

## Research



**Cite this article:** Fokas AS, Dikaios N, Kastis GA. 2021 Covid-19: predictive mathematical formulae for the number of deaths during lockdown and possible scenarios for the post-lockdown period. *Proc. R. Soc. A* **477**: 20200745.

<https://doi.org/10.1098/rspa.2020.0745>

Received: 25 September 2020

Accepted: 23 April 2021

**Subject Areas:**

mathematical modelling

**Keywords:**

mathematical modelling of epidemics, Covid-19, Riccati equation, integrable systems, inverse problems

**Author for correspondence:**

Athanasios S. Fokas

e-mail: [tf227@cam.ac.uk](mailto:tf227@cam.ac.uk)

# Covid-19: predictive mathematical formulae for the number of deaths during lockdown and possible scenarios for the post-lockdown period

Athanasios S. Fokas<sup>1,2,3</sup>, Nikolaos Dikaios<sup>2,4</sup> and George A. Kastis<sup>2</sup>

<sup>1</sup>Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, UK

<sup>2</sup>Mathematics Research Center, Academy of Athens, Athens, Greece

<sup>3</sup>Department of Biomedical Engineering, University of Southern California, Los Angeles, CA, USA

<sup>4</sup>Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, UK

ASF, 0000-0002-5881-802X; ND, 0000-0001-9865-0260; GAK, 0000-0002-1283-0883

In a recent article, we introduced two novel mathematical expressions and a deep learning algorithm for characterizing the dynamics of the number of reported infected cases with SARS-CoV-2. Here, we show that such formulae can also be used for determining the time evolution of the associated number of deaths: for the epidemics in Spain, Germany, Italy and the UK, the parameters defining these formulae were computed using data up to 1 May 2020, a period of lockdown for these countries; then, the predictions of the formulae were compared with the data for the following 122 days, namely until 1 September. These comparisons, in addition to demonstrating the remarkable predictive capacity of our simple formulae, also show that for a rather long time the easing of the lockdown measures did not affect the number of deaths. The importance of these results regarding predictions of the number of Covid-19 deaths during the post-lockdown period is discussed.

## 1. Introduction

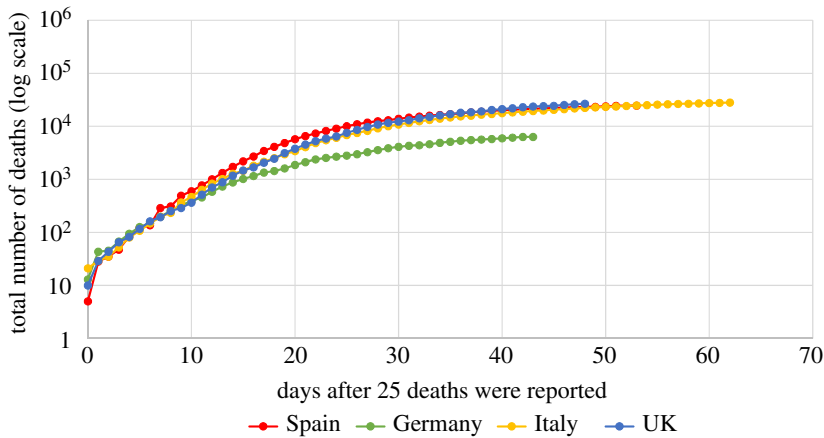
Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the third coronavirus to appear in the human population in the past two decades; the first was the severe acute respiratory syndrome coronavirus (SARS-CoV) that caused the 2002 outbreak; the second was the Middle East syndrome coronavirus (MERS-CoV) responsible for the 2012 outbreak. SARS-CoV-2 has now caused a pandemic, which poses the most serious global public health threat since the devastating 1918 H1N1 influenza pandemic that killed approximately 50 million people (in proportion to today's population, this corresponds to 200 million people). The unprecedented mobilization of the scientific community has already led to remarkable progress towards combating this threat, such as understanding significant features of the virus at the molecular level; see for example [1,2]. In addition, international efforts have intensified towards the development of specific pharmacological interventions, including clinical trials using old or relatively new medications and the employment of specific monoclonal antibodies, as well as novel approaches for the production of an effective vaccine. In particular, the US Food and Drug Administration has granted a conditional approval to the anti-viral medication remdesivir.<sup>1</sup> Unfortunately, the combination of the anti-viral medications lopinavir and ritonavir that are effective against the human immunodeficiency virus has not shown any benefits [3]; similarly, for the combination of the anti-malarial medication hydroxychloroquine and the antibiotic azithromycin [4]. There are ongoing clinical trials testing the synthetic protein tocilizumab that binds interleukin-6 (this medication is often used in rheumatoid arthritis), as well as clinical trials involving the infusion of infected patients with plasma from individuals who have recovered from coronavirus disease (Covid-19) [5].

The scientific community is also playing an important role in advising policy-makers about possible non-pharmacological approaches to limit the catastrophic impact of the pandemic. For example, following the rigorous analysis in [6] of two possible strategies, called mitigation and suppression, for combating the epidemic, the UK switched from mitigation to suppression. Within this context, in order to design a long-term strategy, it is necessary to be able to predict important features of the Covid-19 epidemics, such as the final accumulative number of deaths. Clearly, this requires the development of predictive mathematical models.

In a recent paper [7], we presented a model for the dynamics of the accumulative number of individuals in a given country that are reported at time  $t$  to be infected by SARS-CoV-2. This model is based on a particular ordinary differential equation of the Riccati type, which is specified by a *constant parameter* denoting the final total number of infected individuals and by a *time-dependent function*. Remarkably, although the above Riccati equation is a nonlinear equation that is conditional on time-dependent coefficients, it can be solved in closed form. Its solution depends on the above parameter and function, as well as on a parameter related to the constant arising in integrating this equation. In the particular case that the associated time-dependent function is a constant, the explicit solution of the above Riccati equation becomes the classical *logistic formula*. It was shown in [7] that, although this formula provides an adequate fit for the given data, it does *not* yield sufficiently accurate predictions. In order to provide more accurate predictions, two novel formulae were introduced in [7], called *rational* and *birational*. The predictive capabilities of these formulae were established, first by comparing their predictions with real data and second by showing that a complicated deep learning algorithm did not perform better than these explicit formulae.

