicted total number of positives (top panel) and the deceased (bottom panel). Countries, which did not provide the number of daily tests (Spain, China), were not analyzed including the UK because of the large increase in testing especially at later times in corresponding to the time of their Prime Minister's hospitalization. At later times, some cases were added (UK) and updated to June 23, 2020.

From the updated data, we may notice the overall good agreement with the model with some exceptions so far. France is the most notable with the model over-predicting the data. The model under-predicts the USA case, while it is in good agreement with the states of Texas (for the deceased) and New York. The problem is that the USA did not react to the pandemic as a whole but state by state; this explains why the prediction for the deceased in Texas is good while the positive is underestimated: reopening too early! Recall that there is a time delay between the positive and the deceased. Very soon economic and political reasons influenced heavily the response. As an example, wearing a mask when in public became a political problem [13], which would be ridiculous if not tragic [22]. In Figs. 21 and 22, we discuss different cases for different party affiliations for each US state.

Sweden decided to follow a different path of non-imposing the quarantine (herd immunization or natural selection), a choice that could be justified under the assumption that the vaccine will not be available soon enough. In Fig. 14, we plot the probabilities for Sweden, Finland and Norway since they are bordering countries. The probabilities are quite different especially regarding the death rate. We predict for Sweden about 74,750 ($\pm 1.1\text{e}4$), 8195 (± 1230) positives and deceased, respectively, on June 28, 2020. Since there is no quarantine, we are not able to estimate t_{crit} and the decay rate. On the same day, using Eq. (1), we predict for the other countries the values 18,271 ($\pm 2.7\text{e}3$) and 1495 (± 224) for Finland, 15,225 ($\pm 2.3\text{e}3$) and 391 (± 59) for Norway. Thus, we see that herd immunization takes a heavy toll not explained by the larger Swedish population (a factor of 2 respect to the other countries considered), and it will be very difficult to explain this choice to the relatives of the victims and their lawyers. We do not have any explanation for the difference in the number of deceased for Norway and Finland since the number of positives is practically the same. Authorities of those countries should investigate this difference further. In Fig. 23, we update the results to June 23, 2020, see also Table 2.

We can repeat the analysis rescaling the data as we did for Italy and some US states in the previous section. This is interesting for Sweden because of the herd immunization adopted strategy. In Fig. 15, we plot the relevant probabilities as function of time. No t_{max} can be defined; thus, we have kept t_{crit} , 2, 4 and 6 weeks shifts. The positive probabilities are rather constant and similar to the TX case, see Fig. 12. Interestingly and fortunately, the deceased probability is decreasing with time, which may indicate that the weaker part of the population is decreasing in number (natural selection) and/or remedies to mitigate the virus are improving.

We can compare the Sweden case to Norway. The positive probabilities quickly decrease while there are no recorded deaths in the last two weeks prior to August 3, 2020, see Fig. 16.

One feature worth noticing from Fig. 14 is the time delay and the slow spread of the Covid-19; this could be due to the extremely cold weather in the winter and early spring for these countries. There is some hope and common belief that the warmer season will help to normalize the situation, as for flu. Other reasons might be put forward, for instance if the virus is somehow adapted to bats, we can naively assume that it will be more deadly for temperatures higher than 10 °C, since below such value most bats hibernate. Temperature difference might explain the spread delay in countries like France, UK, Germany and Russia respect to Italy. Of course, other ingredients must be considered such as people flows from/to infected places, population density, etc. [13, 14]. No matter what the reasons may be for a temperature dependence of the spread, it is clear that some systems perform better if it is not too hot or too cold. We can test these hypotheses using the 50 US states data since they cover a wide range of temperatures in the spring season. In Fig. 17, we display the results obtained using the data on May 3, 2020. Different states values were averaged if their temperatures differ about 1 °C in order to have better statistics. Even though a smooth behavior is not observed, we enforced (optimistic) Gaussian fits, which give 10 °C at the maximum and a similar variance, see the inset in Fig. 17. The large error bars and the discrepancies respect to the fit at higher temperatures refer to touristic places: Florida, Hawaii and Louisiana, particularly popular during spring break. If we take this result at face value, it predicts about 240 people per million inhabitants to be positive to the virus in the summer with 35 °C average temperatures. Even if this value seems small, it is a seed d_0 to restart the pandemic. We can already see this in Fig. 17 from the large increases over the Gaussian fit corresponding to the high population density states and large touristic flows. It might suggest that low temperatures in hospitals may decrease the virus aggressive spreading, keeping in mind that a vaccine is the only definitive solution. Until then we can only aggressively test and isolate positives similarly to the S. Korean approach to the pandemic. If we compare this prediction to the scaled values of the positive probabilities for Italy where the average temperature in the summer is of the order of 30 °C, we get about 0.5% which is a higher value than predicted by the fit.

