



HAL
open science

Common trends in the epidemic of Covid-19 disease

Milad Radiom, Jean-Francois Berret

► **To cite this version:**

Milad Radiom, Jean-Francois Berret. Common trends in the epidemic of Covid-19 disease. The European Physical Journal Plus, 2020. hal-02554301v2

HAL Id: hal-02554301

<https://hal.science/hal-02554301v2>

Submitted on 29 Apr 2020 (v2), last revised 18 Jun 2020 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Common trends in the epidemic of Covid-19 disease

Milad Radiom, Jean-François Berret

*Matière et Systèmes Complexes, UMR 7057 CNRS Université Denis Diderot Paris-VII, Bâtiment
Condorcet, 10 rue Alice Domon et Léonie Duquet, 75205 Paris, France*

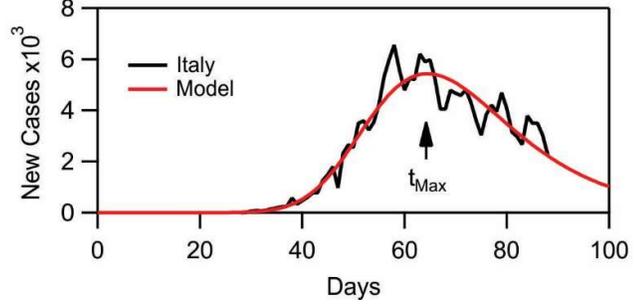
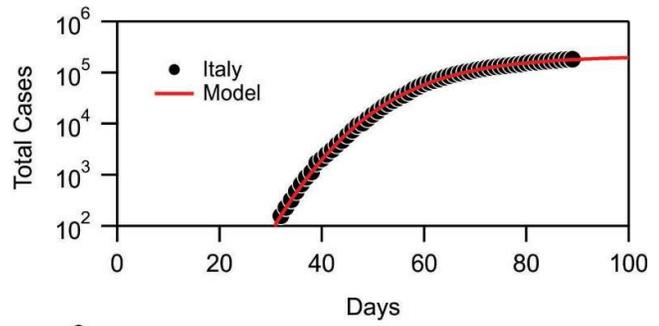
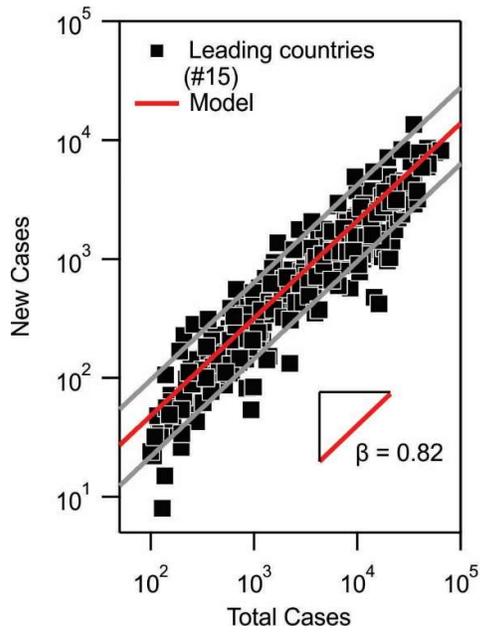
Correspondence: milad.radiom@univ-paris-diderot.fr, jean-francois.berret@univ-paris-diderot.fr

Abstract. The discovery of SARS-CoV-2, the responsible virus for the Covid-19 epidemic, has sparked a global health concern with many countries affected. Developing models that can interpret the epidemic and give common trend parameters are useful for prediction purposes by other countries that are at an earlier phase of the epidemic; it is also useful for future planning against viral respiratory diseases. One model is developed to interpret the fast-growth phase of the epidemic and another model for the full interpretation during the fast-, slow- and plateau phases. Both models are shown to agree reasonably with the data. It is shown that during the fast phase, the number of new infected cases in several countries depends on the total number of cases by a power-law relation with a scaling exponent equal to 0.82. During this phase, the duplication time is 1 – 3 days. The duplication time together with another parameter that controls the phase change from the fast phase to the slow and plateau phases, varying in the range of 0.1 – 0.5, may be used for data interpretation and for guiding predictions regarding this disease, *e.g.* the onset of the maximum in the number of new cases.

