CoViD-19 in Italy: a mathematical model to analyze the epidemic containment strategy and the economic impacts

di Fabio Verachi[1,] Luca Trussoni[2] (LTlogics), Luciano Lanzi (Freelance)[3]

Abstract

The objective of this paper is to evaluate the potential costs deriving from the adoption of the CoViD-19 epidemic management strategy. For this purpose, we developed a specific methodology that combines an epidemiological model, known in the literature as "SIR" (Susceptible - Infected - Recovered), and a probabilistic state model, also known as "multi-state". The model was then parameterized using the dataset published by the Italian Government through the Civil Protection and the Istituto Superiore di Sanità. Leveraging the identified parameters, we estimated the total amount of ICU hospital, non-ICU hospital and home isolation days that needed to extinguish a single-wave pandemic. Then we proceeded estimating the related hospital costs in the ongoing pandemic management strategy. We furthermore analyzed different scenarios that can be useful to assess choices the Government could adopt to face unforeseen events and outlined a way to exploit the results of the model (and in particular, the total amount of home isolation days and the overall timespan of the pandemic wave). These considerations represent a first step to assess the overall direct impact on Gross Domestic Product (GDP) of the pandemic.

1. Introduction

As recently in the spotlight, the CoViD-19 epidemic refers to the so-called coronavirus disease, technically caused by SARS-CoV-2 virus (coronavirus 2 from severe acute respiratory syndrome), found to be similar to at least 70% of its gene sequence to that of SARS-CoV. Initially identified in the city of Wuhan in December 2019, the capital of the Chinese province of Hubei, it subsequently spreads to more than 210 countries in the world[4].

The first confirmed cases of the virus in human patients were found towards the end of November 2019. From the middle of January 2020, the first cases were also found outside of China, brought by international travelers, mainly to the major trading partner nations of the country. The following table shows the spread of CoViD-19 for the main countries affected by the infection (as of April 19th, 2020)[5].

Country	Confirmed Cases	Deaths	Recovered
United States	715,105	37,889	63,841
Italy	175,925	23,227	44,927
Spain	191,726	20,043	74,797
France	147,969	18,681	34,420
United Kingdom	114,217	15,464	N/A
Belgium	37,183	5,453	8,348
Iran	80,868	5,031	55,987
China	82,719	4,632	77,029
Germany	142,569	4,405	85,400
Netherlands	31,589	3,601	250
Brazil	34,913	2,201	14,026
Turkey	78,546	1,769	8,361

---

[1] IntesaSanpaolo, Enterprise Risk Manager. Mail: fabio.verachi@intesasanpaolo.com. The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policies or position of Intesa Sanpaolo Group.

[2] LTlogics, Owner. Mail: lg.trussoni@ltlogics.it.

[3] Freelance, Risk Manager. Mail: luciolan@hotmail.com.

[4] Further details in David S. Hui, Esam EI Azhar, Tariq A. Madani, Francine Ntoumi, Richard Kock, Osman Dar, Giuseppe Ippolito, Timothy D. Mchugh, Ziad A. Memish, Christian Drosten e Alimuddin Zumla "The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health – The latest 2019 novel coronavirus outbreak in Wuhan, China" in International Journal of Infectious Diseases, vol. 91, 14 January 2020, pp. 264–266.

[5] For confirmed cases, deaths and recovered in each country see "Confirmed Cases and Deaths by Country, Territory, or Conveyance" in Worldometer.

Country	Confirmed Cases	Deaths	Recovered
🇸🇪 Sweden	13,822	1,511	550
🇨🇦 Canada	32,412	1,346	10,543
🇨🇭 Switzerland	27,404	1,336	16,400

Table 1 – The spread of CoViD-19 pandemic in April 2020

2. Analysis

The model we propose consists of two sub-models that can be independently defined: (i) an epidemiological model, and (ii) a state model that describes the management of the infected people. The first sub-model is a Susceptible, Infected, Recovered (SIR) one, that estimates the speed of spread and the size of the epidemic (in line with Batista, [15] and [21]), while the state model is a proposal from the authors, analogous to models used in the credit risk assessment, through which we will estimate the epidemic management needs (in terms of hospitalization costs).