Here, we will show that similar expressions can also be used for determining the time evolution of the number,  $N(t)$ , of deaths in a given country caused by Covid-19. The Riccati equation formulated in [7] is now specified by the parameter  $N_f$ , denoting the final total number of deaths, and by the function  $\alpha(t)$ ; this function and  $N_f$  model the effect of the basic characteristics of Covid-19 as well as the cumulative effect of the variety of different measures taken by the given country for the prevention and the treatment of this viral infection. In the case that  $\alpha(t)$

<sup>1</sup><https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-issues-emergency-use-authorization-potential-covid-19-treatment>.



**Figure 1.** The total cumulative number of deaths in Spain, Germany, Italy and the UK, due to Covid-19, reported up to 1 May 2020, as a function of days after the day that 25 deaths were reported. (Online version in colour.)

is a constant, denoted by  $k$ , the solution of the above Riccati equation becomes the well-known logistic formula,

$$N(t) = \frac{N_f}{1 + \beta e^{-kt}}, \quad (1.1)$$

where  $\beta$  is the constant of integration. In the case that  $\alpha(t)$  is the rational function  $kd/(1 + dt)$ , the Riccati equation yields the rational formula; this expression is obtained from the logistic formula by simply replacing the exponential function with the function  $(1 + dt)^{-k}$ , where  $d$  and  $k$  are constants. The birational formula is similar to the rational formula, but the associated parameters are different before and after a fixed time  $X$ ; this constant parameter is either  $T$  or in the neighbourhood of  $T$ , where  $T$  denotes the time that the rate of deaths reaches a maximum. The point on the curve depicting  $N$  as a function of  $t$  corresponding to  $t = T$  is known as the *inflection point*.

It turns out that, in general, the birational formula yields better predictions than the rational, which in turn provides better predictions than the logistic. However, the birational formula can be employed only after the curve approaches its sigmoidal part (a precise criterion of when the birational can be used is discussed in §4). For this reason, since we only used data until 1 May 2020, at which time the epidemic in Germany had not yet reached the sigmoidal part of the curve, the birational formula could not be used for the epidemic in Germany.

Figure 1 depicts the total number of deaths as a function of time after the day that 25 deaths had occurred, for the epidemics in Spain, Germany, Italy and the UK. The reason for analysing the epidemics after the time that 25 deaths are reached is that the data are quite noisy before this stage (this is analogous to the fact that usually models for the number of individuals reported infected with SARS-CoV-2 consider data after the number of reported infected reaches 500).

In this work, we implemented the following tasks. (i) We determined the parameters specifying the logistic, rational and birational formulae (for Germany only the first two formulae were used) using the above data, namely until 1 May 2020. In this way, it became evident that each of these models *can* fit the data. (ii) We compared the predictions of these models for the following 122 days, namely until 1 September 2020. These comparisons show that the predictions for Spain, Italy and Germany via our novel formulae are truly outstanding, and for the UK quite good. (iii) We computed the time of the plateau as well as the value of  $N$  at the plateau using the logistic and our two formulae (for Germany only the logistic and the rational formulae were used), where the *plateau* is defined as the point when the rate of deaths is 5% of the maximum rate.

By implementing the above tasks, we found the following estimates for the dates that the plateau is reached as well as for the number of deaths when this occurs: Spain plateaus on 25

May 2020 with 27 090 deaths; Germany plateaus on 16 June 2020 with 8702 deaths; Italy plateaus on 19 June 2020 with 34 399 deaths; and the UK plateaus on 25 June 2020 with 38 487 deaths. As will be discussed below, these estimates are very accurate for the characteristics of the actual plateaux in the above four countries.

## 2. Results

Table 1 presents the model parameters and the plateau characteristics for the three different formulae, for Spain, Germany, Italy and the UK (for the epidemic in Germany, only the parameters and characteristics of the logistic and rational formulae are presented). The confidence intervals for each fitting parameter are also presented. The calculation of the inflection point requires the time that the derivative of  $N$  becomes maximum. To compute this point, we used the logistic formula (the other formulae yield similar results). For Spain, Germany, Italy and the UK the inflection point occurred at  $t=28$ ,  $t=27$ ,  $t=35$  and  $t=32$ , respectively. This corresponds to 6 April, 15 April, 4 April and 15 April 2020, respectively.

Figures 2–5 present the actual versus predicted cumulative number of deaths due to Covid-19 as a function of days after 25 deaths were reported, for Spain, Germany, Italy and the UK, respectively. The three formulae were trained with data up to 1 May 2020, which corresponds to  $T+27$ ,  $T+25$ ,  $T+16$  and  $T+16$ , respectively. Regarding the birational formula, we used  $X=T$ . The predictions of the three formulae were compared for the following 122 days, namely until 1 September 2020. It is evident from these curves that, whereas each of the explicit logistic, rational and birational formulae *can* fit the data quite well, the predictive capacities of these formulae are *quite different*. The predictions of our two novel formulae for the 122 days following 1 May are remarkably accurate for all four counties analysed here. Regarding Spain, it is interesting to note that early on it appeared that our predictions were slightly off, but after the real data were corrected at the official site of the European Centre for Disease Prevention and Control, the curves depicting our predictions are essentially indistinguishable from the curve depicting the real data. Similarly, for the UK, our predictions differed substantially from the dashed curve (online version in red) in figure 5, which depicts the *earlier* data, but, again, after these data were officially corrected it became clear that our predictions are quite accurate.

For the epidemic in Spain (figure 2), the logistic formula predicts a plateau on 7 May 2020 (day 59 after the day that 25 deaths were reported) with 23 795 deaths; the rational and birational formulae predict a plateau on 25 May 2020 (day 77) with 26 581 and 27 090 deaths, respectively. From the actual reported data until 1 September 2020, we were able to determine that the actual plateau occurred on approximately 27 May 2020 (day 79) with approximately 27 940 deaths (owing to the repeated corrections imposed by the official agency reporting the number of Covid-19 deaths, it is not possible to identify the above numbers exactly). That is, only 2 days and 850 deaths over the estimate of the birational formula.

