



A Simulation of a COVID-19 Epidemic Based on a Deterministic SEIR Model

José M. Carcione¹, Juan E. Santos^{2,3,4}, Claudio Bagaini⁵ and Jing Ba^{2*}

¹ National Institute of Oceanography and Applied Geophysics - OGS, Trieste, Italy, ² School of Earth Sciences and Engineering, Hohai University, Nanjing, China, ³ Departamento de Energía, IGPUBA, Universidad de Buenos Aires, FIUBA, Buenos Aires, Argentina, ⁴ Department of Mathematics, Purdue University, West Lafayette, IN, United States, ⁵ Independent Researcher, Haywards Heath, United Kingdom

OPEN ACCESS

Edited by:

Zisis Kozlakidis,
International Agency for Research on
Cancer (IARC), France

Reviewed by:

Francesco Mainardi,
University of Bologna, Italy
Afshin Montakhab,
Shiraz University, Iran

*Correspondence:

Jing Ba
jba@hhu.edu.cn

Specialty section:

This article was submitted to
Infectious Diseases - Surveillance,
Prevention and Treatment,
a section of the journal
Frontiers in Public Health

Received: 22 March 2020

Accepted: 15 May 2020

Published: 28 May 2020

Citation:

Carcione JM, Santos JE, Bagaini C
and Ba J (2020) A Simulation of a
COVID-19 Epidemic Based on a
Deterministic SEIR Model.
Front. Public Health 8:230.
doi: 10.3389/fpubh.2020.00230

An epidemic disease caused by a new coronavirus has spread in Northern Italy with a strong contagion rate. We implement an SEIR model to compute the infected population and the number of casualties of this epidemic. The example may ideally regard the situation in the Italian Region of Lombardy, where the epidemic started on February 24, but by no means attempts to perform a rigorous case study in view of the lack of suitable data and the uncertainty of the different parameters, namely, the variation of the degree of home isolation and social distancing as a function of time, the initial number of exposed individuals and infected people, the incubation and infectious periods, and the fatality rate. First, we perform an analysis of the results of the model by varying the parameters and initial conditions (in order for the epidemic to start, there should be at least one exposed or one infectious human). Then, we consider the Lombardy case and calibrate the model with the number of dead individuals to date (May 5, 2020) and constrain the parameters on the basis of values reported in the literature. The peak occurs at day 37 (March 31) approximately, with a reproduction ratio R_0 of 3 initially, 1.36 at day 22, and 0.8 after day 35, indicating different degrees of lockdown. The predicted death toll is approximately 15,600 casualties, with 2.7 million infected individuals at the end of the epidemic. The incubation period providing a better fit to the dead individuals is 4.25 days, and the infectious period is 4 days, with a fatality rate of 0.00144/day [values based on the reported (official) number of casualties]. The infection fatality rate (IFR) is 0.57%, and it is 2.37% if twice the reported number of casualties is assumed. However, these rates depend on the initial number of exposed individuals. If approximately nine times more individuals are exposed, there are three times more infected people at the end of the epidemic and IFR = 0.47%. If we relax these constraints and use a wider range of lower and upper bounds for the incubation and infectious periods, we observe that a higher incubation period (13 vs. 4.25 days) gives the same IFR (0.6 vs. 0.57%), but nine times more exposed individuals in the first case. Other choices of the set of parameters also provide a good fit to the data, but some of the results may not be realistic. Therefore, an accurate determination of the fatality rate and characteristics of the epidemic is subject to knowledge of the precise bounds of the parameters. Besides the specific example,

the analysis proposed in this work shows how isolation measures, social distancing, and knowledge of the diffusion conditions help us to understand the dynamics of the epidemic. Hence, it is important to quantify the process to verify the effectiveness of the lockdown.

Keywords: COVID-19, epidemic, lockdown, SEIR model, infection fatality rate (IFR), reproduction ratio (R_0), Lombardy (Italy)

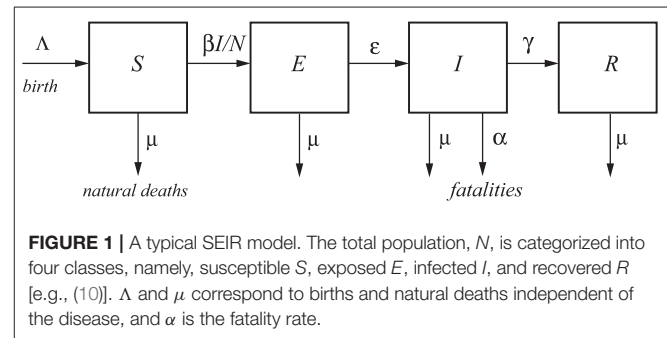
1. INTRODUCTION

The most abundant species in nature are viruses; they are parasites, since they cannot replicate themselves. Upon replication, some viruses cause serious infectious diseases in human and/or animals and are medically, socially, and economically important (1, 2). One of these species is the coronavirus. An outbreak of pneumonia caused by a novel coronavirus (COVID-19) began (officially) on February 24, 2020, in Northern Italy, and the number of newly reported cases is still increasing. Approximately 29,000 casualties are reported in Italy at the time of writing (May 5). The serious danger COVID-19 poses is reflected in the high number of cases of transmission to healthcare workers, more than 20% in Italy. The experience in China showed that the use of relative extreme isolation measures in conjunction with rapid diagnosis has a strong impact on the dynamics of the epidemic, hence the importance of understanding and quantifying the process to verify the effectiveness of the isolation measures [e.g., (3)].

There is a long history of mathematical models in epidemiology, going back to the eighteenth century. Bernoulli (4) used a mathematical method to evaluate the effectiveness of the techniques of variolation against smallpox, with the aim of influencing public health policy. Most of the models are compartmental models, with the population divided into classes and with assumptions being made about the rate of transfer from one class to another (5, 6). We consider a Susceptible-Exposed-Infected-Removed (SEIR) model to describe the spread of the virus and compute the number of infected and dead individuals. The SEIR model has many versions, and mathematical treatments can be found, for instance, in Hethcote (5), Keeling and Rohani (7), and Diekmann et al. (8), among others. The goal is to compute the number of infected, recovered, and dead individuals on the basis of the number of contacts, probability of disease transmission, incubation period, recovery rate, and fatality rate. The epidemic disease model predicts a peak of infected and dead individuals per day as a function of time and assumes that births and natural deaths are balanced, since we are dealing with a very short period of time. The population members solely decrease due to the disease as dictated by the fatality rate of the disease. The differential equations are solved with a forward Euler scheme.

2. THEORY AND DIFFERENTIAL EQUATIONS

When no vaccine is available, the isolation of diagnosed infectives and social distancing are the only control measures available.



We consider an SEIR epidemic disease model [e.g., (5, 7–9)]. The total (initial) population, N_0 , is categorized into four classes, namely, susceptible, $S(t)$, exposed, $E(t)$, infected-infectious, $I(t)$ and recovered, $R(t)$, where t is the time variable. The governing differential equations are

$$\begin{aligned}\dot{S} &= \Lambda - \mu S - \beta S \frac{I}{N}, \\ \dot{E} &= \beta S \frac{I}{N} - (\mu + \epsilon)E, \\ \dot{I} &= \epsilon E - (\gamma + \mu + \alpha)I, \\ \dot{R} &= \gamma I - \mu R,\end{aligned}\quad (1)$$

where $N = S + E + I + R \leq N_0$ in this case, and a dot above a variable denotes time differentiation. Equations (1) are subject to the initial conditions $S(0)$, $E(0)$, $I(0)$, and $R(0)$. The parameters are defined as:

Λ : Per-capita birth rate.

μ : Per-capita natural death rate.

α : Virus-induced average fatality rate.

β : Probability of disease transmission per contact (dimensionless) times the number of contacts per unit time.

ϵ : Rate of progression from exposed to infectious (the reciprocal is the incubation period).

γ : Recovery rate of infectious individuals (the reciprocal is the infectious period).

having units of $(1/T)$, with T : time. The scheme is illustrated in **Figure 1**. The choice $\Lambda = \mu = 0$ and $\epsilon = \infty$ gives the classical SIR model [e.g., (11)], while if Λ and μ are not zero, the model is termed an endemic SIR model [e.g., (12)]. However, the SIR model has no latent stage (no exposed individuals), and then it is inappropriate as a model for diseases with an ϵ such as that of COVID-19.

Let us better clarify the meaning of each quantity. N is the total number of live humans in the system at time t . S is the number of humans susceptible to be exposed, and E is the actual number of exposed individuals (a class in which the disease is latent; they are infected but not infectious); people go from S to E depending on the number of contacts with I individuals, multiplied by the probability of infection (β) (see **Figure 1**, where $\beta I/N$ is the average number of contacts with infection per unit time of one susceptible person). The other processes taking place at time t are: the exposed (E) become infectious (I) with a rate ϵ and the infectious recover (R) with a rate γ . Recovered means an individual who does not flow back into the S class, as lifelong immunity is assumed, but it remains to be seen whether patients recovered from COVID-19 will develop antibodies and achieve lifelong protection. The reciprocals ϵ^{-1} and γ^{-1} are the average disease incubation and infectious periods, respectively.