The choice of a SIR-type model is motivated both by the literature findings and by the fact that there is a broad consensus that even during the incubation period, people exposed to SARS-Cov2 virus should be considered infectious.

3. Methodology

3.1 SIR model

The methodology proposed in this paper for the epidemiological component is a variant of the SIR model. The use of an epidemiological model is of fundamental importance to quantitatively define the infection dynamics.

It is also necessary to determine the possible strategies for identifying, preventing and managing the disease as well as, as will be seen, to estimate the direct and indirect economic impacts.

The SIR model was first proposed in the 1920s by Kermack and McKendrick to explain the rapid growth and subsequent decrease in the number of infected people observed in some epidemics such as the London plague in 1665-1666, Bombay in 1906 or London cholera in 1865.

The logic behind the methodology is to distinguish, among the population, the susceptible individuals (the uninfected but susceptible to infection) and the infected individuals (infected people who are able to transmit the disease).

The underlying mathematical model is a system of three differential equations:

$$\begin{cases} \dfrac{dS}{dt} = -\beta \dfrac{SI}{N} \\ \dfrac{dI}{dt} = \beta \dfrac{SI}{N} - \gamma I \\ \dfrac{dR}{dt} = \gamma I \end{cases}$$

where S represents the population susceptible to the disease, I the infected population and R the "removed" population that is no longer able to spread the disease (composed by the healed and immunized people and the deceased people). The constants β and γ identify the evolution of the disease (respectively an infectivity parameter and the reciprocal of the average time of evolution of the disease). The SIR model is characterized by the following dynamics:

$$\begin{cases} \dfrac{dS}{dt} = -\beta \big(1 - \delta H(t - t^*)\big) \dfrac{SI}{N} \\ \dfrac{dI}{dt} = \beta \dfrac{SI}{N} - p_{catch} \gamma_1 I - (1 - p_{catch}) \gamma_2 I \\ \dfrac{dR_1}{dt} = p_{catch} \gamma_1 I \\ \dfrac{dR_2}{dt} = (1 - p_{catch}) \gamma_2 I \end{cases}$$

with the following meaning of the variables:

- N represents the consistency of the population in which the epidemic develops;
- S represents, in every moment of time, the "susceptible" population, which has not been infected by the disease;
- I represents the population that has been infected with the disease and that is active in the spreading phenomenon;
- p_{catch} represents the probability that a case of an infected person will be caught by the disease management system;
- γ_1 and γ_2 respectively represent the reciprocals of the average times during which the infected spread the disease. In the model, it will be assumed $1/\gamma_1 = 5.2$ days and $1/\gamma_2 = 15$ days (in line with the indications in Li et al. [2]);
- R_1 and R_2 represent the population of the "removed", or those who are no longer affected by the epidemic (because they are isolated, healed or deceased), respectively if they have been intercepted or not by the disease management system;
- β represents the speed of free spread of the epidemic;

- δ represents the effectiveness of the lockdown in limiting the spread of the epidemic, starting from the time t^* corresponding to March 8th, 2020. The function H is the classic Heaviside function ($H(x) = 1$ se $x \geq 0$, $H(x) = 0$ se $x < 0$).

A classic parameter of epidemiological models is the "basic reproduction number" that represents the average number of susceptible individuals converted into infected for each individual infected, usually indicated with the acronym r_0.

The proposed model is characterized by three r_0 parameters, one for the "catched" population $r_{0,catched} = \beta/\gamma_1$, one for the "non-catched" population $r_{0,non\ catched} = \beta/\gamma_2$, and an overall estimation $r_0 = \beta/(p_{catch}\gamma_1 + (1 - p_{catch})\gamma_2)$.