Keywords: Covid-19; Respiratory diseases; Epidemic; Modeling; Duplication Time

Abbreviation: SARS = Severe Acute Respiratory Syndrome; CoV = Corona Virus

Graphical Abstract:



1. Introduction

An outbreak of an infectious respiratory disease emerged in the city of Wuhan, in the Chinese province of Hubei, in December 2019. One month later, the discovery of the responsible virus, SARS-CoV-2, was announced [1]. The respiratory disease, which is termed Covid-19, is an ongoing global health concern with many countries affected so far. While a great momentum in the scientific community is to find therapeutics and vaccines against Covid-19 [2], it seems beneficial to develop *simple* models to understand the trends of the epidemic [3, 4]. Previous attempts have been made to this end. New models are generally adopted from the previous models of avian and swine flu epidemics [5, 6], and may be used to make predictions regarding the Covid-19 epidemic, *e.g.* to simulate spreading to neighboring regions from hot spots, and to evaluate the effectiveness of the mitigating measures [7-9]. Available models have generally estimated a human-to-human transmissibility of the disease, which is called the basic reproduction number, R_o . Estimations of R_o give values larger than 2 denoting a highly contagious virus [5, 10, 11]. Simpler models exist and have been used for example to compare between different SARS epidemics [12, 13].

This letter seeks to interpret the available data using a simple model making use of physically meaningful parameters. In particular, with respect to the several periods for an epidemic, the presented analysis focuses on the first period, before the so-called herd immunity period. In the herd immunity and in the absence of a vaccine, the virus stops expanding when 60% of a population is contaminated [14-16]. A Californian study has shown that the number of contaminated people in this region was 50 times more than the total number of reported cases [17]. Applying this factor to France's numbers, with a current total cases of about 1.8×10^5 , this factor gives about 7.5 million contaminated cases. This number of cases is equivalent to 12% of the French population. Recently, another study stated that this number was about 5.7% on

average and about 12% in the region Ile-de-France [18]. Thereby, the available data and the model presented in this work is for the first period of the epidemic before the herd immunity is acquired.

The focus of this letter is initially placed on the data that is available for China and South Korea because at the time of writing this letter these countries have *apparently* reached the end of the first period of their epidemic. The data from these countries together with a few others, such as Iran, Turkey, Italy, Spain, France, and the United States are then used to develop common trend parameters using a simple physical model. The developed models in this work may be used for prediction purposes by other countries or with respect to future respiratory viral infections.

2. Results and discussion

The data of temporal variation of total number of cases is gathered from an online source [19] from 22-02-2020 to 20-04-2020 and reproduced in Figs. 1a and 1b for China and South Korea, respectively. It is expected that the trends of the other countries before the herd immunity period will be similar. In particular for China and South Korea, in the first period of the epidemic, three phases may be assigned to the temporal variation of the total number of cases data, whereby an initial fast-growing phase is followed by a slow-growing phase and then a plateau. These phases are respectively marked f (fast), s (slow) and p (plateau) on Figs. 1a and 1b. These phases may be discerned from the apparent variations in the growth exponent of the total number of cases versus time. For example, in the f phase of China, the apparent growth exponent $(1/n)(dn/dt)$, where n is the total number of cases and t the time, is about 0.096, while it is 0.006 and < 0.001 respectively in the s phase and the p phase. For South Korea, the apparent growth exponents are 0.137, 0.006 and 0.001 respectively during the f phase, the s phase, and the p phase. Since the data are presented as a function of the number of days starting

from 22-02-2020, a note should be made for the data-days prior to this date. This data is not available; however, it may be assumed that the temporal variation of the total number of cases prior to 22-02-2020 follows the behavior of the f phase. The onset of the s phase and the p phase is about two to four weeks for both China and South Korea.