Λ is the rate of birth and μ is the natural rate of death, both per unit time. The reciprocal μ^{-1} , interpreted as the normal life expectancy (e.g., 83 years), refers to the average normal deaths (e.g., natural deaths, by normal flu, accidents, etc.) not related to the infectious disease. These quantities describe a model with vital dynamics (endemic model), which has an inflow of births into the S class at rate Λ and deaths into the other classes at rates μS , μI , and μR (see **Figure 1**). If $\Lambda = \mu N$, the deaths balance the newborns. The number of live people at time t is $N(t) = S(t) + E(t) + I(t) + R(t)$, which can be lower or higher than N_0 depending on the values of Λ and μ . In this case, it is lower than N_0 .

One of the key parameters, besides β , is α , which represents the disease-related fatality rate (3, 13). In a very fast pandemic, we may assume that there are no births and normal deaths (or that they balance and $\Lambda = \mu N$), but there are deaths due to the fatality rate of the disease. This rate is an average, because the model does not take into account the age (a far higher portion of old people die from the disease than young people), the patients' preexisting conditions, and the healthcare quality.

In summary, susceptible persons enter the exposed class with a rate proportional to β and remain there for a mean incubation period ϵ^{-1} , i.e., those already infected with the disease but not able to transmit it are in the exposed class and progress to the infectious class, to recover at the rate γ and die at the rate α . It is important to recall that the E class does not have the symptoms of the disease, because they are incubating it. They will have symptoms when they pass to class I . Individuals in class I may not have symptoms (asymptomatic), but they are infectious, while those in class E are not. Moreover, individuals in class E can move to R without showing symptoms, but they become infectious when they are in class I .

The dead population as a function of time is $D(t) = N_0 - N(t)$, whereas the curve giving the dead people per unit time is

$$\dot{D}(t) = -\dot{N}(t) = -(\dot{S} + \dot{E} + \dot{I} + \dot{R})(t). \tag{2}$$

Another equivalent approach is an SEIDR model [e.g., (14, 15)], where we have to add

$$\dot{D}(t) = \alpha I(t) \tag{3}$$

to Equations (1). In Keeling and Rohani [(7), section 2.2], $\alpha/(\gamma + \mu) = \rho/(1 - \rho)$, where ρ is the per capita probability of dying from the infection. It can easily be shown that Equations (2) and (3) are equivalent if births and natural deaths compensate.

2.1. Reproduction Ratio

The basic reproduction ratio, R_0 , is the classical epidemiological measure associated with the reproductive power of the disease. For the SEIR model, it is

$$R_0 = \frac{\beta \epsilon}{(\epsilon + \mu)(\gamma + \alpha + \mu)} \tag{4}$$

(8, 13). It gives the average number of secondary cases of infection generated by an infectious individual. Therefore, it is used to estimate the growth of the virus outbreak. R_0 provides a threshold for the stability of the disease-free equilibrium point. When $R_0 < 1$, the disease dies out; when $R_0 > 1$, an epidemic occurs. The behavior of SEIR models as a function of R_0 can be found, for instance, in Al-Sheikh (16).

2.2. Infection and Case Fatality Rates

The infection fatality rate (IFR) is based on all the population that has been infected, i.e., including the undetected individuals and asymptomatic. In terms of the recovery and fatality rates, we have

$$\text{IFR} (\%) = 100 \cdot \frac{D_\infty}{R_\infty + D_\infty}, \tag{5}$$

since the total humans that have been infected is the sum of the recovered and dead individuals, where the subscript refers to the end of the epidemic ($t \rightarrow \infty$). It can easily be shown that, using the last Equations (1) and (3), we obtain

$$\text{IFR} (\%) = 100 \cdot \frac{\alpha I_\infty}{(\alpha + \gamma)I_\infty - \mu R_\infty} \approx 100 \cdot \frac{\alpha}{\alpha + \gamma} \approx 100 \cdot \frac{\alpha}{\gamma}, \tag{6}$$

since the term containing μ is much smaller, because $\mu \ll \alpha \ll \gamma$, and Equation (6) holds approximately at all times, not only at the end of the epidemic. On the other hand, the case fatality rate (CFR) considers the number of deaths related to the diagnosed individuals, and $\text{CFR} > \text{IFR}$ is always true, since the number of diagnosed individuals is lower than the denominator of Equation (5). The CFR is time-dependent and is the value that is usually reported.

3. NUMERICAL ALGORITHM

We solve the differential Equations (1) by using a forward Euler finite-difference scheme [e.g., (17)], discretizing the time variable as $t = ndt$, where n is a natural number and dt is the time step.

After discretization, Equations (1) and (2) become:

$$\begin{aligned} S^{n+1} &= S^n + dt \left(\Lambda - \mu S^n - \beta S^n \frac{I^n}{N^n} \right), \\ E^{n+1} &= E^n + dt \left[\beta S^n \frac{I^n}{N^n} - (\mu + \epsilon) E^n \right], \\ I^{n+1} &= I^n + dt \left[\epsilon E^n - (\gamma + \mu + \alpha) I^n \right], \\ R^{n+1} &= R^n + dt (\gamma I^n - \mu R^n), \\ \dot{D}^n &= -(\dot{S}^n + \dot{E}^n + \dot{I}^n + \dot{R}^n)(t), \end{aligned} \quad (7)$$

where \dot{D}^n is the number of dead people in only the specific day n . This algorithm yields positive and bounded solutions [e.g., see (6) and Problem 1.42(iv) in (8)], and the system converges to an equilibrium, i.e., $S^n + R^n + D^n = S_\infty + R_\infty + D_\infty = N_0$ for $t \rightarrow \infty$.

4. RESULTS

Let us consider the following base parameters as an example so as to analyze the results by varying some of them. $N_0 = 10$ million, $\alpha = 0.006/\text{day}$, $\beta = 0.75/\text{day}$, $\gamma = (1/8)/\text{day}$, $\epsilon = (1/3)/\text{day}$, $\Lambda = \mu N$ (balance of births and natural deaths), and initial conditions: $S(0) = N_0 - E(0) - I(0)$, $E(0) = 20,000$, $I(0) = 1$ and $R(0) = 0$. These data are taken from Chowell et al. [(3), **Table 1**] for SARS and imply an average disease incubation (latent period) of 3 days and an infectious period of 8 days. The data correspond to no isolation conditions among individuals and an epidemic situation (high β , $R_0 = 5.72 > 1$).

The time step of the Euler scheme to solve the discretized Equations (7) is $dt = 0.01$ day. **Figure 2** shows the number of individuals in the different classes (**Figure 2A**) and also the total number of dead people (D) and the number of dead people per specific day (\dot{D}) (**Figure 2B**). As can be seen, the peak of dead individuals per day is reached at day 30. The high values in **Figure 2B** do not consider complete home isolation and social distancing measures (or “suppression”). The maximum number of infected individuals is almost 4 million. According to data from China, around 5% of people who tested positive for COVID-19 experience severe symptoms and require admission to an intensive-care unit, almost 200,000 individuals in this case. Under these conditions, the health system would be completely overwhelmed, with very high death rates and an inability to provide intensive care. A partial “mitigation” strategy involving social distancing (home isolation of suspect cases and social distancing of the elderly) would not be enough, and a severe lockdown is required in order to make it possible to decrease R_0 to less than 1 (20).

Hereafter, we vary the parameters and plot the infected (I) individuals, i.e., excluding those who are incubating the disease (E). In order for the process to start, there should be at least one exposed or one infectious individual. **Figure 3** shows the number of infected individuals for $R_0 > 1$ (a) and $R_0 \leq 1$ (b), where

all the other parameters are kept constant except β , which takes the value

$$\beta \approx (\gamma + \alpha)R_0, \quad (8)$$

for μ much smaller than γ and α ($\mu^{-1} \approx 83$ years in Italy). We recall here that β is the probability of transmission times the number of contacts per unit time. Basically, with a reduction in β (and R_0), the peak decreases in intensity but moves to later times for R_0 higher than 1 (**Figure 3A**), although it is wider. There is a significant reduction in the number of infected individuals for $R_0 \leq 1$, meaning that strict home isolation is very effective below a given threshold.