The model is adapted to observations provided in the Protezione Civile GitHub repository [23], applying the method described in the Appendix. The following values are found:

Parameter	Value	Parameter	Value
N	1,995,998	δ	60.0%
β	0.2770	t^*	38
		p_{catch}	5.38%

Table 2 – Parameters of the SIR submodel

In the values above we can implicitly infer the "basic reproductive numbers", equal to $r_0 = 4.01$; $r_{0,catched} = 1.44$; $r_{0,non\ catched} = 4.16$ before the lockdown and $r_0 = 2.41$; $r_{0,catched} = 0.86$; $r_{0,non\ catched} = 2.50$ after it. The values, although high, do not disagree with the literature which in particular reports, for the pre-lockdown values, different ranges referring to different geographical areas[6]:
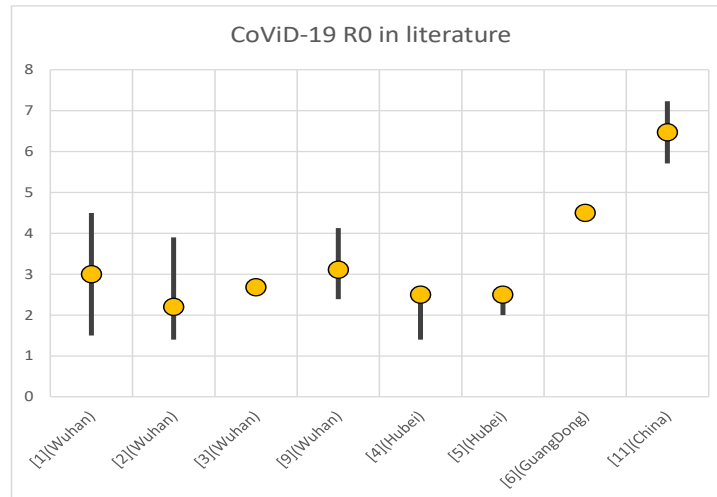
Chart 1 – The R0 of CoViD-19 pandemic in different location as available in literature in April 2020

Below we report some charts to compare the results of the model with observations:
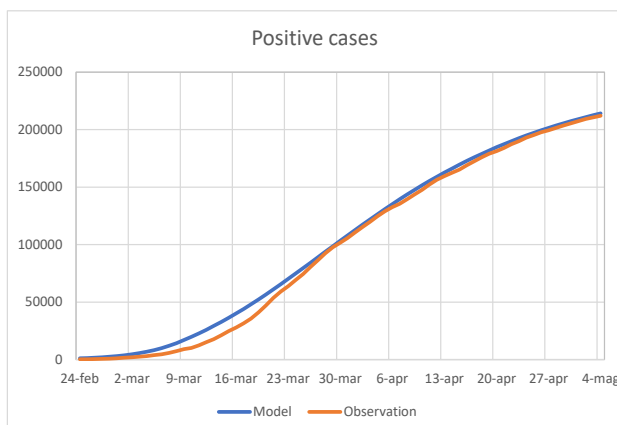
2a – Dynamics of total infected cases vs observations
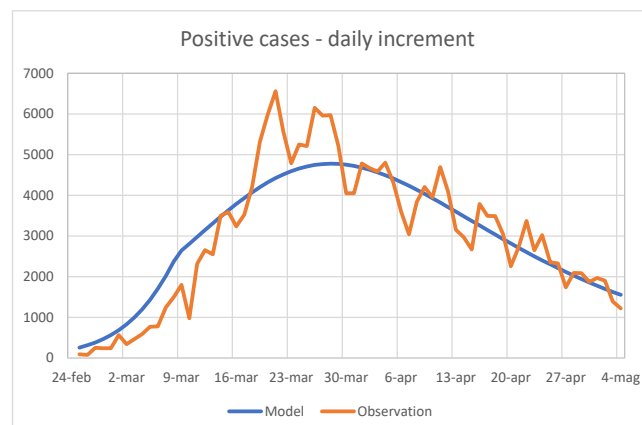
2b – Dynamics of daily increment of positive cases vs observations

Chart 2 – SIR submodel performance vs observations

---

[6] For example, in the Diamond Princess the r_0 of the epidemic spread is estimated at 14.8 [7]. The authors wish to cite the Gabriel Goh epidemic simulator, https://gabgoh.github.io/COVID/index.html, in whose documentation can be found the values that have been used for the reported graph.

We point out that the model seems to support the conclusion that the lockdown, however effective in reducing the r_0, was not enough to reduce it below value one in not-catched cases. This phenomenon can be explained by the fact that the main consequence of the lockdown is to have locked the value of N at two million people, instead of allowing it to rise up to the entire Italian population of 60 million people.