The preceding discussion and Fig. 17 suggest that the ‘common belief’ that the summer season weakens the virus is (unfortunately) not supported by the data.

Another popular argument widely discussed in the press [23] is the beginning time of the virus spread. The first reported cases to the World Health Organization date December 31, 2019 (hence the name Covid-19 [13]) from the Wuhan region in China. In Ref. [8], the Wuhan case was analyzed and it was concluded that the time it took for the virus to peak is of the order of 15 days, thus around the middle of January 2020. We have seen the devastating effect of the virus on the society with tens of thousands of positives and deceased in short times. Thus to hide the pandemic is impossible. China does not provide the number of daily tests and we cannot derive the typical times entering our model as discussed in Sect. 2. However, we can derive these times for most of the countries/states/regions discussed in the paper. In Fig. 18, we plot t_{\max} & t_{crit} versus $1/\lambda$ (the Lyapunov time) and for completeness the corresponding numerical values are reported in Table 3. The Lyapunov time gives the rate of the virus propagation; the shorter it is the faster the peak probability is reached. Sweden, the country, which adopted herd immunization, displays the largest values of t_{\max} and Lyapunov time thus the longest time duration for the pandemic. The value of the Lyapunov time decreases depending on the efficacy of the quarantine and other factors (population density, etc.). Interestingly enough, two almost parallel lines for t_{crit} can be seen the lowest near 50 days and the other below 100 days; the corresponding countries are reported in Table 3. These values are sensitive to the beginning times of the quarantine in each country. A rather random increase in t_{\max} may be noticed and attributed to early reopening in different countries. This

result supports the first report of the virus spread at the end of 2019 with a possible error of 2 weeks at most [23].

In conclusion, in this paper we have discussed the predictive power of a two-step model based on chaos theory. A comparison among different countries suggests that it would be safe to release the quarantine when the probability for positive is lower than 4%, the maximum value for S. Korea. This implies that, if the quarantine is dismissed, then the same measures, as for the Koreans, should be followed by the other countries: careful testing, backtracking and isolation of positives and quarantine again if needed. Herd immunization or natural selection is very difficult to justify from the data available so far, especially since we are dealing with thousands of human lives no matter the age or other nonsense. No real dependence on seasonal temperatures is observed, maybe with the only exception of very low temperatures, below 10 °C. The model suggests that the beginning time for the pandemic spread is at most as early as the middle of December 2019.

Acknowledgements We thank Professor J. Natowitz for a critical reading of the manuscript. This work was supported by the Fundamental Research Funds for the Central Universities No. GK201903022. The data analysis code for the model proposed in the paper is available from the authors upon request.

Appendix A: Updated data

In this appendix, we update some of the relevant figures discussed in the text to the end of June and August, 2020 (Figs. 19, 20, 21, 22, 23 and Tables 1, 2, 3).

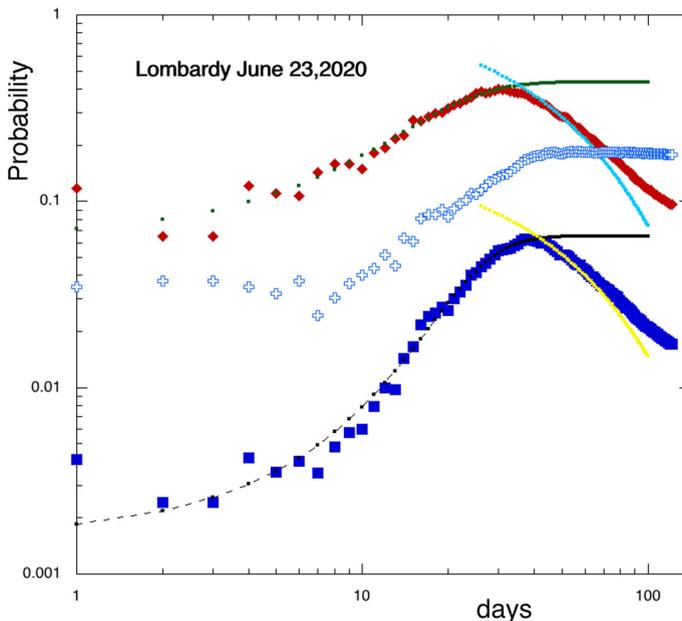


Fig. 19 Same as Fig. 4 for the Lombardy case. The original model predictions are given by the full lines. Notice the data increase respect to the prediction at later times due to the reopening of normal activities: a situation to monitor attentively