The effect of the initial number of exposed individuals is shown in **Figure 4** for two sets of values of R_0 , greater (**Figure 4A**) and less (**Figure 4B**) than 1. **Figure 4A** indicates that more exposed people does not mainly affect the intensity of the peak; rather, it precipitates the spread of the epidemic, so that the location of the peak is highly dependent on $E(0)$. On the other hand, **Figure 4B** shows that for $R_0 < 1$, the peak location does not change, but its intensity changes significantly, indicating an effective “suppression” of the epidemic, with more exposed leading to more infectious. **Figure 5** indicates that the incubation period ($1/\epsilon$) also has an impact on the results. If $R_0 > 1$ (**Figure 5A**), increasing the period from 3 to 9 days decreases the maximum number of infected individuals by almost half and delays the spread of the epidemic, but the peak is wider. If $R_0 < 1$, the curves behave similarly, but there are much fewer infected cases. The initial number of infectious individuals (from 1 to 10,000) has no apparent effect on the results, as can be seen in **Figure 6**, but this is not the case when we deal with the real case history (see next section). The effects of the infectious period are shown in **Figure 7**, where, as expected, increasing this quantity delays the epidemic when $R_0 > 1$. Below $R_0 = 1$, the number of infected individuals decreases substantially.

Let us now assume that isolation precautions have been imposed and that after day 22, R_0 changes from 5.72 to 0.1 [a change of β according to Equation (8)], and we consider the same parameters to produce **Figure 2**. The results are shown in **Figure 8**, where the peak has moved from day 30 to day 25, with a significant slowing in the number of new cases. The total number of dead individuals has decreased, and the number of dead individuals per day at the peak has decreased from 22 to 13 K, approximately. Extreme isolation after imperfect isolation anticipates the process. **Figure 9** shows the results if the isolation measures start two days earlier, at day 20 instead of day 22. The number of casualties decreases from 220 to 155 K.

4.1. The Lombardy Case

Next, we attempt to model the COVID-19 epidemic in Lombardy (Italy), for which data are available at <https://github.com/pcm-dpc/COVID-19>. The time of writing is day 72 (May 5), and the available data allow us to perform a relatively reliable fit to the total number of casualties from day 1 to date. On day 69 (May 2), 329 casualties were reported, of which 282 are equally distributed in April since this number is a late report of the hospitals, corresponding to the whole month of April. Predicting the behavior of the epidemic with high accuracy is nearly

TABLE 1 | Constraints and initial–final values of the inversion algorithm.

Case	Variable →	α (day ⁻¹)	β_1 (day ⁻¹)	β_2 (day ⁻¹)	β_3 (day ⁻¹)	ϵ^{-1} (day)	γ^{-1} (day)	$E(0)$	I_∞ (M) L (day) D_∞
	Lower bound	10 ⁻⁵	0.5	10 ⁻⁶	10 ⁻⁶	3	3	10 ³	
	Upper bound	10 ⁻¹	0.9	10 ³	10 ³	6	6	2 × 10 ⁵	
	Initial value	0.006	0.75	0.5	0.5	5	5	10 ⁴ /10 ⁵	
1.1	Final value	0.00144	0.75	0.34	0.2	4.25	4.02	11,460	2.69
1.2	IFR	0.57%							262–Nov 11
1.3	R_0		3.00	1.36	0.80				15,652
2.1	Final value (*)	0.0051	0.702	0.29	0.132	5.45	4.75	9,460	1.33
2.2	IFR	2.37 %							264–Nov 13
2.3	R_0		3.25	1.34	0.61				31,934
3.1	Final value (*)	0.00142	0.75	0.57	0.395	5.79	3.31	99,500	6.49
3.2	IFR	0.47%							236–Oct 16
3.3	R_0		2.47	1.87	1.30				30,544
	Lower bound	10 ⁻⁵	0.5	10 ⁻⁶	10 ⁻⁶	2	2	10 ³	
	Upper bound	10 ⁻¹	0.9	10 ³	10 ³	20	20	2 × 10 ⁵	
	Initial value	0.006	0.75	0.5	0.2	5/15	5/15	10 ³ /10 ⁵	
4.1	Final value	0.00436	0.59	0.29	0.094	6.10	5.28	8,800	0.69
4.2	IFR	2.25%							239–Oct 19
4.3	R_0		3.04	1.50	0.48				15,652
5.1	Final value	0.0011	0.81	0.33	0.03	13	5.53	91,900	2.49
5.2	IFR	0.60%							247–Oct 27
5.3	R_0		4.45	1.81	0.16				15,345
6.1	Final value	0.0073	0.755	0.23	0.125	4.87	5.11	170	0.44
6.2	IFR	3.59%							269–Nov 18
6.3	R_0		3.72	1.13	0.61				16,401
7.1	Final value	0.09	0.9	0.28	0.175	2.99	6.15	32	0.04
7.2	IFR	35.62%							212–Sep 22
7.3	R_0		3.56	1.11	0.69				16,112
8.1	Final value	0.00674	0.83	0.006	0.01	12.79	14.93	1,270	0.18
8.2	IFR	9.15%							272–Nov 21
8.3	R_0		11.2	0.08	0.13				16,681
9.1	Final value	0.0055	0.506	0.044	0.01	11.08	14.97	8,960	0.22
9.2	IFR	7.60%							268–Nov 17
9.3	R_0		7.00	0.61	0.14				16,653

Several cases that honor the data. $I(0) = 1000$, except Case 7, with $I(0) = 1$.

(*) Doubling the number of casualties.

The values of β refer to the periods (in days): [1, 22], [22, 35], and [35, ∞] (in days).

I_∞ (in millions) indicates the total number of infected individuals at the end of the epidemic.

L denotes the day of the last infected individual, obtained when $I < 1$.

D_∞ is the death toll at the end of the epidemic.

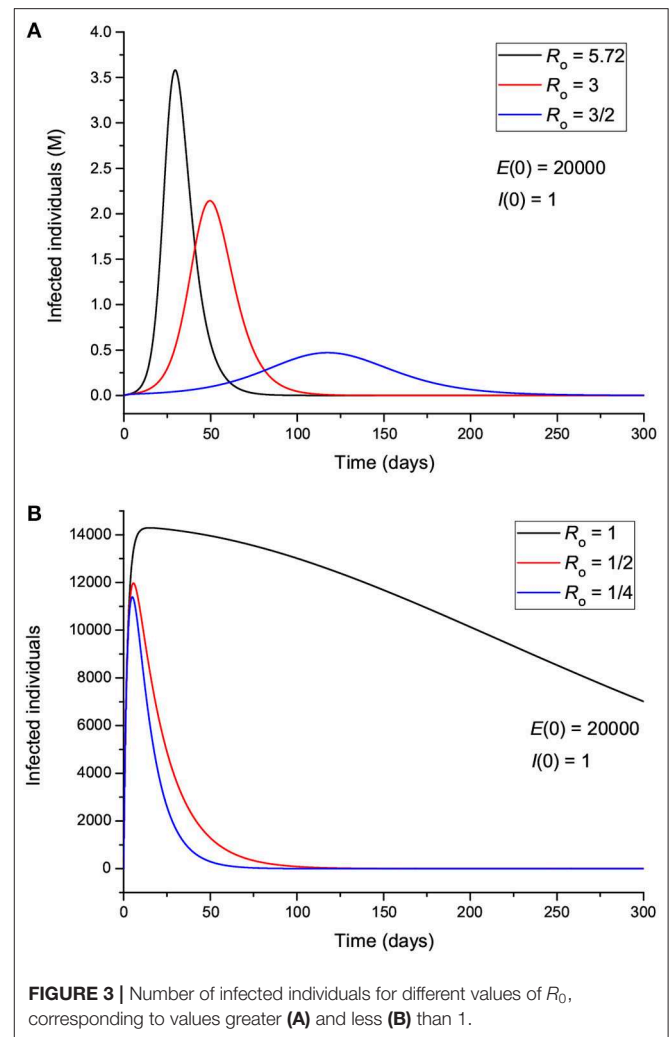
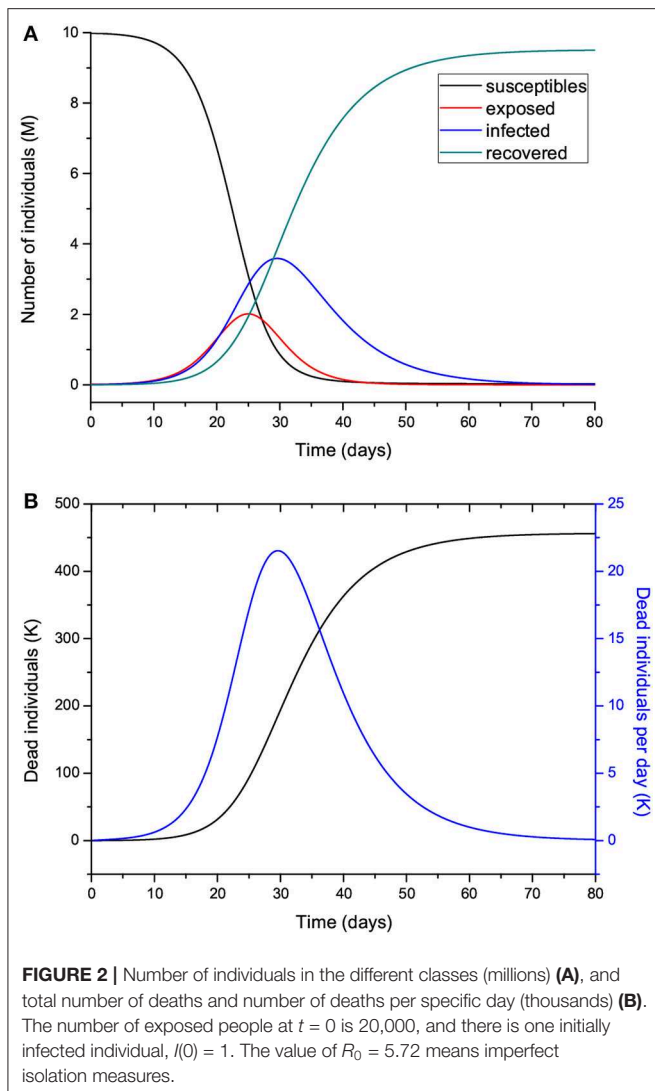
Read et al. (13) report the mean values $\epsilon^{-1} = 4$ days and $\gamma^{-1} = 3.6$ days.