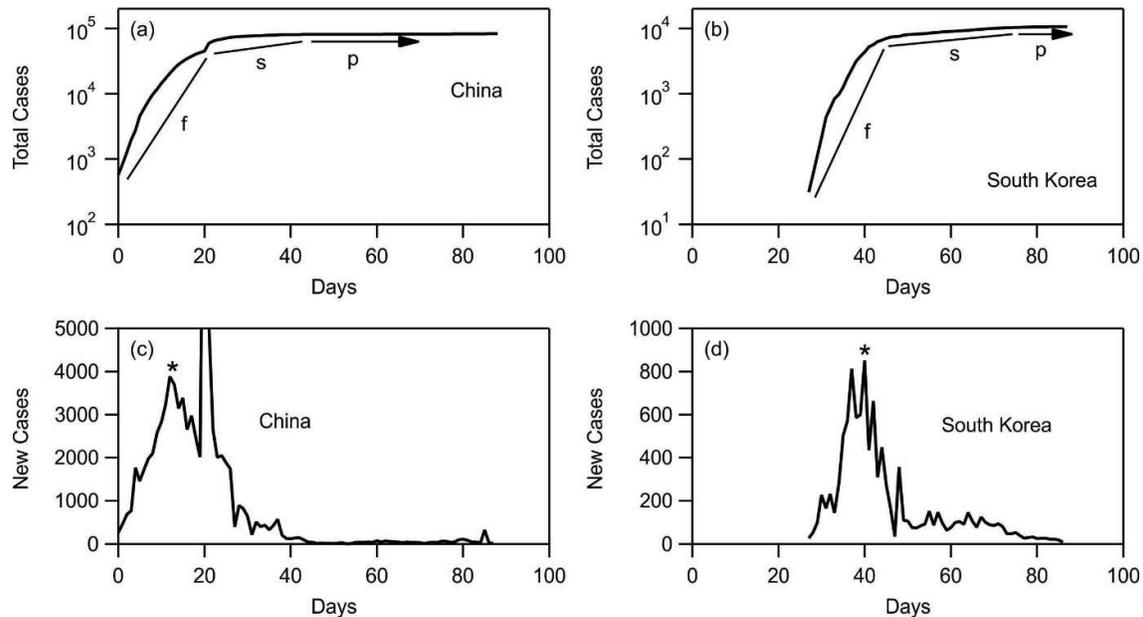


Figure 1: (a, b) Total number of cases as a function of the number of days starting from 22-01-2020 for China (a) and South Korea (b). The fast-growing phase (f), slow-growing phase (s) and plateau phase (p) are marked. These phases may be discerned from the apparent variations in the growth exponent. (c, d) Number of new cases as a function of the number of days starting from 22-01-2020 for China (c) and South Korea (d). Peak in the number of new cases is marked with asterisks. All data are collected until 20-04-2020.

The number of new cases is calculated from the previous data and presented in Figs. 1c and 1d respectively for China and South Korea. In particular, these data show peaks which are marked with asterisks on Figs. 1c and 1d. The onset of the peak corresponds approximatively to the start of the s phase discussed earlier [14].

The models presented hereafter in this work address the first period of the epidemic (before the herd immunity), when typically the proportion of the infected population is inferior to 10% [18]. One common trend is obtained by plotting the number of new cases as a function of the

total number of cases in the f phase. Here we include the f phase data from 15 different countries with leading total number of infected cases [19, 20]. As shown in Fig. 2, this data presentation leads to the expression:

$$dn = \lambda n^\beta, \quad (1)$$

where dn is the number of new cases, λ the prefactor and β a scaling exponent. It is found that Eq. 1 can reasonably interpret the data with an exponent $\beta = 0.82$. The prefactor λ is found to vary from 0.5 to 2.2 for all the investigated countries, including China, South Korea, Iran, Turkey, Italy, Spain, France, the United States, Germany, Switzerland, India, the United Kingdom, Canada, Brazil and Singapore. The high value of the exponent β signifies the wide spread of the virus during the f phase. It also shows that the number of new cases is strongly dependent on the total number of cases during the f phase. However, as $\beta < 1$, the spread during this phase is sub-exponential.

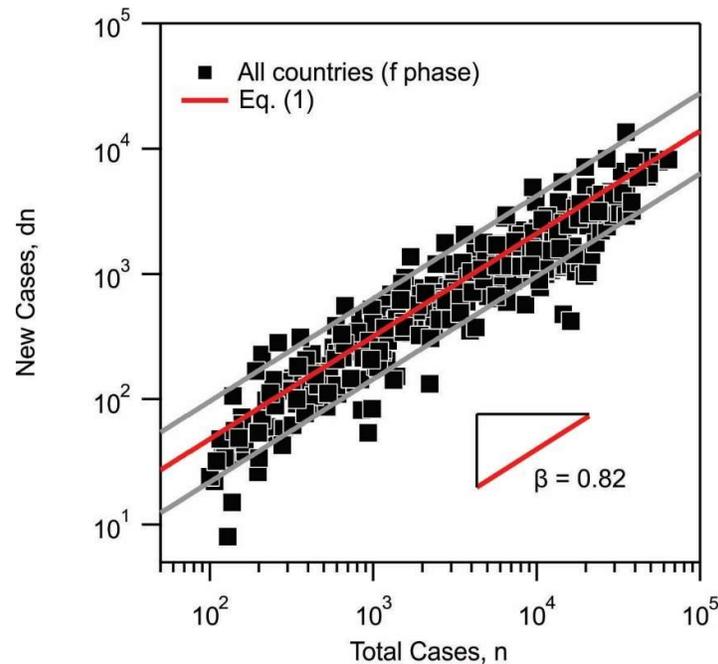


Figure 2: Number of new cases as a function of the total number of cases showing the power-law dependence of Eq. 1. The exponent $\beta = 0.82$ is found to interpret the data reasonably. The

prefactor $\lambda = 0.5$ gives the lower bound and $\lambda = 2.2$ the upper bound. Data from all countries (15 countries) are included [19, 20].