It follows that the circulation of the disease remains high within the "confined" but not isolated population. This conclusion seems plausible to the authors and in line with other studies: an extreme case was the Diamond Princess cruise ship, where the illness circulation exhibited an r_0 equal to about 15 (Rocklov et al. [7]).

The authors consider a high level of r_0 reasonable for the spread of the disease within the confined population, also in light of the preliminary study of the Istituto Superiore di Sanità which highlights how the contagions in the period 1st - 23rd April 2020 in Italy are for 2/3 occurred in rest homes and in the family environment (ISS [26]).

This population must be considered as the "circle" of the infected, within which the lockdown was only partially effective (think of the case of those people infected in home isolation who certainly no longer infect their work colleagues but expose their family members to a significant risk of contagion).

The value of p_{catch} is also remarkable, the interpretation of which is that the model implies the presence of about 18.6 not symptomatic/not recognized infection cases for each case intercepted by the national health service. Different estimates have been proposed for asymptomatic to symptomatic infection rate ratio: for example, a 5:1 ratio for the Wuhan province has been proposed by Day [24], while Li et al. [25] propose a range between 5:1 and 10:1. On the other hand, Wu et al. [27] find a 94% asymptomatic cases for the Chinese outbreak, in line with our result.

Finally, the value t^* implies that the model estimates the beginning of the circulation of the disease in Italy on January 17th (initiation level at 10 infected).

3.2 SEIR model

Many variations have been introduced over time, such as the SEIR models, which explicitly consider the class of "exposed" people: this model is particularly valid for diseases with a significant incubation time, during which the infected are not yet contagious:

$$\begin{cases} \dfrac{dS}{dt} = -\beta \dfrac{SI}{N} - \epsilon E \\[2mm] \dfrac{dE}{dt} = \epsilon E - \gamma I \\[2mm] \dfrac{dI}{dt} = \beta \dfrac{SI}{N} - \gamma I \\[2mm] \dfrac{dR}{dt} = \gamma I \end{cases}$$

where the parameter ϵ represents the reciprocal of the incubation time. Further variants arise from the explicit consideration of time delays. For the CoViD-19 epidemic, researchers have proposed both the direct application of the SIR model (Nesteruk [22]) and the SEIR model (Zhou et al. [16]), variants of the SEIR with additional components (Tang et al. [11]), the explicit modeling of external factors (such as injection or subtraction of population due to travel Wu et al. [3]) or in which those exposed contribute to the spread of the disease ([16]).

For the sake of completeness, alternative modeling approaches not based on deterministic differential equations should be mentioned. These models, based on stochastic approaches, exhibit better performance in explaining the disease transmission dynamics when the size of population is small (see for example Allen et al. [28]). In consideration of the fact that we are dealing with CoViD-19 modeling at a nation-wide scale, these approaches have not been used in this paper.

3.3 Patient management model

The patient management model focuses on the evolution of the $R_1(t)$ aggregate of the SIR model. The basic idea stems from the observation that during the accumulation phase, visually stable relationships were observed between the deaths and the number of patients in intensive care on the previous day.

Such empirical observations suggested that patient management could be modeled using approaches known in the literature as "multi-state". A multi-state model describes how a single individual moves between a series of states in discrete time. Movements in a space of discrete states is governed by the transition intensity matrix which represents the instantaneous risk of moving from one state to another state.

In our application, this means that the transition of the health status of the subjects takes place through successive states of illness depending on its timespan.

The following model was therefore formulated:

- for each head that enters in the aggregate R_1 at time t three periods are identified $M_1 = [t, t + T_1]$, $M_2 = [t + T_1, t + T_2]$, $M_3 = [t + T_2, +\infty]$;
- at any instant of simulated time the head can be in one of five states ($R_{1G}, R_{1D}, R_{1Hm}, R_{1Hs}, R_{1ICU}$), which respectively represent the patients recovered, deceased, in home isolation, hospitalized or in intensive care;

- for each period of illness M_i we define the probabilities $p_i^{Hm}, p_i^{Hs}, p_i^{ICU}$ that the head (not healed or deceased) is in one of the three states on any day of the period;
- for each of the periods M_i is defined a differentiated probability of daily death $p_{D,Mi}$; the heads pass to the deceased state according to the length of the period (in particular, in each period the deaths are calculated by applying a first in first out logic);
- for the periods M_2 and M_3 two daily probabilities of recovery $p_{H,Mi}$ are specifically defined; the heads pass to the healed state according to the length of the period (again utilizing the same first in first out logic).