Lauer et al. (19) report $\epsilon^{-1} = 5.1$ days.

Ferguson et al. (20) estimate an average IFR = 0.9%.

impossible due to there being many unknown factors, e.g., the degree of spatial distancing, probability of disease transmission, characteristics of the disease, and parameters of the epidemic. Uncertainties are related to parameter β , which varies with time, while the others are assumed to lie between certain bounds and also contribute to the error. Relative predictions of the trend require an analysis of the data, particularly to define the variation of β and R_0 with time. We do not assume a specific continuous

function, but a general approach should consider a partition into discrete periods, $[t_0, t_1], [t_1, t_2] \dots [t_{L-1}, \infty]$, guided by the measures taken by the state and the behavior of the population. In this case, $t_0 = \text{day } 1$, $t_1 = \text{day } 22$, and $t_2 = \text{day } 35$, i.e., $L = 3$, since after t_1 (March 16), home isolation, social distancing, and partial national lockdown started to be effective, as indicated by an inflection point in the curve of casualties per day (see below), although it is debatable whether the Italian government



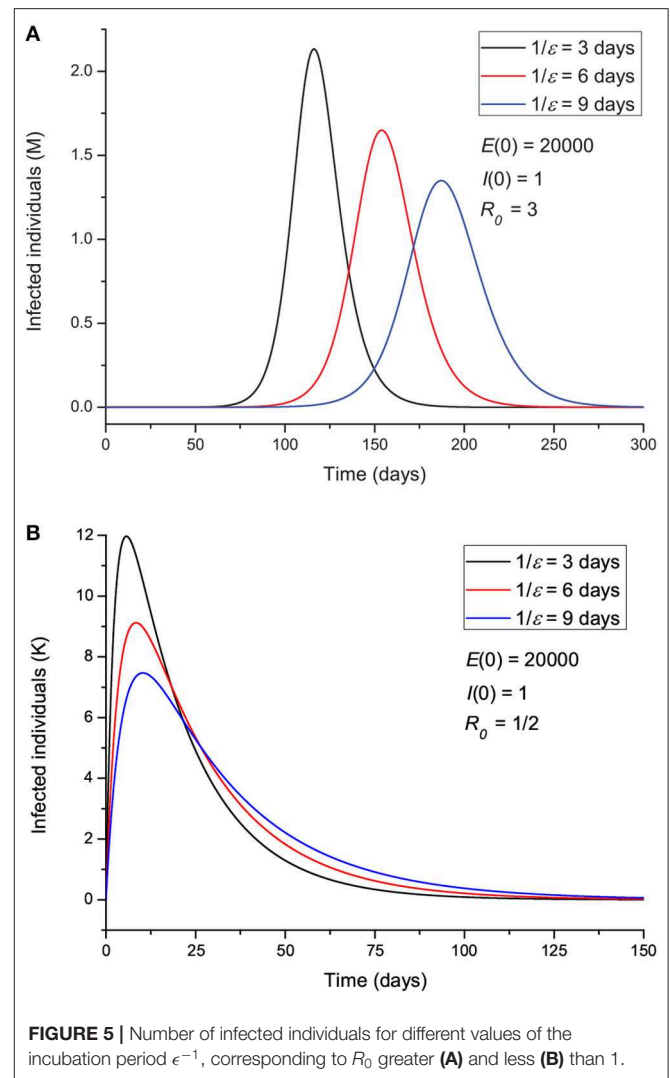
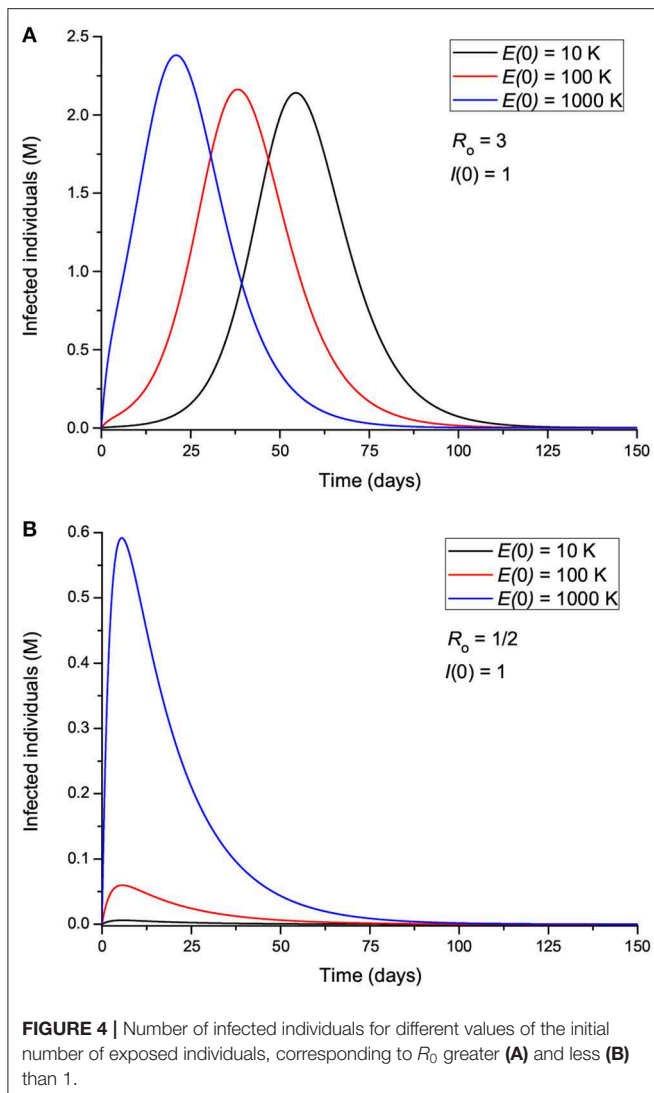
followed the same rules as in Wuhan, China. We also observe that at t_2 (March 29), the curve starts to bend downwards and reach a “peak.” This partition into three periods is valid to date, but the trend can have unpredictable behavior due to the factors mentioned above, too early removal of the lockdown conditions, etc.

The reported infected people cannot be used for calibration because these data cannot be trusted. The hospitalization numbers cannot be considered to be representative of the number of infected people, and, at present, the number of asymptomatic, undiagnosed infections is largely unknown. However, we are aware that even using the number of casualties is uncertain, since there can be an under-ascertainment of deaths, but the figures cannot vary as much as the error related to the infected individuals. Hence, the reported number of deceased people could possibly be underestimated due to undeclared cases. This number depends on the country (quality of the health system) and average age of the population, but it is certain that this novel virus is more deadly and spreads more quickly than seasonal

flu. Moreover, authorities make a distinction between a death that occurred “with the co-action” of the virus and a death “caused by” the virus. Indeed, only a small percentage of the casualties were in a healthy condition prior to the infection, and most of the patients were already affected by other illnesses (e.g., diabetes, dementia, cancer, stroke). Therefore, we also consider cases where 100% more people actually died per day compared to the official figures.

In order to accomplish the fit, we use the simulated annealing algorithm developed by Goffe et al. (21). The Fortran code can be found at: <https://econwpa.ub.uni-muenchen.de/econ-wp/prog/papers/9406/9406001.txt>. The fit is based on the L^2 -norm and yields α , β_1 (before t_0), β_2 (after t_0), β_3 (after t_1), ϵ , $E(0)$, and γ from the beginning of the epidemic (day 1, February 24) to date (day 72, May 5), i.e., seven free parameters. We use the total number of deaths for the calibration.