Table 1 gives predictions for the number of new cases from the total number of cases during the f phase. The values are calculated for the exponent $\beta = 0.82$, and the prefactor $\lambda = 0.5$ (resp. $\lambda = 2.2$) to obtain a minimum (resp. maximum) value. The results show that by controlling the total number of cases at the start of the epidemic, significant reduction in the number of new cases can be obtained. It further shows that 1 to 2 orders of magnitude increase in the total number of cases in the f phase, results in about 7 – 44 times higher number of new cases.

Table 1: Predictions for the number of new cases from the total number of cases during the f phase. Calculations are based on Eq. 1, for $\beta = 0.82$, , and $\lambda = 0.5$ to obtain a minimum and $\lambda = 2.2$ to obtain a maximum.

Total cases, n	Minimum new cases, dn_{\min}	Maximum new cases, dn_{\max}
100	22	96
1000	140	630
10000	950	4190
50000	3570	15690

Next, a common trend for the temporal evolution of the total number of cases is sought. The following expression is used to interpret the latter data [21]:

$$n = 2^{t/\tau} \left[\frac{1}{n_o^\alpha} + \frac{1}{n_\infty^\alpha} (2^{\alpha t/\tau} - 1) \right]^{-1/\alpha}, \quad (2)$$

where n_o and n_∞ correspond to the total number of cases at the start and the end of epidemic respectively (*i.e.* the boundaries), τ is similar to the duplication time, and α is a parameter that controls the temporal variation of the curve. For example, a high value of α corresponds to a fast transition from the f phase to the s phase and the p phase. The form of Eq. 2 is inspired from the growth of eukaryotic cells in Petri dishes, where the boundary conditions slow the

initial exponential growth [21]. For cell culture, it was found that $\alpha = 2$ gave good results. Here, α has been left as an adjustable parameter controlling the different measures taken by the states to mitigate the virus expansion, including social distancing and barrier gestures. Although the effects of social confinement are different from Petri dish boundary conditions, it is shown below that Eq. 2 can reasonably interpret the data. It is noted that other functions may be used for data interpretation [11, 14, 22]; the preference in the use of Eq. 2 is its simplicity and physically meaningful parameters. It is noted however that cell division is fundamentally different from disease spreading in that, in the simplest form of argument, one cell divides to two related daughter cells while one individual can infect more than one related or non-related individuals. The applicability of Eq. 2 is in its form and should be not taken as a parallel between cell growth process and disease spreading process.

However, it is not possible to fit all available data to Eq. 2 as many countries are in their f phase and a value of n_{∞} cannot be assigned without overestimation or even underestimation. Data from selected countries including China, South Korea, Iran, Turkey, Italy, Spain, France, and the United States were then fitted to Eq. 2. These fits for China, South Korea, Iran, and Turkey are shown in Fig. 3. It is noted that fitting started from $n \geq 100$, and the value of n_0 was fixed. Thereby, the number of the fitting parameters was reduced to 3. Moreover, wherever the duplication time was obtained to be less than 1 day, its value was fixed to be equal to 1 day. In these cases, the number of the fitting parameters was reduced to 2. From Fig. 3, it is evident that Eq. 2 provides a reasonable agreement with the data, especially considering occasional non-uniformities in the trend, *e.g.* in the case of Iran's data. The best fit parameters are presented in Table 2. One observes an almost similar duplication time among these countries but different α parameters. The differences in the latter parameter show different time periods to change from the f phase to the s phase, which is *e.g.* two time longer for Iran as compared with South Korea.

Furthermore, the number of new cases is calculated from the total number of cases as discussed previously. Thereby, Fig. 3 further shows a comparison between the latter data and calculations from the derivative of Eq. 2 using the same fit parameters. The reasonable agreement obtained from the latter comparison further shows that Eq. 2 is adequate for data interpretation. Similarly, fits to the total of number of cases for Italy, Spain, France, and the United States, together with the reproduction of the number of new cases from data and from derivative of Eq. 2 are shown in Fig. 4. The agreement is again reasonable.