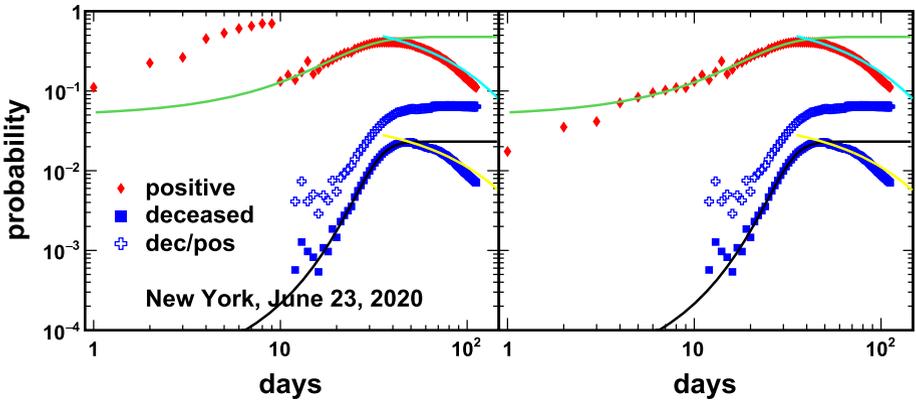


Fig. 20 Same as Fig. 8 updated to June 23, 2020. A small decrease respect to the prediction from Eqs. (1) and (2) is observed at later times

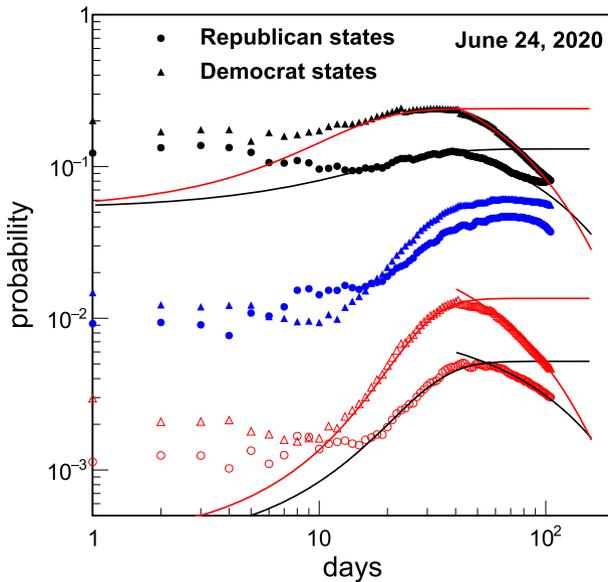


Fig. 21 Positive (full symbols) and deceased (open symbols) probabilities and their ratios (blue symbols) for US states with different party governor. The democratic case is dominated by the high population density state of New York, see Fig. 8. The striking different behavior explains the prediction discussed in Fig. 13 regarding the USA. Notice an increase at later times for the republican states suggesting a too early reopening. Recall that there is a time delay for the deceased respect to positive

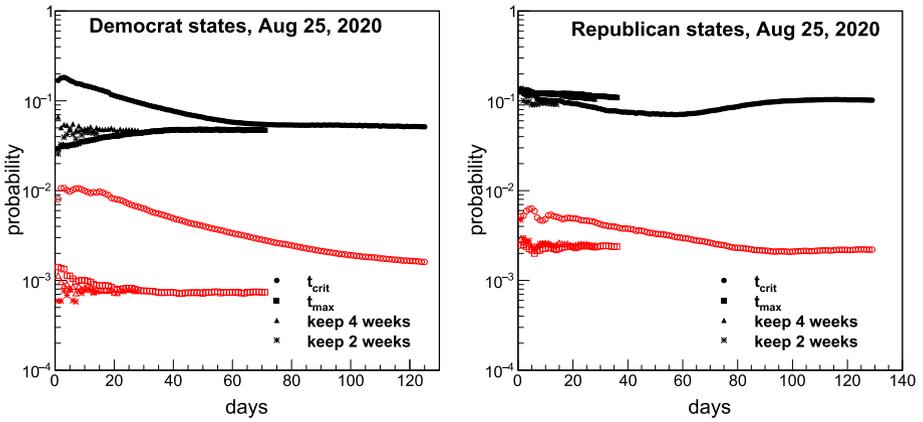


Fig. 22 Same as Fig. 21 but for the rescaled probabilities, compare to Figs. 11 and 12