Table 1 shows the constraints, initial values, and results for different cases, where Cases 1 and 2 correspond to approximately nine times fewer exposed individuals at the beginning of the epidemic, and Cases 2 and 3 assume double casualties. Cases 4 and 5 consider a wider range of the lower and upper bounds

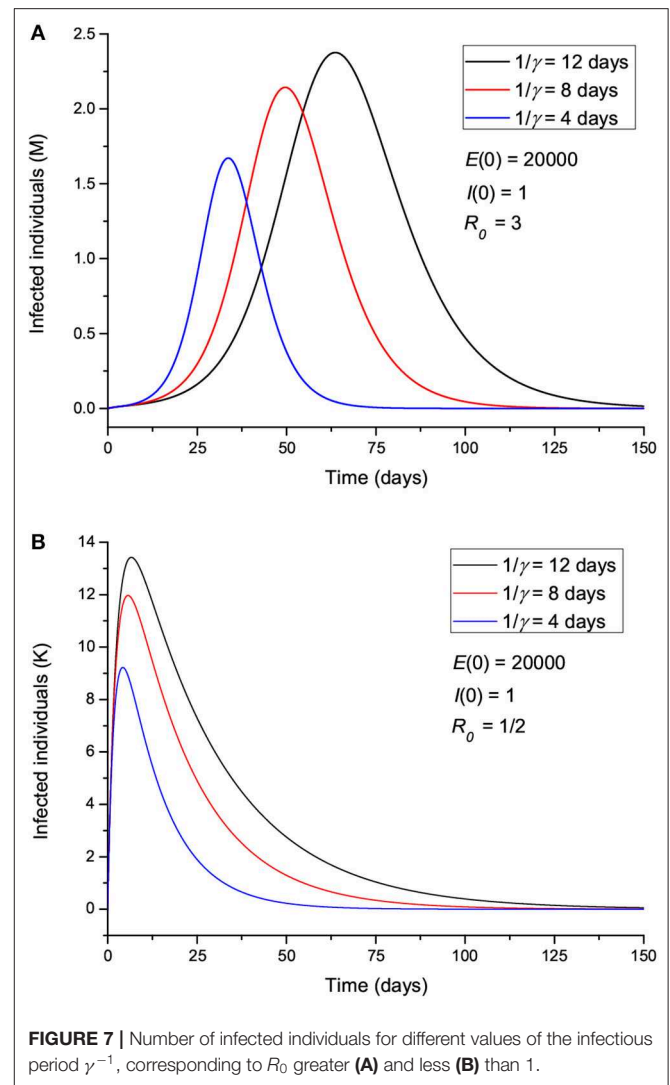
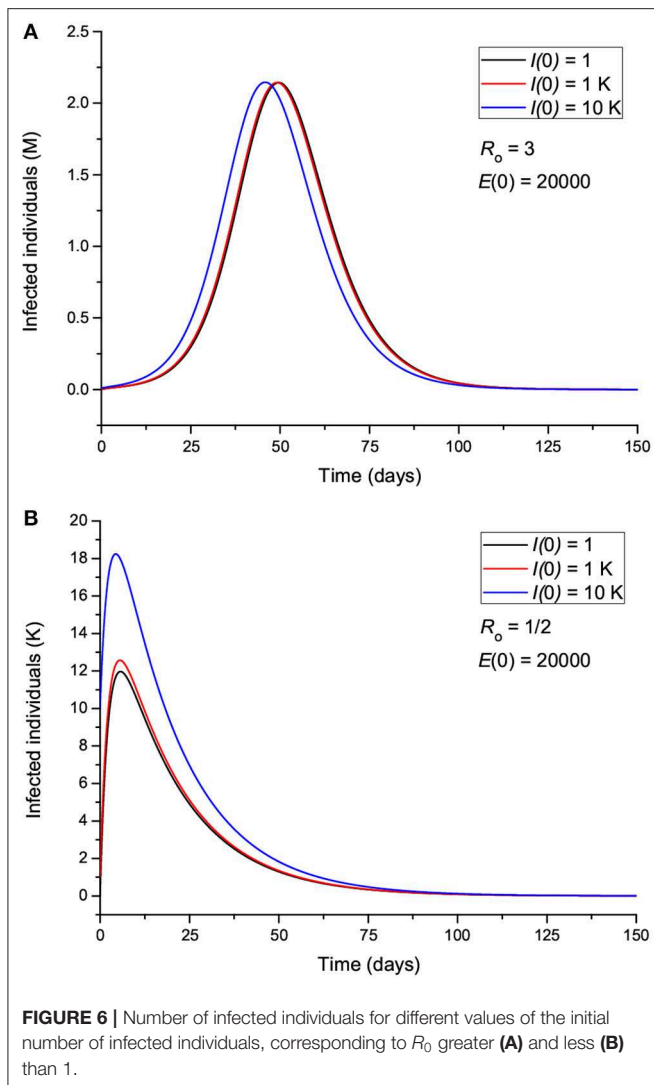


for the incubation and infectious periods (ϵ^{-1} and γ^{-1}). The last column does not correspond to variables but indicates the number of infected individuals at the end of the epidemic, i.e., $I_\infty = R_\infty + D_\infty \approx R_\infty$, the day of the last infected individual (the end of the epidemic in theory), and the death toll D_∞ . The results are very sensitive to variations in parameter β , and, consequently, those of R_0 , mostly due to the impact of the intervention strategies performed.

Figure 10 shows the curves of Case 1 compared to the data (black dots), with IFR = 0.57% and R_0 decreasing from 3 to 0.8 by the end of the epidemic. The final number of infected individuals is 2.69 million (see **Figure 11A**, **Table 1**). The peak value of the I class is 0.3 M or 300,000 individuals. If 5% of these people require admission to an intensive-care unit (ICU), this amounts to 15,000 individuals and substantially exceeds the capacity of Lombardy, which was approximately 1,000 ICU beds on March 16. **Figure 11B** compares the infectious and dead individuals (per day) and, as expected, the two curves are synchronous, since a proportion α of infectious individuals die. The inflection point at day 22 (**Figure 10B**) indicates that the isolation measures started

to be effective. Strict isolation could not be achieved by day 22 for several reasons, and there is a reasonable delay of a few days before it can be implemented (day 35). The total number of casualties is approximately 15,600, and the effective duration of the epidemic is about 100 days. However, see the last column indicating the day when the last individual is infected, obtained with the condition $I < 1$. Recent data reveal that the effective duration of the Wuhan epidemic was almost 60 days [(22), **Figure 1B**], a shorter period that was favored by the very strict isolation measures applied in that city. Case 2, which considers twice as many casualties and the results for which are shown in **Figure 12**, has a high fatality rate, IFR = 2.37%, but 1.33 million infected people. If the number of exposed individuals is much higher (Case 3), we obtain IFR = 0.47% and 6.5 million infected people (see **Figures 13, 14**, **Table 1**), but in this case, the fit is not optimal at the beginning of the epidemic. The calculations indicate the uncertainty related to the initial number of exposed individuals, i.e., those that are incubating the disease.

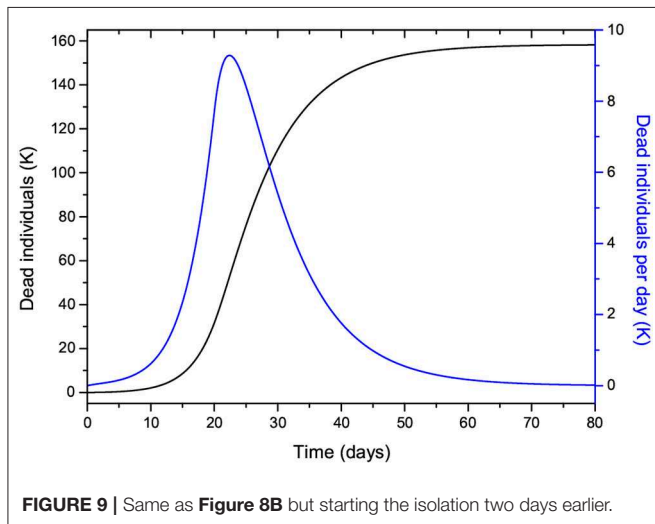
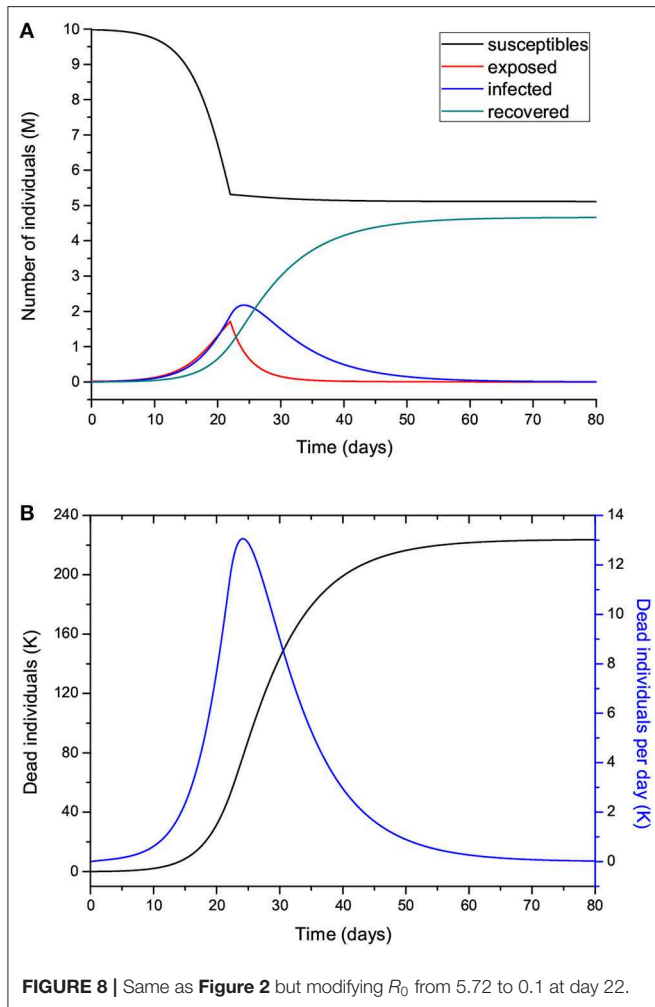
In the following, we do not show the plots, but the results honor the data. If we modify the constraints and use a wider range of lower and upper bounds, the results are those of Cases