For the epidemic in Germany (figure 3), the logistic formula predicts a plateau on 17 May 2020 (59 days after the day that 25 deaths were reported) with 6830 deaths; the rational formula predicts a plateau on 16 June 2020 (day 89) with 8702 deaths. From the actual reported data until 1 September 2020, we were able to determine that the actual plateau for Germany occurred on 20 June 2020 (day 93) with 8882 total deaths. That is, only 4 days and 180 deaths over the estimate of the rational formula.

For the epidemic in Italy (figure 4), the logistic formula predicts a plateau on 12 May 2020 (73 days after the day that 25 deaths were reported) with 27 390 deaths; the rational formula predicts a plateau on 9 June 2020 (day 101) with 31 976 deaths; and the birational formula predicts a plateau on 19 June 2020 (day 111) with 34 399 deaths. From the actual reported data until 1 September 2020, we were able to determine that the actual plateau for Italy for the number of deaths due to Covid-19 occurred on 22 June 2020 (day 114) with 34 634 total deaths. That is, only 3 days and 235 deaths over the estimate of the birational formula.

For the epidemic in the UK (figure 5), the logistic formula predicts a plateau on 15 May 2020 (62 days after the day that 25 deaths were reported) with 27 433 deaths; the rational formula predicts a

**Table 1.** Model parameters and plateau characteristics for the logistic, rational and birational formulae for the epidemics in Italy, Spain and the UK. For the epidemic in Germany, only the parameters and characteristics of the logistic and rational formulae are presented. The confidence intervals for each fitting parameter are also presented.

	Italy	Spain	UK	Germany
<b>logistic model</b>				
$N_f$	27 687 (26 979, 28 395)	24 016 (23 410, 24 622)	27 708 (27 129, 28 287)	6906 (6680, 7124)
k	0.1188 (0.1020, 0.1258)	0.1481 (0.1384, 0.1578)	0.1533 (0.1488, 0.1578)	0.143 (0.1357, 0.1503)
$\beta$	63.4044 (50.0735, 76.7353)	65.7474 (49.2186, 82.2762)	134.8143 (129.2112, 140.4174)	51.2444 (43.7599, 58.7289)
T	35	28	32	27
plateau (days)	73	59	62	59
plateau (deaths)	27 390	23 765	27 433	6830
$R^2$	0.9955	0.9955	0.9975	0.9979
RMSE	678	607	475	102
<b>rational model</b>				
$N_f$	33 153 (31 764, 34 542)	27 524 (26 532, 28 516)	34 920 (34 267, 35 573)	9021 (8197, 9845)
k	3.6299 (3.1668, 4.0930)	3.8035 (3.2815, 4.3255)	4.0954 (3.8357, 4.3551)	3.7379 (2.9106, 4.5652)
$\beta$	4892.02 (4317.43, 5466.60)	6372.60 (5824.45, 6920.80)	23 496.07 (22 976.40, 24 016.0)	451.2287 (315.4700, 586.9874)
d	0.2455 (0.2120, 0.2790)	0.2996 (0.2630, 0.3362)	0.2997 (0.2546, 0.3448)	0.12828 (0.1073, 0.1493)
plateau (days)	101	77	88	89
plateau (deaths)	31 976	26 581	33 886	8702
$R^2$	0.9994	0.9994	0.9996	0.9994
RMSE	262	233	199	55
<b>birational model</b>				
k	6.7118 (5.4890, 7.9346)	5.0583 (3.5627, 6.5539)	4.4465 (3.265, 5.628)	
b	953.44 (617.10, 1289.80)	6532.56 (4187.6, 8877.5)	39 666 (36 475, 42 857)	
c	23 310 (20 820, 25 800)	21 434 (19 329, 23 539)	29 724 (28 914, 30 534)	
d	0.0572 (0.0085, 0.1059)	0.1814 (0.0837, 0.2791)	0.3000 (0.2662, 0.3338)	
$k_1$	4.0093 (3.6935, 4.3251)	4.6349 (4.1008, 5.1690)	3.5263 (1.7963, 5.2563)	

(Continued.)

**Table 1.** (Continued.)

	Italy	Spain	UK	Germany
$b_1$	478.1700 (385.2837, 571.0563)	97 887.64 (7034, 12 541)	27 864.22 (20 965.32, 34 763.12)	
$c_1$	31 737 (30 303, 33 171)	21 766 (20 367, 23 165)	38 492 (31 449, 45 535)	
$d_1$	0.0893 (0.0779, 0.1007)	0.1796 (0.1338, 0.2254)	0.4354 (0.1081, 0.7627)	
plateau (days)	111	77	103	
plateau (deaths)	34 399	27 090	38 487	
$R^2$	1.0000	0.9999	0.9998	
RMSE	60	82	148	

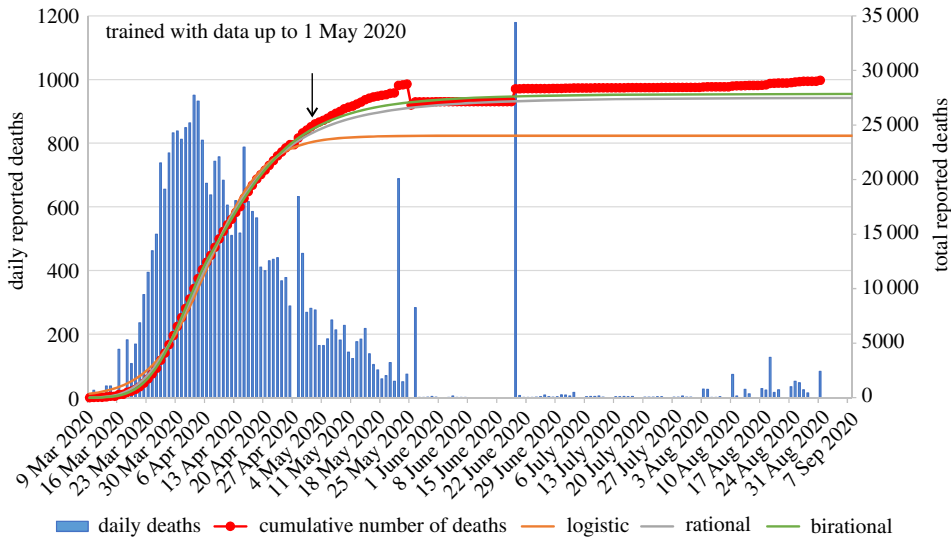
plateau on 10 June 2020 (day 88) with 33 886 deaths; and the birational formula predicts a plateau on 25 June 2020 (day 103) with 38 487 deaths. From the actual reported data until 1 September 2020 (corrected on 18 August), we were able to determine that the actual plateau for the UK for the number of deaths due to Covid-19 occurred on 28 June 2020. That is, only 3 days and 1802 deaths over the estimate of the birational formula.