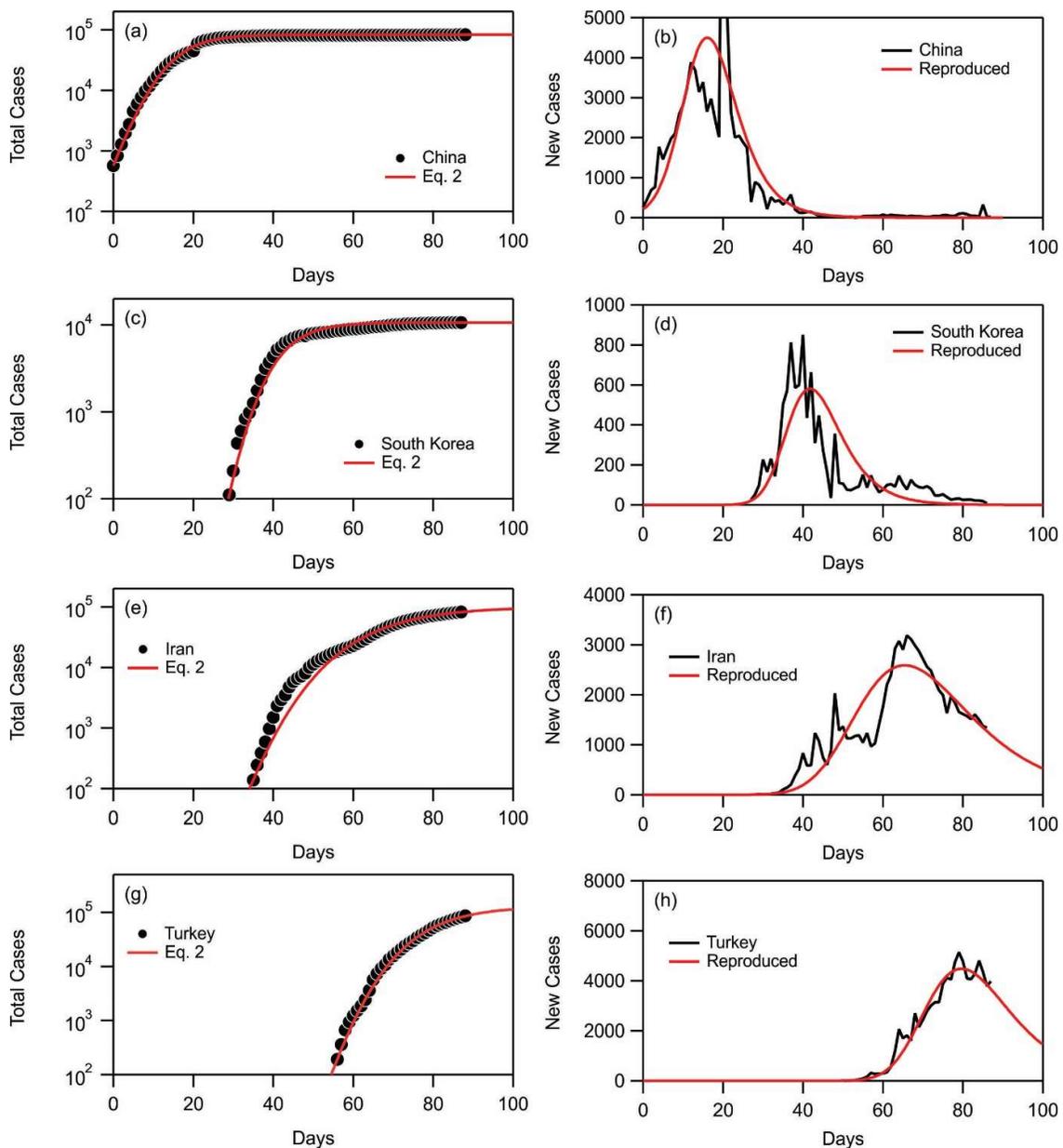


Figure 3: (a, c, e, g) Total number of cases as a function of the number of days starting from 22-01-2020 together with fits to Eq. 2. (a) China, (c) South Korea, (e) Iran, and (g) Turkey. The fit parameters are given in Table 2. (b, d, f, h) Number of new cases as a function of the number of days starting from 22-01-2020 together with reproduced curves from the derivative of Eq. 2 using the fitted values given in Table 2. (b) China, (d) South Korea, (f) Iran, and (h) Turkey. All data are collected until 20-04-2020.