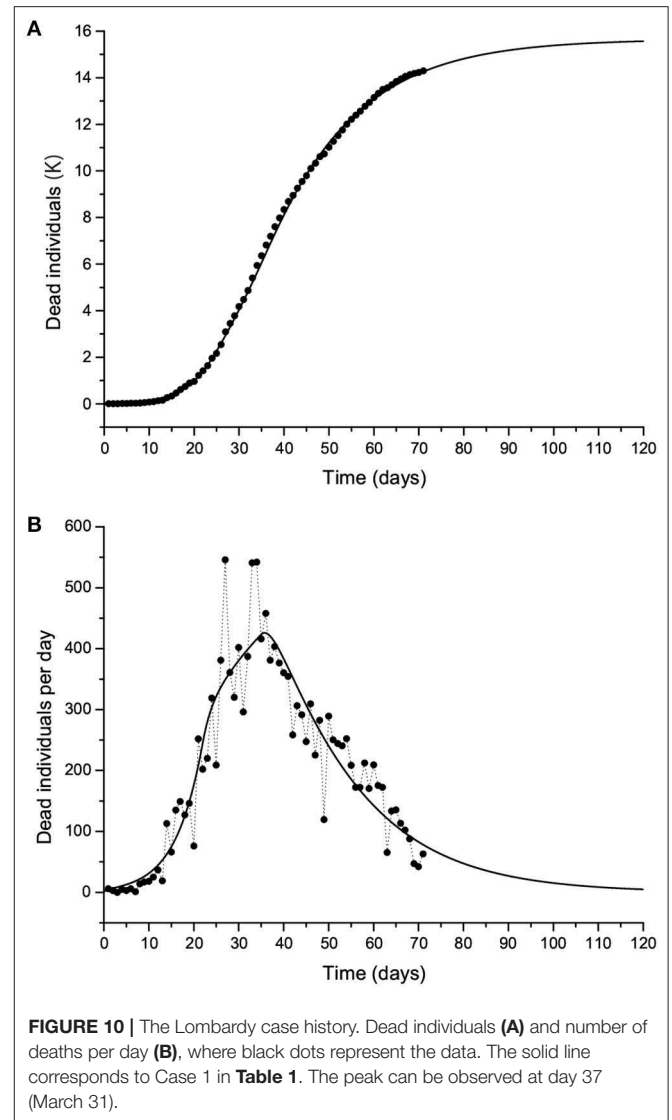
4 and 5 in **Table 1**. Case 4 has a slightly higher incubation and infectious periods compared to Case 1 but a higher IFR (2.25 vs. 0.57%), whereas there being more exposed individuals yields an incubation period of 13 days and lower IFR (Case 5). Case 6 considers that, initially, there are a few exposed individuals (we start with one). The algorithm achieves a very good fit to the data with IFR = 3.6%, comparable periods to Case 1, and 0.44 M infected individuals. Case 7, which considers $I(0) = 1$ and starts with one exposed individual, shows a good fit, but IFR is too high and so possibly wrong, indicating that at day 1, there were more exposed and infectious individuals. Fewer initially exposed and infectious individuals requires a higher IFR to fit the curve, but a higher R_0 could also have the same effect if the IFR is kept within a realistic range. Finally, we constrain the incubation and infectious periods to between 10 and 20 days, and the results are those of Cases 8 and 9, assuming different initial numbers of exposed individuals. The calculations yield fatality rates comparable to that of SARS (3), as in Case 6. These calculations indicate the uncertainty in the determination

of the parameters of the epidemic, but the solutions have to be restricted to reasonable values of the properties of the disease and parameters of the epidemic.

The values in **Table 1** can be compared to figures reported in the literature. The fatality rate and IFR depend on the age of the population. Verity et al. [(23), **Table 1**] estimate an IFR = 0.657% for China but a rate of 3.28% for those over the age of 60. If the number of infected people is several times higher than the number of reported cases, the actual fatality rate could be considerably lower than the official one, suggesting that this disease is less deadly than SARS and MERS, although much more contagious. Read et al. (18) report a mean value $R_0 = 4$, while Wu et al. (22) obtain values between 1.8 and 2. According to Chowell et al. (3), IFR = 4.8% for SARS, and Verity et al. (23) state that the average case fatality rate (CFR) of SARS is higher than that of COVID-19, with the latter being approximately 1.38% (their IFR is 0.657%). However, this virus seems to be much more contagious. The meaning of α^{-1} is the life expectancy of



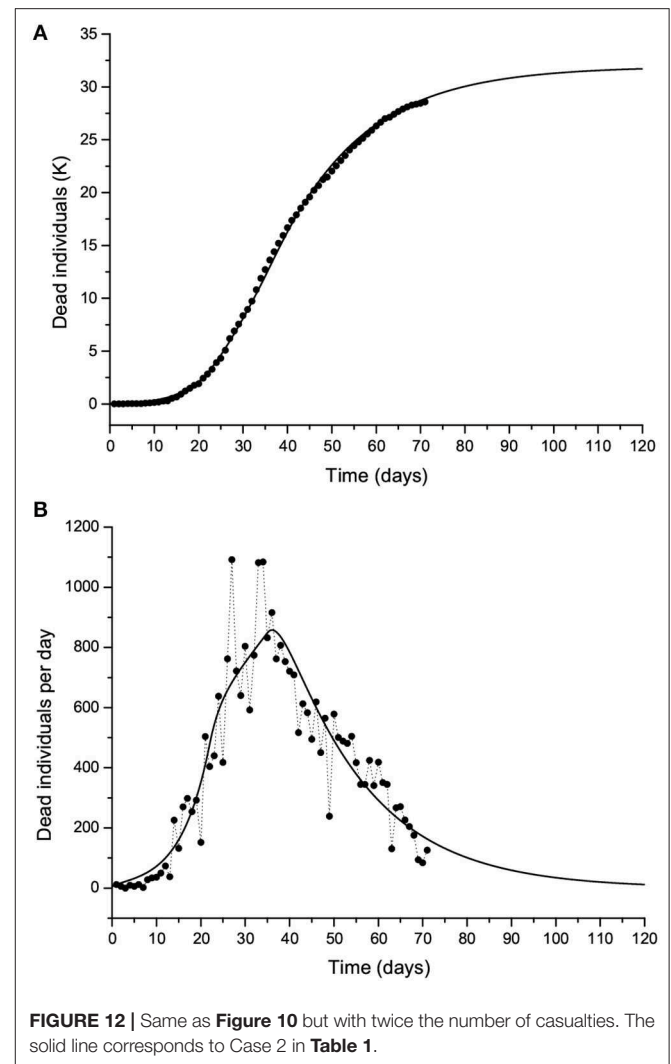
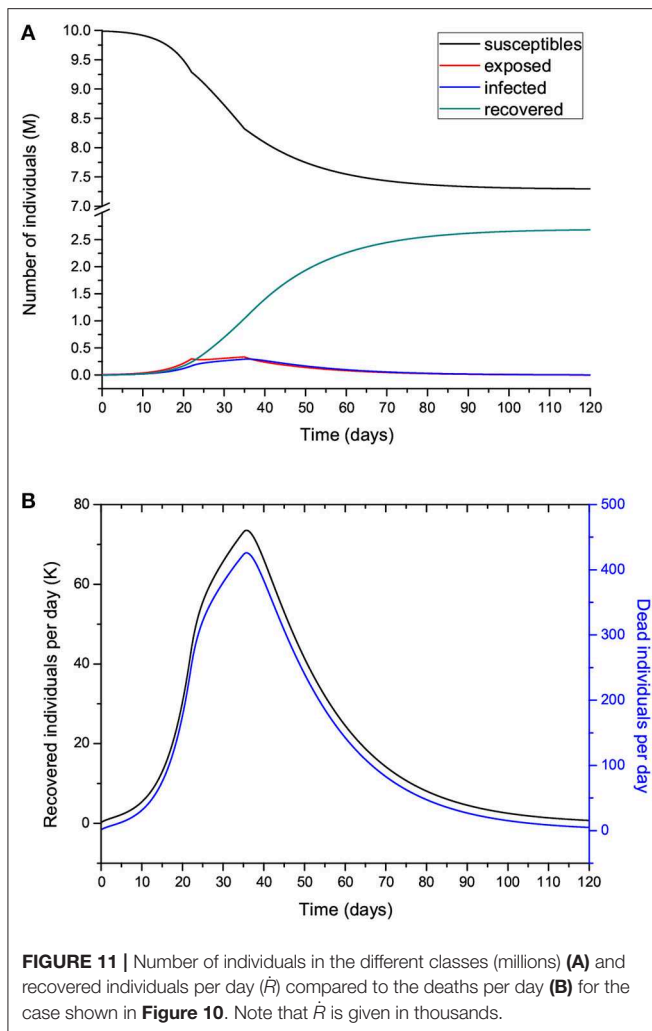
an individual in the infectious class, i.e., if $\alpha = 0.00144/\text{day}$ (Case 1), the expectancy is 694 days.