### 3. Discussion

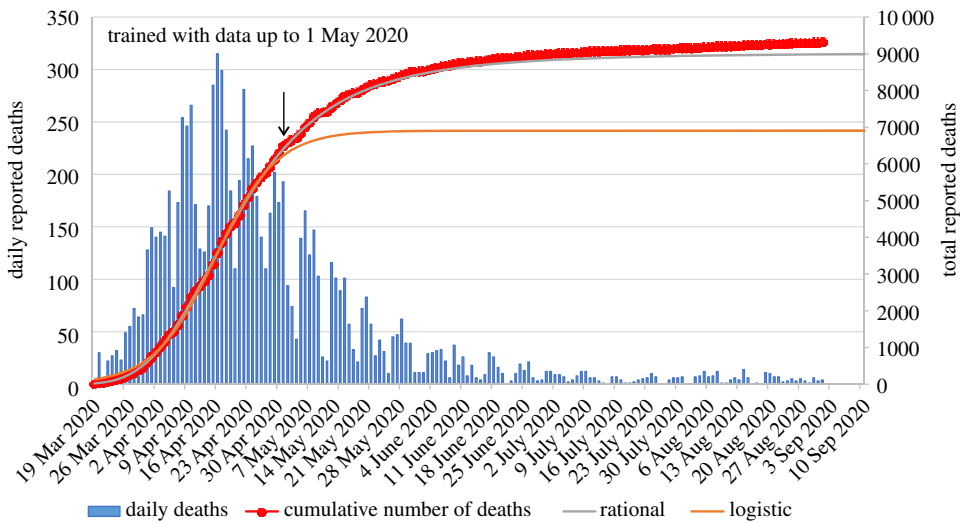
Several useful models elucidating aspects of the SARS-CoV-2 pandemic have already appeared in the literature, including [8–17]. Also, Roosa *et al.* [18,19] investigated phenomenological models to perform real-time and short-term predictions of the reported cases for the Covid-19 pandemic in China. Most of these models depend only on the time variable  $t$ , but in [20], where a useful analogy is presented between the spreading of an infection and the process of chemical reaction, the dependence of a space variable of the spread of infection is also taken into consideration. In [21], a novel approach is presented for reducing the cost of testing without compromising accuracy: combining subsamples and testing them in groups. The above references represent only a tiny fraction of more than 3000 publications that have appeared in the last four months in arXiv, medRxiv and bioRxiv.

Here, following [7], we modelled the cumulative number  $N$  of deaths caused by Covid-19 in a given country as a function of time, in terms of a Riccati equation. This equation is specified by the constant  $N_f$  and the function  $\alpha(t)$ . Although this Riccati equation is a nonlinear ordinary differential equation containing time-dependent coefficients, it can be solved in closed form. For appropriately chosen functions  $\alpha(t)$ , the solution provides a flexible generalization of the classical logistic formula given by equation (1.1) that has been employed in a great variety of applications, including the modelling of infectious processes.

The fact that  $\alpha$  is now a function of  $t$  has important implications. In particular, it made it possible to construct the rational and birational formulae. These two formulae, as well as the logistic formula, provide good fits for the available data. However, as discussed in detail above, the rational and birational formulae provide much more *accurate predictions*. As shown in figure 2, the success of our explicit formulae for three out of the four counties analysed here is truly remarkable. The fact that for the UK our predictions are not as good as in the other three countries suggests that our model may not be as accurate for countries which delayed the implementation of the lockdown measures. The overall unexpected success of our formulae suggests that the few constant parameters specifying the rational and birational formulae capture the essence of the epidemiological characteristics associated with the deaths caused by Covid-19. This is

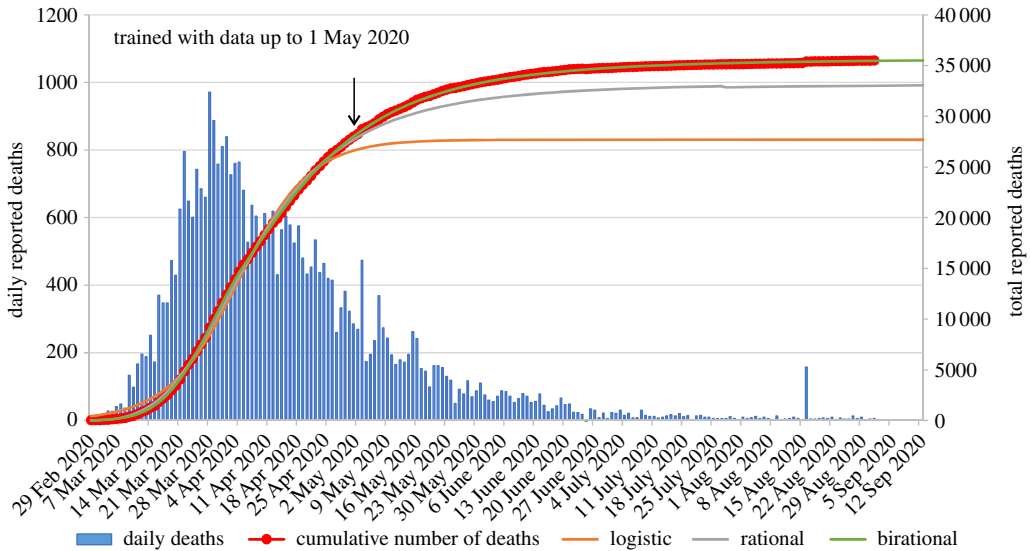


**Figure 2.** Comparison of the predictions of our formulae for the number of reported deaths for Spain using data only until 1 May 2020, versus the actual data until 1 September 2020 (predictions for 122 days). The data points with circle markers depict the cumulative number of deaths. The arrow indicates the last data point used for calibrating the formulae. Negative daily reported deaths are not displayed for clarity. The inflection point for Spain occurred on 6 April 2020. (Online version in colour.)