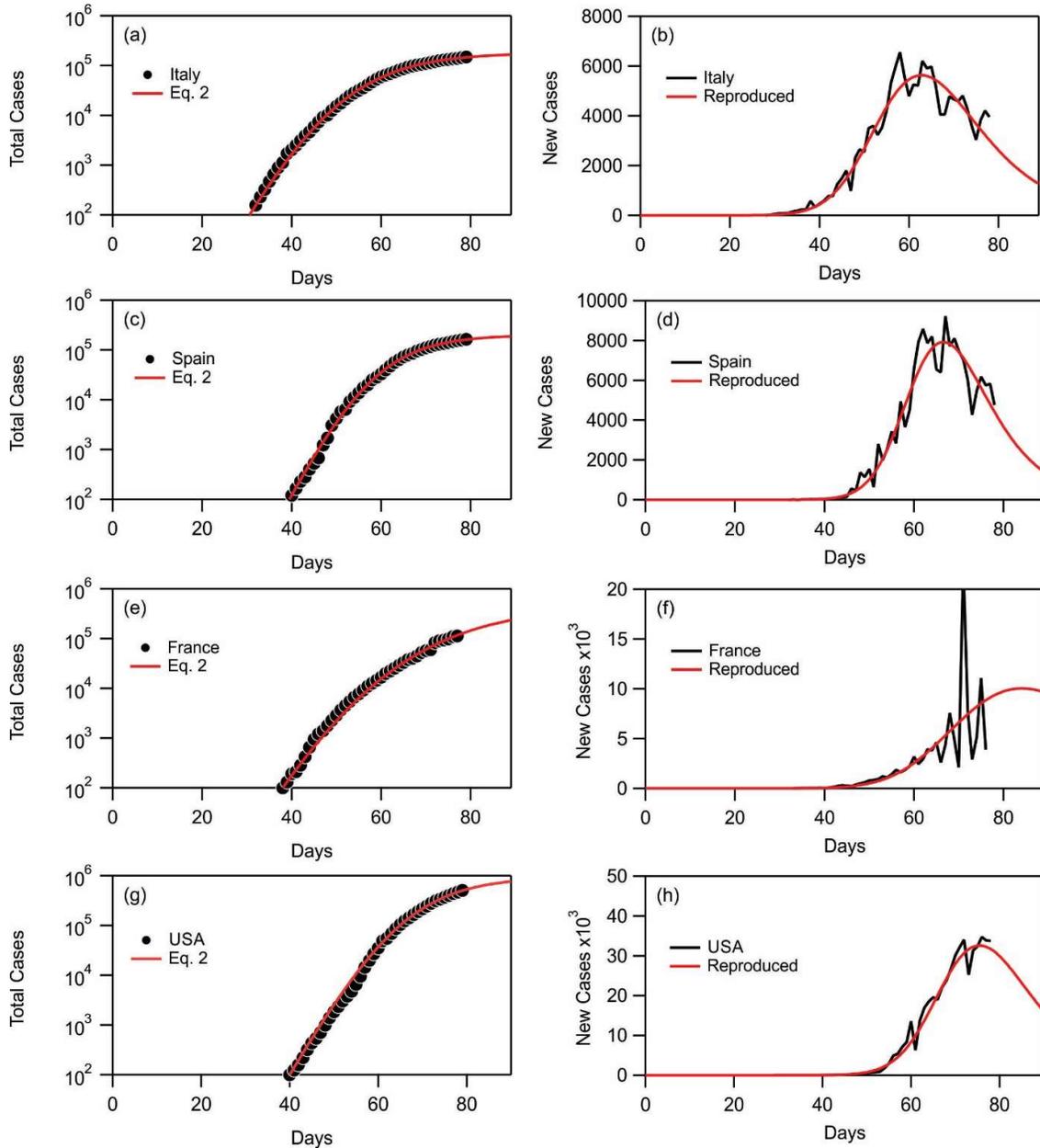


Figure 4: (a, c, e, g) Total number of cases as a function of the number of days starting from 22-01-2020 together with fits to Eq. 2. (a) Italy, (c) Spain, (e) France, and (g) USA. The fit parameters are given in Table 2. (b, d, f, h) Number of new cases as a function of the number of days starting from 22-01-2020 together with reproduced curves from the derivative of Eq. 2

using the fitted values given in Table 2. (b) Italy, (d) Spain, (f) France, and (h) USA. All data are collected until 20-04-2020.