4.2. Further Comments

There are more complex versions of the SEIR model such as for instance, including a quarantine class and a class of isolated (hospitalized) members (24) or generalizing the diffusion (Equation 1) with the use of temporal fractional derivatives. The replacement of the first-order temporal derivative by a Caputo fractional derivative of non-natural order provides an additional parameter to fit the data [e.g., (25–27)]. Furthermore, the model can be made two-dimensional by including the spatial diffusion of the virus [e.g., (28)]. An alternative to spatial diffusion models is to use contact networks. The actual compartmental network through which the disease spreads is a very important part of epidemic spreading. The model used in this study is a homogeneous approximation to these network models (29–31).

Moreover, the model can be improved by including other classes. De la Sen et al. (14) propose an SEIADR model, where A are asymptomatic infectious and D are dead-infective. In other models, recovered can become susceptible again [e.g., (32)],



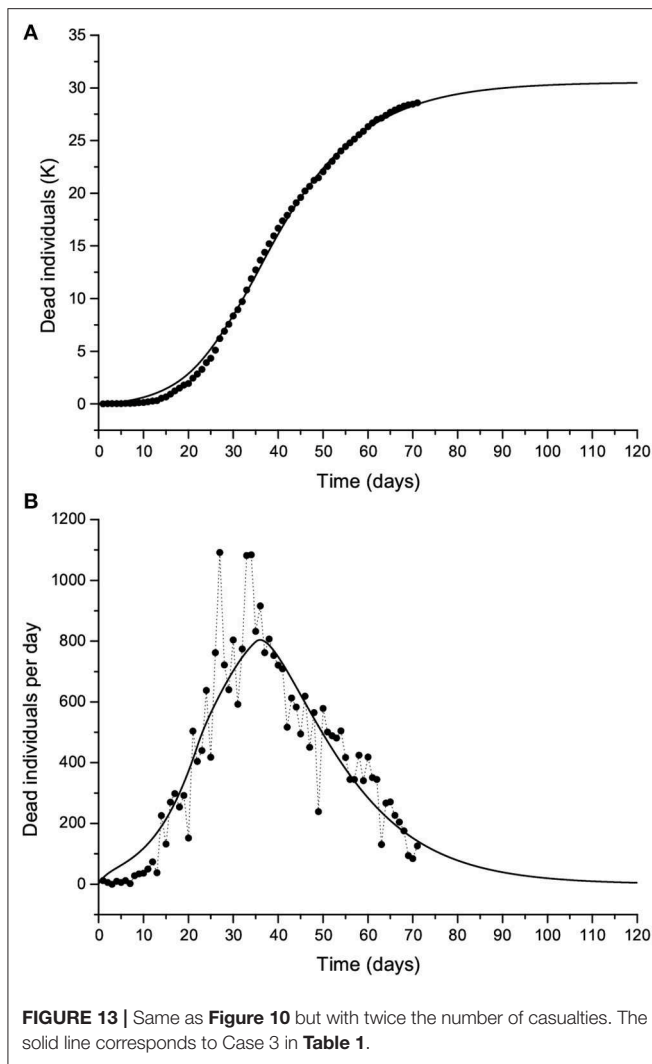
and, in addition, there are stochastic models (12), although the calibration becomes extremely difficult with the incomplete data provided by the authorities and the high number of parameters to be found. Finally, since signals propagate instantaneously in diffusion equations, the model predicts that there are more infectious humans (I class) than actual before the latent period and at late stages of the epidemic. Solutions to this problem can be found, for instance, in Keeling and Rohani [(7), Section 3.3]. At the end of the epidemic, more precise information about the parameters will be available, and the complete data can be used to evaluate the development of β (and R_0) with time.

The outbreak of a pandemic can have catastrophic consequences, not only from the point of view of the casualties but also economically. Therefore, it is essential to absolutely avoid it by taking the necessary measures at the right time, something that has not been accomplished in Italy and the rest of the world. According to these calculations, the effective measures are social distancing and home isolation, since there is no health system designed for ordinary circumstances that can be prepared for a pandemic, when the number of infected individuals grows exponentially. As can be seen, the pandemic can develop in a

few days and the number of casualties can be extremely high if the fatality rate and contagiousness of the disease are high. The difference of only a few days in taking action can make a big difference in the prevention of this disaster. The pandemic and its consequences were predicted in October 2019 by a group of experts (<https://www.politico.com/news/magazine/2020/03/07/coronavirus-epidemic-prediction-policy-advice-121172>), but states ignored the fact and transnational nature of the threat, delaying the necessary measures to avoid the disaster, minimizing in many cases the downsides to their own populations and economies. Moreover, in less than three weeks, the virus has overloaded the healthcare system all over northern Italy, particularly in Lombardy, where the system cannot support this type of emergency and the authorities are not prepared to deal with the epidemic.

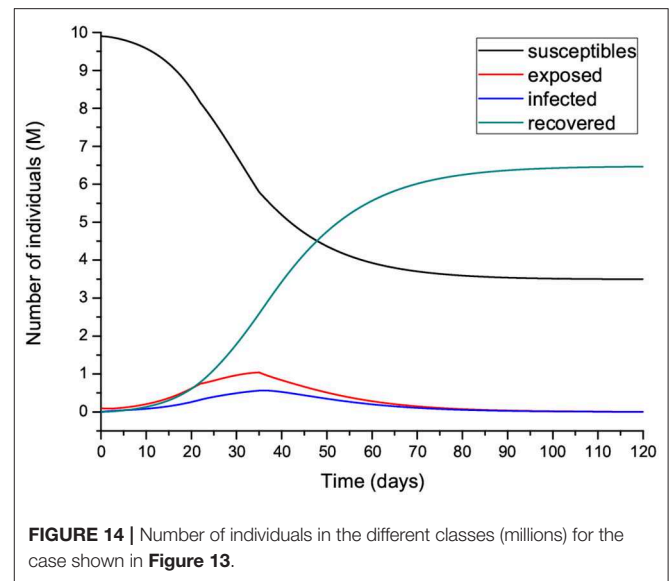
5. CONCLUSIONS

A high number of secondary COVID-19 infections can take place when an infected individual is introduced into a community. It



is essential to simulate the process of infection (and death) in advance so as to apply adequate control measures and mitigate the risk of virus diffusion. One of the most commonly used mathematical algorithms to describe the diffusion of an epidemic disease is the SEIR model, which we have applied to compute the number of infected, recovered, and dead individuals on the basis of the number of contacts, probability of disease transmission, incubation and infectious periods, and disease fatality rate.

A first analysis of the results of the model is based on the parameters of the SARS disease, and we assume that the parameters do not change during the whole epidemic. When the number of contacts is reduced, the peak decreases in intensity but moves to a later time period, although it is wider. Moreover, a larger number of exposed people does not affect the intensity of the peak but precipitates the epidemic. The incubation period also has an impact on the results, with higher values delaying the epidemic. The dependence on the initial number of infected people is apparently weak if R_0 does not change during the epidemic. Increasing the infectious period has the same effect as increasing the incubation period.



Moreover, the day when isolation starts is important, since a difference of only 2 days makes a big difference to the number of casualties.

The Lombardy modeling assumes 10 million individuals and has been calibrated on the basis of the total number of casualties. The results show that the peak occurs after 37 days, with a final number of dead individuals depending on the reproduction ratio R_0 . With the presently available data, this number is approximately 15,600. Up to day 72 (May 5, the day of writing), the reproduction ratio is 3 before March 16 (day 22), 1.36 between March 16 and March 29 (day 35), and 0.8 after March 29, whereas the fatality rate is 0.00144/day (IFR = 0.57%). We have also doubled the number of casualties and obtained IFR = 2.37 and 0.47%, with the second value corresponding to nine times more exposed individuals. These values are obtained by constraining the incubation and infectious periods to values reported in the literature. If we relax these constraints and use a wider range of lower and upper bounds, we obtain slightly higher incubation and infectious periods compared to the first case but a much higher IFR (2.25 vs. 0.57%), while using many more exposed individuals yields an incubation period of 13 days and a lower IFR (0.6%). Of the many solutions that honor the data, we suggest that those that agree with the experimental data published at present, based on the reported incubation and infection periods and IFR, be considered more realistic. The uncertainty is due to the novelty of the virus, whose properties were unknown two months ago, and the initial conditions, i.e., the initial number of exposed and infectious individuals.