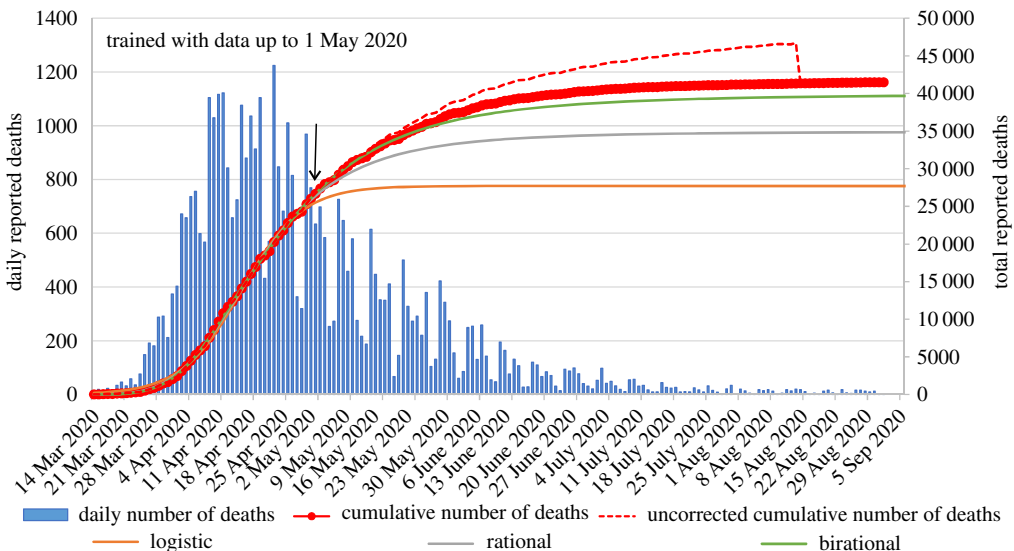


**Figure 3.** Comparison of the predictions of our formulae for the number of reported deaths for Germany using data only until 1 May 2020, versus the actual data until 1 September 2020 (predictions for 122 days). The data points with circle markers depict the cumulative number of deaths. The arrow indicates the last data point used for calibrating the formulae. Negative daily reported deaths are not displayed for clarity. The inflection point for Germany occurred on 15 April 2020. (Online version in colour.)

reminiscent of the well-known Anderson–May model, where a single parameter characterizing a constant coefficient Riccati equation captures the effect of the combination of several different biological mechanisms [22]. An additional explanation of the above success is that our approach is ‘data driven’. Indeed, the given data are used to fix the unknown parameters occurring in our mathematical formulae, and then the resulting algorithms (just as happens with the algorithms in artificial intelligence) can be used for predictions.



**Figure 4.** Comparison of the predictions of our formulae for the number of reported deaths for Italy using data only until 1 May 2020, versus the actual data until 1 September 2020 (predictions for 122 days). The data points with circle markers depict the cumulative number of deaths. The arrow indicates the last data point used for calibrating the formulae. Negative daily reported deaths are not displayed for clarity. The inflection point for Italy occurred on 4 April 2020. (Online version in colour.)



**Figure 5.** Comparison of the predictions of our formulae for the number of reported deaths for the UK using data only until 1 May 2020, versus the actual data until 1 September 2020 (predictions for 122 days). The data points with circle markers depict the cumulative number of deaths. The arrow indicates the last data point used for calibrating the formulae. Negative daily reported deaths are not displayed for clarity. The dashed curve depicts the earlier data before the correction made on 14 August 2020. The inflection point for the UK occurred on 15 April 2020. (Online version in colour.)

It is noted that our approach has the capacity for increasing continuously the accuracy of the predictions: as soon as the epidemic in a given country passes the time  $T$ , the rational model can be used; furthermore, when the sigmoidal part of the curve is approached, the rational formula can be supplemented with the birational formula. It should be noted that all three models examined



in this work, namely the logistic, rational and birational formulae, do not perform well if they are calibrated with data before the epidemic peak (maximum rate or deaths, inflection point T).

Following this expected decline of the ‘first wave’ of infections after the lockdown, several European countries began easing their lockdown measures. In particular, in Spain, the initial easing of the Covid-19 measures occurred on 13 April, when workers in some non-essential sectors, such as construction and industry, were allowed to return to work; on 28 April, the government announced further plans for easing lockdown restrictions [23]. In Germany, initial easing of the Covid-19 measures occurred on 20 April, when smaller shops (shops with a retail space of up to 800 m<sup>2</sup>), as well as bookshops, bike stores and car dealerships, were allowed to reopen to the public [24,25]; schools as well as hair salons opened on 4 May; by 9 May, all statewide curfews had been lifted. Italy started easing the lockdown on 4 May 2020, when millions of Italians went back to work (manufacturing industries and construction sites reopened after seven weeks of restrictive measures); swimming pools, sports centres and gyms reopened on 25 May; theatres and cinemas reopened on 15 June [26]. In the UK, the government eased the lockdown measures on 10 May, by changing the slogan from ‘Stay at Home’ to ‘Stay Alert’ [27]; people could exercise more than once daily outdoors and could interact with others while maintaining social distancing. However, the situation in the UK worsened, leading the Prime Minister to postpone some lockdown easing measures scheduled to begin in England on 1 August [28].