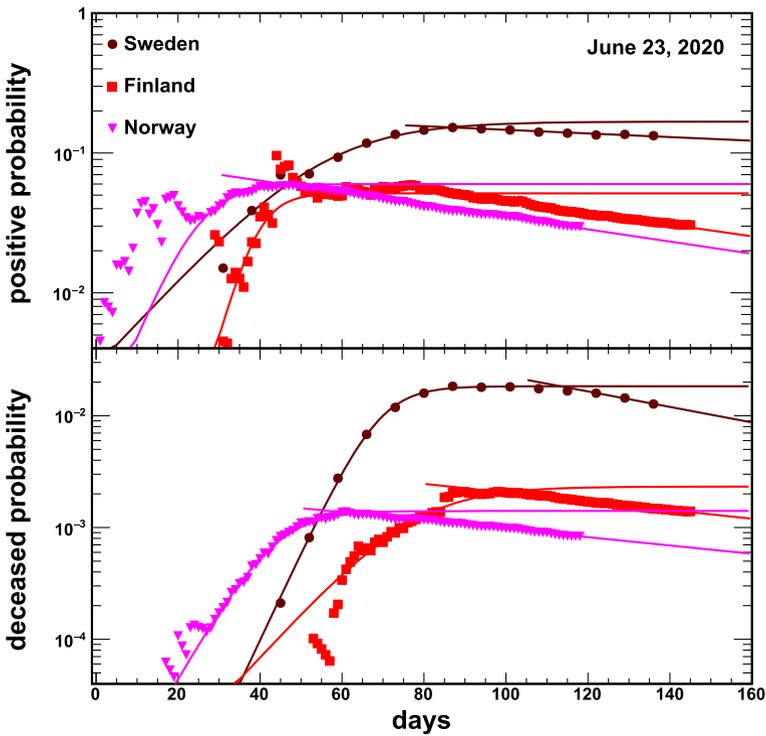


Fig. 23 Same as Fig. 14. Notice the slow decay for Sweden. The data points for Norway have been corrected [<https://ourworldindata.org/grapher/full-list-total-tests-for-covid-19>], compare to Fig. 14. The updated predictions are given in Table 2

Table 1 Model predictions compared to data for different countries corresponding to Fig. 13

Country	Positive	Positive (quarantine)	Data May 6	Data June 28	Deceased	Deceased (quarantine)	Data May6	Data June 28
S. Korea	39,855 ± 5978	10,880 ± 1632	10,806	12,715	479 ± 72	292 ± 44	255	282
Japan	67,305 ± 10,096	23,457 ± 3519	15,354	18,390	4332 ± 650	708 ± 106	543	971
Germany	555,076 ± 83,261	189,479 ± 28,422	164,897	193,499	19,350 ± 2902	10,615 ± 1592	6996	8957
Italy	1,994,297 ± 299,145	250,188 ± 37528	213,013	240,136	188,467 ± 28,270	40,351 ± 6053	29,315	34,716
France	1,055,857 ± 158,379	457,352 ± 68,603	131,292	156,156	209,604 ± 31,441	114,135 ± 17,120	25,491	29,700
New York	1,418,146 ± 212,721	398,913 ± 59,837	323,978	392,539	68,931 ± 10,340	26,361 ± 3954	19,877	24,835
Texas	181,520 ± 27,228	55,075 ± 8261	34,422	148,728	5518 ± 828	2623 ± 393	948	2393
USA	4,490,440 ± 673,566	1,487,103 ± 223,065	1,171,185	2,452,048	260,530 ± 39,080	87,963 ± 13,194	68,081	124,811
UK	3,164,339 ± 474,651	287,538 ± 43,131	194,994	310,254	433,073 ± 64,961	45,713 ± 6857	29,427	43,514

The France discrepancy is discussed in the text

Table 2 Updated results on June 23, 2020 for Sweden, Norway and Finland

Country	Positive-model	Data positive	Deceased-model	Data deceased
Sweden	58,270 ± 8740	65,137	4730 ± 710	5280
Finland	7331 ± 1100	7191	338 ± 51	328
Norway	8518 ± 1278	8815	242 ± 36	249