The resulting fit parameters are summarized in Table 2. One finds that the duplication time τ varies from 1 – 3 days. The significance of this range is that during the f phase, every 1 – 3 days the total number of cases is almost duplicated. This observation coincides with the strong dependence of the number of new cases on the total number of cases in the f phase discussed earlier. A high value of α parameter means that a turnover from the f phase to the s phase and the p phase occurs over a shorter time period as is the case for China (Table 2 & Fig. 1). For Spain, which has about the same duplication time as China, a lower value of the α parameter means that the turnover to the s and p phases is longer. It is noted that the fitted n_∞ values for Iran, Turkey, Italy, Spain, France and the United States may be used as predictions since these countries have not yet reached their p phases.

Table 2: Best fit parameters obtained from fits of Eq. 2 to total number of cases data of several countries. Highlighted columns denote estimations.

Country	n_o^* (number)	n_∞ (number)	τ (days)	α	t_{Max}^{**} (days)
China	571	$(0.83 \pm 0.10) \times 10^5$	1.6 ± 0.1	0.40 ± 0.04	16
South Korea	111	$(0.11 \pm 0.01) \times 10^5$	1.0^*	0.24 ± 0.01	13
Iran	139	$(1.00 \pm 0.02) \times 10^5$	1.0^*	0.11 ± 0.01	30
Turkey	191	$(1.29 \pm 0.03) \times 10^5$	1.0^*	0.15 ± 0.01	24
Italy	157	$(2.09 \pm 0.03) \times 10^5$	1.1 ± 0.1	0.12 ± 0.01	32
Spain	120	$(2.24 \pm 0.03) \times 10^5$	1.3 ± 0.1	0.18 ± 0.02	28
France	100	$(1.81 \pm 0.06) \times 10^5$	2.8 ± 0.1	0.53 ± 0.06	35
USA	100	$(10.78 \pm 0.15) \times 10^5$	1.7 ± 0.1	0.23 ± 0.01	37

* Fixed parameter in the fit. ** Counts from the start of the epidemic in each state, and not from 22-01-2020.

The form of Eq. 2 may then be used to give general trends of the epidemic using the fitted values in Table 2. To this end, the number of cases at the two extremities were selected to be $n_o = 10^2$ and $n_\infty = 10^6$, the duplication time was varied in the range $\tau = 1-3$ days and $\alpha = 0.1-0.5$. In the Supporting Information Fig. SI 1, several predictions are plotted using these values. It is shown that for a fix value of α parameter, a higher value of the duplication time τ broadens the curves whereby the turnover period to the s phase and the p phase is also prolonged. Similarly, the onset of the maximum in the number of new cases is delayed. It is otherwise found that the α parameter has a reverse effect. For example, at a fix value of the duplication time τ , a higher value of the α parameter shifts the curves to the left whereby the s phase and the p phase occur earlier, and similarly the onset of the maximum in the number of new cases is shortened. Thereby the α parameter gives approximately the time range between the initial growth phases and the final p phase. This variation may be abrupt when α is large (*e.g.* this is the case in the cell model) or may be smooth when α is small. These curves suggest that both α and τ control the trend behavior from n_o to n_∞ . These two parameters can then be used for prediction purposes and planning regarding the trends in the other countries.

Furthermore, from the estimated values of τ , α and n_∞ in Table 2, it is possible to estimate the duration from the start of the epidemic to the onset of the maximum in the number of new cases. This duration which is denoted t_{Max} is expressed as follows:

$$t_{Max} = \frac{\tau}{\alpha \ln 2} \ln \left\{ \frac{1}{\alpha} \left[\left(\frac{n_\infty}{n_o} \right)^\alpha - 1 \right] \right\}. \quad (3)$$

The form of Eq. 3 shows that this time period is proportional to the duplication time τ ; however, it is inversely related to the α parameter. This relation agrees with our previous statement on the general trend behaviors. Furthermore, Eq. 3 shows that the dependence on the

ratio of the total number of cases to the initial number of cases, n_{∞}/n_o is weak. The estimated t_{Max} values are provided in Table 2. It is clear that China and South Korea have reached to this maximum over a shorter period of time as compared to the other countries. However, in general, 2 – 5 weeks is the period over which these countries have reached the maximum of the number of their new cases. The onset of this maximum and the number of new cases at this maximum, which is generally compared with the medical capacity, are important parameters for planning and placing mitigating measures by the states.