Taking into consideration the devastating impact of Covid-19 on the economy, as well as its implications for the psychological health of the population [29], the decision for easing the lockdown measures was considered necessary. A rigorous and computationally effective approach for calculating the number of deaths following the easing of the lockdown measures is presented in [30] (the employment of the Riccati equation studied here was an important ingredient of the algorithm presented in [30]). The analysis of [30] is based on the *assumption* that the characteristics of the virus remain the same in the post-lockdown period. Under this assumption, it is shown that the numbers both of the infected reported individuals and of deaths will begin to grow in the post-lockdown period. Actually, the number of reported infected cases began growing in Spain in the second week of July, in Germany at the end of July, in Italy in the first week of August and in the UK in the second week of July. Thus, the number of reported infected cases began to grow only approximately two to three months after these countries lifted the lockdown measures. Furthermore, the increase in the number of infected individuals was *not* accompanied by an analogous increase in the number of deaths: in Italy, there is no deviation between the curve depicting the number of deaths and the curve of our predictions, whereas in Spain and Germany a small deviation began to occur in the first week of August; in the UK, where our predictions are not as accurate as in the other three countries, there is *no* increase in the deviation between data and predictions since the second week of June. The long delay in the increase in the number of reported infected individuals as well the relatively small number of deaths suggest that the virus responsible for the current epidemics in these countries may have mutated to one which is less virulent. Indeed, it is shown in [31] that the virus mutated to a variant which has the G614 form of the spike protein instead of the D614 form that was originally identified from the first human cases in Wuhan, China. However, the question of whether this form is less virulent remains open, despite a report in *The New York Times* reproduced in many newspapers (see for example [32]), where it is erroneously stated that in [31] it is actually proven that the mutated virus is less virulent. An alternative explanation for the observation that the increase in infected individuals was not immediately followed by an increase in deaths is that the early infections occurred in ‘younger’ persons; taking into consideration that the SARS-CoV-2 infection in ‘younger’ individuals rarely leads to death, and that an unknown time period is required for ‘younger’ individuals to infect ‘older’ once it follows that there should be a time delay before an increase in the number of reported infected individuals and an increase in the number of deaths. In this connection, it is noted that by extending the algorithm developed in [30] into two sub-populations consisting of ‘younger’ (below 40) and ‘older’ (above 40), it was shown

in [33] regarding the epidemic in Greece that, if the number of contacts between asymptomatic ‘younger’ persons increases, the number of reported infected cases increases but not the number of deaths; on the other hand, the increase in contacts involving ‘older’ persons leads to a dramatic increase in the number of deaths. Interestingly, the increase in the number of deaths in Spain and Germany began to occur after approximately four weeks from the time that the number of reported infected cases began to grow. Independently of whether the virus has mutated to a less virulent form or not, a most beneficial factor in the fight against the pandemic is the adherence to protective measures of the vast majority of older and vulnerable individuals. The next couple of weeks following submission of our manuscript are of crucial importance for the epidemics in Spain and Germany, and hence, by analogy, for the epidemics in other European countries: if the deviations observed in figures 2 and 3 become pronounced, then there may be a second substantial wave of deaths, which will be consistent with the predictions of [30]; on the other hand, if the deviation remains small, it will imply that the epidemics will remain under control. It is natural to hypothesize that the different tempo in the four countries examined here simply reflects the fact that the lifting of lockdown measures was implemented first in Spain and then was followed by Germany, Italy and the UK.

## 4. Material and methods

We obtained the time-series data for Covid-19 for Spain, Germany, Italy and the UK from the official site of the European Centre for Disease Prevention and Control.<sup>2</sup> We arranged the data in the form of deaths,  $N$ , over time measured in days, after the day that the number of deaths reached 25.

Throughout this paper, the unknown parameters were determined by employing the *simplex algorithm*. This is based on an iterative procedure that does not need information regarding the derivative of the function under consideration. The simplex algorithm is particularly effective for cases where the gradient of the likelihood functions is not easy to calculate. The algorithm creates a ‘random’ simplex of  $n + 1$  points, where  $n$  is the number of model parameters that need to be estimated. The constrained variation of the simplex algorithm [34,35] available in MATLAB<sup>®</sup> was used for all models; the  $L_1$ -norm was employed in the likelihood function to improve robustness [36]. Random parameter initializations were used to avoid local minima. The simplex algorithm was chosen because it performed better than certain nonlinear least-squares curve-fitting algorithms evaluated in this work, namely the Levenberg–Marquardt [37] and trust-region-reflective algorithms [38].

The stability of the fitting procedure was established by using the following simple criterion: different fitting attempts based on the use of a fixed number of data points must yield curves which have the same form beyond these fixed points. In this way, it was established that the rational formula could be employed provided that data were available until around the time  $T$ , whereas the birational formula could be used only for data available well beyond  $T$ .

The fitting accuracy of each formula was evaluated by fitting the associated formula on *all* the available data until 1 May 2020. The relevant parameters specifying the logistic, rational and birational formulae are given in table 1. Confidence intervals for each fitting parameter were calculated based on [39,40].

We assumed that the function  $N(t)$  satisfies the Riccati ordinary differential equation

$$\frac{dN}{dt} = \alpha(t) \left( N - \frac{N^2}{N_f} \right). \quad (4.1)$$

The general solution of this equation is given by [7]

$$N = \frac{N_f}{1 + \beta e^{-\tau}}, \quad \tau = \int \alpha(t) dt, \quad (4.2)$$

<sup>2</sup><https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases>.