The present data fit and consequent prediction of the epidemic does not take into account the second phase established by the state, which started on May 4. After the partial opening of the economy and under a less stringent lockdown, the reproduction number could increase and induce a second outbreak of the epidemic. Therefore, a precise determination of the fatality rate is subject to knowledge of the parameters of the epidemic and characteristics of the disease, and it is clear from these calculations that the usefulness of simple models for prediction

is limited and that their main role is to help in our understanding of the dynamics of the epidemic.

Models can be used to predict and understand how an infectious disease spreads in the world and how various factors affect the dynamics. Even if the predictions are inaccurate, it has been clear to scientists from many decades that quarantine, social distancing, and the adoption of very strict health and safety standards are essential to stop the spread of a virus. Quarantine was even implemented in medieval times to fight the black death before there was knowledge of the existence of viruses. In this sense, this pandemic reveals the failure of policy-makers, since it is well-known from basic modeling results that earlier adoption of those measures can save thousands of lives and even prevent the pandemic. The interface of science, society, and politics is still uneasy, even in highly developed countries, revealing a disregard for scientific evidence. Moreover, one of the consequences is that some of these countries do not invest sufficiently in R&D and must acquire the new technology from overseas at a much higher cost.

REFERENCES

- Spinney L. *Pale Rider: The Spanish Flu of 1918 and How It Changed the World*. London: Jonathan Cape (2017).
- Adachi A. Grand challenge in human/animal virology: unseen, smallest replicative entities shape the whole globe. *Front Microbiol.* (2019) 11:431. doi: 10.3389/fmicb.2020.00431
- Chowell G, Fenimore PW, Castillo-Garsow MA, Castillo-Chavez C. SARS outbreak in Ontario, Hong Kong and Singapore: the role of diagnosis and isolation as a control mechanism. *J Theor Biol.* (2003) 224:1–8. doi: 10.1016/S0022-5193(03)00228-5
- Bernoulli D. *Essai d'une Nouvelle Analyse de la Mortalité causée par la Petite vérole et des Avantages de l'inoculation pour la prévenir*. Paris: Mémoires de Mathématiques et de Physique; Académie Royale des Sciences (1760). p. 1–45.
- Hethcote HW. The mathematics of infectious diseases. *SIAM Rev.* (2000) 42:599–653. doi: 10.1137/S0036144500371907
- Brauer F. Mathematical epidemiology: past, present, and future. *Infect Dis Model.* (2017) 2:113–27. doi: 10.1016/j.idm.2017.02.001
- Keeling MJ, Rohani P. *Modeling Infectious Diseases in Humans and Animals*. Princeton, NJ; Oxford: Princeton University Press (2008).
- Diekmann O, Heesterbeek H, Britton T. *Mathematical Tools for Understanding Infectious Disease Dynamics*. Princeton Series in Theoretical and Computational Biology. Princeton, NJ: Princeton University Press (2013).
- Al-Showaiikh F, Twizell E. One-dimensional measles dynamics. *Appl Math Comput.* (2004) 152:169–94. doi: 10.1016/S0096-3003(03)00554-X
- Chitnis N, Mac Hyman J, Cushing JM. Determining important parameters in the spread of malaria through the sensitivity analysis of a mathematical model. *Bull Math Biol.* (2008) 70:1272–96. doi: 10.1007/s11538-008-9299-0
- d'Onofrio A, Manfredi P, Salinelli E. Dynamic behaviour of a discrete-time SIR model with information dependent vaccine uptake. *J Diff Equat Appl.* (2015) 22:485–512. doi: 10.1080/10236198.2015.1107549
- Allen LJS. A primer on stochastic epidemic models: formulation, numerical simulation, and analysis. *Infect Dis Model.* (2017) 2:128–42. doi: 10.1016/j.idm.2017.03.001
- Zhang LJ, Li Y, Ren Q, Huo Z. Global dynamics of an SEIRS epidemic model with constant immigration and immunity. *WSEAS Trans Math.* (2013) 12:630–40.
- De la Sen M, Ibeas A, Alonso-Quesada S, Nistal R. On a new epidemic model with asymptomatic and dead-infective subpopulations with feedback

DATA AVAILABILITY STATEMENT

Datasets used in this study can be retrieved at the time of the online publication from <https://github.com/pcm-dpc/COVID-19>.

AUTHOR CONTRIBUTIONS

JC wrote the theoretical section and performed the analysis. JS verified the figures and contributed to the analysis. CB aided in the solution of the differential equations. JB contributed to the discussion.

ACKNOWLEDGMENTS

This manuscript has been released as a pre-print at <https://arxiv.org/> (33). JC dedicates this work to Antonio Buonomo, who passed away of COVID-19 on March 21, 2020.

- controls useful for Ebola disease. *Discr Dyn Nat Soc.* (2017) 2017:1–22. doi: 10.1155/2017/4232971
- Sameni R. *Mathematical Modeling of Epidemic Diseases; A Case Study of the COVID-19 Coronavirus* (2020). Available online at: <https://arxiv.org/abs/2003.11371>
- Al-Sheikh S. Modeling and analysis of an SEIR epidemic model with a limited resource for treatment. *Glob J Sci Front Res Math Decis Sci.* (2012) 12:56–66.
- Carcione JM. *Wave Fields in Real Media. Theory and Numerical Simulation of Wave Propagation in Anisotropic, Anelastic, Porous and Electromagnetic Media*. 3rd Edn. Amsterdam: Elsevier (2014).
- Read JM, Bridgen JRE, Cummings DAT, Ho A, Jewell CP. Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. (2020) 1–11. doi: 10.1101/2020.01.23.20018549
- Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Ann Intern Med.* (2020) 4232971. doi: 10.7326/M20-0504
- Ferguson NM, et al. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. (2020) 1–20. doi: 10.25561/77482
- Goffe WL, Ferrier GD, Rogers J. Global optimization of statistical functions with simulated annealing. *J Econ.* (1994) 60:65–99. doi: 10.1016/0304-4076(94)90038-8
- Wu JT, Leung K, Bushman M, Kishore N, Niehus N, de Salazar PM, et al. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nat Med Lett.* (2020) 26:506–10. doi: 10.1038/s41591-020-0822-7
- Verity R, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis (2020) 1–9. doi: 10.1016/S1473-3099(20)30243-7
- Brauer F, Castillo-Chavez C. *Mathematical Models in Population Biology and Epidemiology*. New York, NY: Springer (2012).
- Caputo M, Carcione JM, Cavallini F. Wave simulation in biological media based on the Kelvin-Voigt fractional-derivative stress-strain relation. *Ultrasound Med Biol.* (2011) 37:996–1004. doi: 10.1016/j.ultrasmedbio.2011.03.009
- Mainardi F. *Fractional Calculus and Waves in Linear Viscoelasticity*. London: Imperial College Press (2010).

27. Chen Y, Cheng J, Jiang X, Xu X. The reconstruction and prediction algorithm of the fractional TDD for the local outbreak of COVID-19. *arXiv[Preprint].arXiv:2002.10302* (2020).
28. Naheed A, Singh M, Lucy D. Numerical study of SARS epidemic model with the inclusion of diffusion in the system. *Appl Math Comput.* (2014) 229:480–98. doi: 10.1016/j.amc.2013.12.062
29. Pastor-Satorras R, Vespignani A. Epidemic spreading in scale-free networks. *Phys Rev Lett.* (2001) 86:3200–3. doi: 10.1103/PhysRevLett.86.3200
30. Montakhab A, Manshour P. Low prevalence, quasi-stationarity and power-law behavior in a model of contagion spreading. *Explor Front Phys.* (2012) 99:1–6. doi: 10.1209/0295-5075/99/58002
31. Pastor-Satorras R, Castellano C, Van Mieghem P, Vespignani A. Epidemic spreading in complex networks. *Rev Modern Phys.* (2015) 87:925–79. doi: 10.1103/RevModPhys.87.925
32. Xia W, Kundu S, Maitra S. Dynamics of a delayed SEIQ epidemic model. *Adv Diff Equat.* (2018) 336:1–21. doi: 10.1186/s13662-018-1791-8
33. Carcione JM, Santos JE, Bagaini C, Ba J. A simulation of a COVID-19 epidemic based on a deterministic SEIR model. *arxiv[Preprint].arxiv:2004.03575* (2020). doi: 10.1101/2020.04.20.20072272

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Carcione, Santos, Bagaini and Ba. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.