3. Conclusions

It is concluded by the analysis presented in this letter that the first period of the Covid-19 epidemic, before the herd immunity, may be interpreted in terms of simple models containing physically meaningful parameters. The temporal evolution of the total number of cases may be divided to three phases. In the f phase, it is shown that the number of new cases depends on the total number of cases in a power-law relation with an exponent $\beta = 0.82$. The overall temporal evolution of the total number of cases, *i.e.* all three phases, may also be interpreted by modeling. It is found that the duplication time τ is 1 – 3 days and the α parameter which governs the phase change from the f phase to the s and p phases to vary from 0.1 – 0.5 for the selected countries. The higher is the value of this parameter, the shorter is the period of phase change. The duplication time together with the α parameter may be used to interpret the total number of case as well as the number of new cases. The models presented in this work, and the average values which are obtained by fitting to several countries with leading number of cases, may be used for prediction purposes as well as for guiding the implementation of proper controlling measures during each phase of the epidemic. The presented model and analysis is for when the total number of infected cases is less than 10% of the entire population. In order

to interpret the trends when up to 60% of the population are infected, a different model is required, which is out of the scope of the present work.

References

1. www.who.int/westernpacific/emergencies/covid-19.
2. Kupferschmidt, K. and J. Cohen, *Race to find COVID-19 treatments accelerates*. Science, 2020. **367**(6485): p. 1412.
3. Enserink, M. and K. Kupferschmidt, *With COVID-19, modeling takes on life and death importance*. Science, 2020. **367**(6485): p. 1414.
4. Epstein, J.M., *Modelling to contain pandemics*. Nature, 2009. **460**(7256): p. 687-687.
5. Colizza, V., et al., *Modeling the worldwide spread of pandemic influenza: Baseline case and containment interventions*. Plos Medicine, 2007. **4**(1): p. 95-110.
6. Coburn, B.J., B.G. Wagner, and S. Blower, *Modeling influenza epidemics and pandemics: insights into the future of swine flu (H1N1)*. BMC Medicine, 2009. **7**.
7. Ivorra, B., et al., *Mathematical modeling of the spread of the coronavirus disease 2019 (COVID-19) taking into account the undetected infections. The case of China*. 2020.
8. Huang, R., M. Liu, and Y. Ding, *Spatial-temporal distribution of COVID-19 in China and its prediction: A data-driven modeling analysis*. Journal of Infection in Developing Countries, 2020. **14**(3): p. 246-253.
9. Karako, K., et al., *Analysis of COVID-19 infection spread in Japan based on stochastic transition model*. Bioscience trends, 2020.
10. Zhao, S., et al., *Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak*. International Journal of Infectious Diseases, 2020. **92**: p. 214-217.
11. Anastassopoulou, C., et al., *Data-based analysis, modelling and forecasting of the COVID-19 outbreak*. PLOS ONE, 2020. **15**(3): p. e0230405.
12. Jia, L., et al., *Prediction and analysis of Coronavirus Disease 2019*. arXiv:2003.05447v2, 2020.
13. Ciufolini, I. and A. Paolozzi, *Mathematical prediction of the time evolution of the COVID-19 pandemic in Italy by a Gauss error function and Monte Carlo simulations*. European physical journal plus, 2020. **135**(4): p. 355-355.
14. Kermack, W.O., A.G. McKendrick, and G.T. Walker, *A contribution to the mathematical theory of epidemics*. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 1927. **115**(772): p. 700-721.
15. Li, Q., et al., *Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia*. New England Journal of Medicine, 2020. **382**(13): p. 1199-1207.
16. Riou, J. and C.L. Althaus, *Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020*. Euro Surveill, 2020. **25**(4).
17. Bendavid, E., et al., *COVID-19 Antibody Seroprevalence in Santa Clara County, California*. medRxiv, 2020: p. 2020.04.14.20062463.
18. Salje, H., et al., *Estimating the burden of SARS-CoV-2 in France*. 2020.
19. www.worldometers.info/coronavirus/.
20. *Counties include: China, South Korea, Iran, Turkey, Italy, Spain, France, the United States, Germany, Switzerland, India, UK, Canada, Brazil and Singapore*.
21. Safi, M., et al., *Interactions between sub-10-nm iron and cerium oxide nanoparticles and 3T3 fibroblasts: the role of the coating and aggregation state*. Nanotechnology, 2010. **21**(14): p. 145103.
22. Liang, K., *Mathematical model of infection kinetics and its analysis for COVID-19, SARS and MERS*. Infection, Genetics and Evolution, 2020. **82**: p. 104306.