where  $\beta$  is the constant arising from the integrating equation (4.1). In the particular case that  $\alpha(t) = k$ , equation (4.2) becomes the logistic formula of (1.1). If  $\alpha(t)$  is given by the rational function

$$\alpha(t) = \frac{kd}{1 + kt},$$

then equation (4.2) becomes the rational formula

$$N = \frac{N_f}{1 + \beta(1 + dt)^{-k}}. \quad (4.3)$$

The birational formula is given by

$$N = \begin{cases} \frac{c}{1 + b(1 + dt)^{-k}}, & t \leq X \\ \frac{c}{1 + b(1 + dX)^{-k}} - \frac{c_1}{1 + b_1(1 + d_1X)^{-k_1}} + \frac{c_1}{1 + b_1(1 + d_1t)^{-k_1}}, & t > X, \end{cases} \quad (4.4)$$

where the fixed parameter  $X$  is either  $T$  or in the neighbourhood of  $T$ . For the birational model

$$\alpha(t) = \begin{cases} \frac{kd}{1 + dt}, & t \leq X \\ \frac{k_1 d_1}{1 + d_1 t} \frac{1}{1 + (1 - (c_1/N_f))(1 + d_1 t)^{-k_1}}, & t > X. \end{cases} \quad (4.5)$$

Letting in equation (4.4)  $t \rightarrow \infty$  we find

$$N_f = \frac{c}{1 + b(1 + dX)^{-k}} - \frac{c_1}{1 + b_1(1 + d_1X)^{-k_1}} + c_1. \quad (4.6)$$

If  $b, c, d, k$  are close to  $b_1, c_1, d_1, k_1$ , then  $N_f$  is close to  $c_1$ , and hence the value of  $\alpha(t)$  after  $t = X$  is close to the value of  $\alpha(t)$  before  $t = X$ .

The constant  $T$  can be computed by solving the equation obtained by equating to zero the second derivative of  $N$ . This implies that for the logistic and the rational models  $T$  is given, respectively, by

$$T = \frac{\ln(\beta)}{k}, \quad (1 + dT)^k = b \left( \frac{k-1}{k+1} \right). \quad (4.7)$$

Similarly, for the birational model where the parameters  $b, d$  and  $k$  in the right-hand equation of equations (4.7) are replaced by  $b_1, d_1$  and  $k_1$ , respectively.

Our work uses only the logistic, rational and birational formulae, namely equations (1.1), (4.3) and (4.4) with  $X = T$ , respectively.

**Data accessibility.** The time-series data for Covid-19 were obtained from the official site of the European Centre for Disease Prevention and Control.

**Authors' contributions.** A.S.F. conceived, designed and supervised the study. N.D. and G.A.K. performed the fitting experiments and analysis of the results. A.S.F. wrote the initial draft of the manuscript. All authors critically revised and improved the manuscript in various ways. All authors reviewed the manuscript and gave final approval for publication. All authors agree to be held accountable for the work performed in this paper.

**Competing interests.** We declare we have no competing interests.

**Funding.** A.S.F. acknowledges support from EPSRC, UK (RG79707), in the form of a Senior Fellowship. N.D. acknowledges support from the Royal Society (grant no. INF/R1/191030) in the form of a fellowship.

**Acknowledgements.** A.S.F. is grateful to several colleagues and in particular to the neurologist M. Dalakas and the mathematician S. Kourouklis for encouraging him to be involved with the modelling of SARS-CoV-2.

## References

1. Zhou P *et al.* 2020 A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273. (doi:10.1038/s41586-020-2012-7)
2. Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, Geng Q, Auerbach A, Li F. 2020 Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**, 221–224. (doi:10.1038/s41586-020-2179-y)

3. Cao B *et al.* 2020 A trial of lopinavir–ritonavir in adults hospitalized with severe Covid-19. *New Engl. J. Med.* **382**, 1787–1799. (doi:10.1056/NEJMoa2001282)
4. Magagnoli J, Narendran S, Pereira F, Cummings TH, Hardin JW, Sutton SS, Ambati J. 2020 Outcomes of hydroxychloroquine usage in United States veterans hospitalized with Covid-19. *Med* **1**, 114–127.e3. (doi:10.1016/j.medj.2020.06.001)
5. Bhimraj A *et al.* 2020 Infectious Diseases Society of America guidelines on the treatment and management of patients with Covid-19 infection. See <https://www.idsociety.org/practice-guideline/covid-19-guideline-treatment-and-management> (accessed 20 September 2020).
6. Ferguson NM *et al.* 2020 *Impact of nonpharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand*. London, UK: Imperial College Covid-19 Response Team. See <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-9-impact-of-npis-on-covid-19>.
7. Fokas AS, Dikaivos N, Kastis GA. 2020 Predictive mathematical models for the number of individuals infected with Covid-19. *J. R. Soc. Interface* **17**, 20200494. (doi:10.1098/rsif.2020.0494)
8. Bastos SB, Caljueiro DO. 2020 Modeling and forecasting the early evolution of the Covid-19 pandemic in Brazil. *Sci. Rep.* **10**, 19457. (doi:10.1038/s41598-020-76257-1)
9. Qi C, Karlsson D, Sallmen K, Wyss R. 2020 Model studies on the Covid-19 pandemic in Sweden. (<https://arxiv.org/abs/2004.01575>, q-bio.PE)
10. Gjini E. 2020 Modeling Covid-19 dynamics for real-time estimates and projections: an application to Albanian data. *medRxiv* (doi:10.1101/2020.03.20.20038141)
11. Kaxiras E, Neofotistos G, Angelaki E. 2020 The first 100 days: modeling the evolution of the Covid-19 pandemic. *Chaos Solitons Fract.* **138**, 110114. (doi:10.1016/j.chaos.2020.110114)
12. Yang C, Wang J. 2020 A mathematical model for the novel coronavirus epidemic in Wuhan, China. *Math. Biosci. Eng.* **17**, 2708–2724. (doi:10.3934/mbe.2020148)
13. Tam K-M, Walker N, Moreno J. 2020 Projected development of Covid-19 in Louisiana. (<https://arxiv.org/abs/2004.02859>).
14. Arenas A, Cota W, Gomez-Gardenes J, Gómez S, Granell C, Matamalas JT, Soriano-Panos D, Steinegger B. 2020 A mathematical model for the spatiotemporal epidemic spreading of Covid-19. *Phys. Rev. X* **10**, 041055. (doi:10.1103/PhysRevX.10.041055)
15. Danon L, Brooks-Pollock E, Bailey M, Keeling M. 2020 A spatial model of Covid-19 transmission in England and Wales: early spread and peak timing. *medRxiv*. (doi:10.1101/2020.02.12.20022566)
16. Russo L, Anastassopoulou C, Tsakris A, Bifulco GN, Campana EF, Toraldo G, Siettos C. 2020 Tracing day-zero and forecasting the COVID-19 outbreak in Lombardy, Italy: a compartmental modelling and numerical optimization approach. *PLoS ONE* **15**, e0240649. (doi:10.1371/journal.pone.0240649)
17. Anastassopoulou C, Russo L, Tsakris A, Siettos C. 2020 Data-based analysis, modelling and forecasting of the Covid-19 outbreak. *PLoS ONE* **15**, e0230405. (doi:10.1371/journal.pone.0230405)
18. Roosa K, Lee Y, Luo R, Kirpich A, Rothenberg R, Hyman JM, Yan P, Chowell G. 2020 Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. *Infect. Dis. Model* **5**, 256–263. (doi:10.1016/j.idm.2020.02.002)
19. Roosa K, Lee Y, Luo R, Kirpich A, Rothenberg R, Hyman JM, Yan P, Chowell G. 2020 Short-term forecasts of the COVID-19 epidemic in Guangdong and Zhejiang, China: February 13–23, 2020. *J. Clin. Med.* **9**, 596. (doi:10.3390/jcm9020596)
20. Ramaswamy H, Oberai AA, Yotsos YC. 2021 A comprehensive spatial-temporal infection model. *Chem. Eng. Sci.* **233**, 116347. (doi:10.1016/j.ces.2020.116347)
21. Mutesa L *et al.* 2021 A pooled testing strategy for identifying SARS-CoV-2 at low prevalence. *Nature* **589**, 276–280. (doi:10.1038/s41586-020-2885-5)
22. Anderson RM, May RM. 1981 The population dynamics of microparasites and their invertebrate hosts. *Phil. Trans. R. Soc. Lond. B* **291**, 451–524. (doi:10.1098/rstb.1981.0005)
23. Wikipedia. 2020 COVID-19 pandemic in Spain. See [https://en.wikipedia.org/w/index.php?title=COVID-19\\_pandemic\\_in\\_Spain&oldid=970146896](https://en.wikipedia.org/w/index.php?title=COVID-19_pandemic_in_Spain&oldid=970146896) (accessed 3 August 2020).
24. Wikipedia. 2020 COVID-19 pandemic in Germany. See [https://en.wikipedia.org/w/index.php?title=COVID-19\\_pandemic\\_in\\_Germany&oldid=970982353](https://en.wikipedia.org/w/index.php?title=COVID-19_pandemic_in_Germany&oldid=970982353) (accessed 3 August 2020).
25. BBC News. 2020 Coronavirus: How lockdown is being lifted across Europe. 2 July 2020. See <https://www.bbc.com/news/explainers-52575313> (accessed 3 August 2020).