Notice that the discrepancy between Norway and Finland discussed in the main text was due to a mistake in the data reporting of Norway [<https://ourworldindata.org/grapher/full-list-total-tests-for-covid-19>]

Table 3 Typical times obtained from the model fits to data for different countries corresponding to Fig. 13

Country (region)	$1/\lambda$ (days)	t_{crit} (days)	t_{max} (days)	Population density (km^{-2})
S. Korea	2.4	27.3	124.2	511.6
Japan	8.9	90.2	189.0	334.7
Germany	6.2	78.7	168.6	234.6
Italy	5.4	52.9	103.3	200.6
France	5.8	84.4	152.4	118.3
New York	8.4	41.2	90.8	137.6
Texas	6.2	32.4	148.5	42.4
USA	7.3	90.3	165.1	36.2
UK	7.4	80.5	109.3	279.5
Sweden	13.7	54.9	389.6	24.6
Finland	4.1	86.7	190.4	18.2
Norway	6.0	45.5	145.3	14.8
Florida	4.6	45.5	108.6	129.1
Lombardy	7.2	39.0	76.0	421.7
Spain	4.9	82.9	135.3	92.4

References

1. H.G. Schuster, *Deterministic Chaos* (VCH, Weinheim, 1995)
2. L.D. Landau, E.M. Lifshitz, *Fluid Mechanics* (Elsevier, Amsterdam, 2004)
3. D. Helbing, D. Brockmann, T. Chadefaux et al., *J. Stat. Phys.* **158**, 735 (2015)
4. P. Arena, A. Bonasera, C. Brigante, M. Bucolo, F. Di Grazia, D. Lombardo, F. Sapuppo, M.C. Virzì, in 2008 Proceedings of the Mediterranean Conference on Control and Automation, MED'08 (2008)
5. V. Baran, M. Zus, A. Bonasera, A. Paturca, *Rom. J. Phys.* **60**, 1263 (2015)
6. S. Zhang, A. Bonasera, M. Huang et al., *Phys. Rev. C* **99**, 044605 (2019)
7. A. Bonasera, M. Bruno, C.O. Dorso, P.F. Mastinu, *Rivista Del Nuovo Cimento* **23**(2), 1 (2000)
8. A. Bonasera, S. Zhang, *Front. Phys.* **8**, 171 (2020)
9. A. Bonasera, G. Bonasera, S. Zhang, *Eur. Phys. J. Plus* **135**, 453 (2020)
10. M. Perc et al., *Front. Phys.* **8**, 127 (2020)
11. M.G. Hãncean, M. Perc, J. Lerner, *R. Soc. Open Sci.* **7**, 200780 (2020)
12. B. Povh, K. Rith, C. Scholz, F. Zetsche, *Particles and Nuclei* (Springer, Berlin, 2006)
13. J.J. Van Bavel, K. Baicker, P.S. Boggio et al., *Nat. Hum. Behav.* **4**, 460 (2020)
14. M. Coccia, *Sci. Total Environ.* **729**, 138474 (2020)
15. P-F. Verhulst, *Corresp. Math. Phys.* **10**, 113 (1838)
16. L.R. Zou, F. Ruan, M.X. Huang et al., *N. Engl. J. Med.* **382**, 1177 (2020)
17. C. Carver, N. Jones, Comparative accuracy of oropharyngeal and nasopharyngeal swabs for diagnosis of COVID-19. Oxford COVID-19 evidence service team center for evidence based medicine

- (2020). <https://www.cebm.net/covid-19/comparative-accuracy-of-oropharyngeal-and-nasopharyngeal-swabs-for-diagnosis-of-covid-19/>
18. W.L. Wang, Y.L. Xu, R.Q. Gao et al., JAMA (2020). <https://doi.org/10.1001/jama.2020.3786>
 19. M.J. Loeffelholz, Y.W. Tang, Emerg. Microbes Infect. (2020). <https://doi.org/10.1080/22221751.2020.1745095>
 20. C. Sheridan, Nat. Biotechnol. **38**, 509 (2020)
 21. Z.T. Li, Y.X. Yi, X.M. Luo et al., J. Med. Virol. (2020). <https://doi.org/10.1002/jmv.25727>
 22. <https://www.cnn.com/2020/06/19/politics/nebraska-governor-no-face-masks-requirement/index.html>
 23. <https://www.bbc.com/news/world-us-canada-52975934>