26. Giuffrida A. 2020 Italy unveils plans for easing of coronavirus lockdown restrictions. *The Guardian*. See <https://www.theguardian.com/world/2020/apr/27/italy-unveils-plans-for-easing-of-coronavirus-lockdown-restrictions-conte> (accessed 3 August 2020).
27. Wikipedia. 2020 Timeline of the COVID-19 pandemic in the United Kingdom. See [https://en.wikipedia.org/w/index.php?title=Timeline\\_of\\_the\\_COVID-19\\_pandemic\\_in\\_the\\_United\\_Kingdom&oldid=970899419](https://en.wikipedia.org/w/index.php?title=Timeline_of_the_COVID-19_pandemic_in_the_United_Kingdom&oldid=970899419) (accessed 3 August 2020).
28. BBC News. 2020 Coronavirus: How lockdown is being lifted across Europe. 31 July 2020. See <https://www.bbc.com/news/uk-53609467> (accessed 3 August 2020).
29. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, Rubin GJ. 2020 The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *Lancet* **395**, 912–920. (doi:10.1016/S0140-6736(20)30460-8)
30. Fokas AS, Cuevas-Maraver J, Kevrekidis PG. 2020 A quantitative framework for exploring exit strategies from the COVID-19 lockdown. *Chaos Solitons Fract.* **140**, 110244. (doi:10.1016/j.chaos.2020.110244)
31. Korber B *et al.* 2020 Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell* **182**, 812–827. (doi:10.1016/j.cell.2020.06.043)
32. Pinkstone J. 2020 New mutation of the coronavirus ‘may be a GOOD thing’: Strain common in Europe, America and parts of Asia is more infectious but less deadly, expert claims. MailOnline. See <https://www.dailymail.co.uk/sciencetech/article-8637821/More-infectious-coronavirus-mutation-good-thing-says-disease-expert.html> (accessed 18 August 2020).
33. Fokas AS, Cuevas-Maraver J, Kevrekidis PG. 2021 Easing COVID-19 lockdown measures while protecting the older restricts the deaths to the level of the full lockdown. *Sci. Rep.* **11**, 5839. (doi:10.1038/s41598-021-82932-8)
34. Errico JD. 2005 *Bound constrained optimization, computer program, MATLAB and simulink centre*. The MathWorks, Inc.
35. Lagarias JC, Reeds JA, Wright MH, Wright PE. 1998 Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM J. Optimization* **9**, 112–147. (doi:10.1137/S1052623496303470)
36. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. 1994 *Numerical recipes in C*. Cambridge, NY: Cambridge University Press.
37. Moré JJ. 1977 The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical analysis* (ed. GA Watson), pp. 105–116. Lecture Notes in Mathematics, vol 630. Berlin, Germany: Springer.
38. Coleman TF, Li Y. 1996 An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM J. Optimization* **6**, 418–445. (doi:10.1137/0806023)
39. Chowell G. 2017 Fitting dynamic models to epidemic outbreaks with quantified uncertainty: a primer for parameter uncertainty, identifiability, and forecasts. *Infect. Dis. Model.* **2**, 379–398. (doi:10.1016/j.idm.2017.08.001)
40. Bates DM, Watts DG. 1988 *Nonlinear regression analysis and its applications*. New York, NY: John Wiley & Sons.