The model is therefore identified by a total of 13 parameters (in the appendix the structure of the model and the technique used for parameter fitting are detailed).

With regard to the parameters, the following values are obtained:

|					
$p_{D,M1}$	3.06%	$p_{H,M3}$	1.58%	p_1^{Hs}	53.2%
$p_{D,M2}$	0.13%	p_1^{Hm}	35.0%	p_2^{Hs}	50.6%
$p_{D,M3}$	0.16%	p_2^{Hm}	45.0%	p_3^{Hs}	13.2%
$p_{H,M2}$	3.28%	p_3^{Hm}	86.0%	T_1	4 days
				T_2	5 days

Table 3 – Fitted parameters of the multi-state submodel

From parameters, complementing to 1, we can easily deduce the probabilities of being treated in intensive care in the three periods p_i^{ICU}, respectively 11.8%, 4.4% and 0.8%.

The model exhibits a good fit with the observed values, as highlighted by the following charts (we called Hospital ratio the fraction of hospitalized infected compared to the intercepted infected, and ICU ratio the fraction of hospitalized patients treated in intensive care):
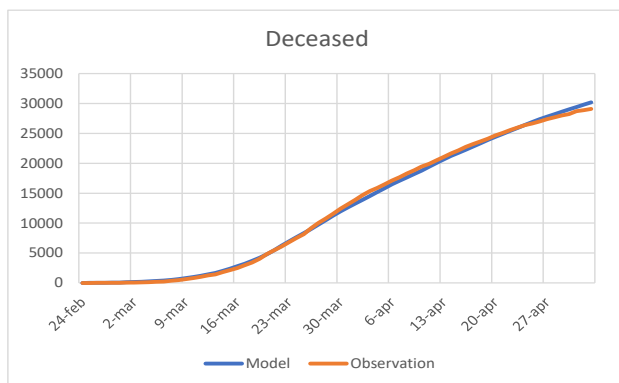
3a – Model implied deceased vs observations
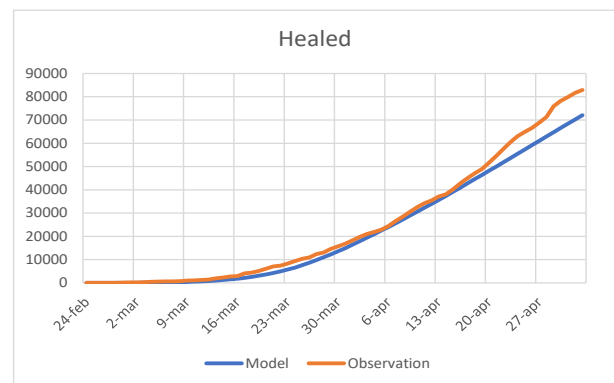
3b – Model implied healed vs observations
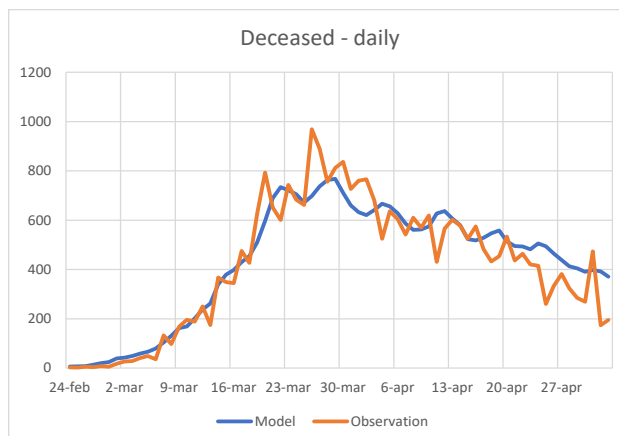
3c – Model implied daily increment of deceased, vs observations

3d – Model implied daily increment of deceased, vs observations

3e – Model implied hospital ratio vs observations

3f – Model implied ICU ratio vs observations
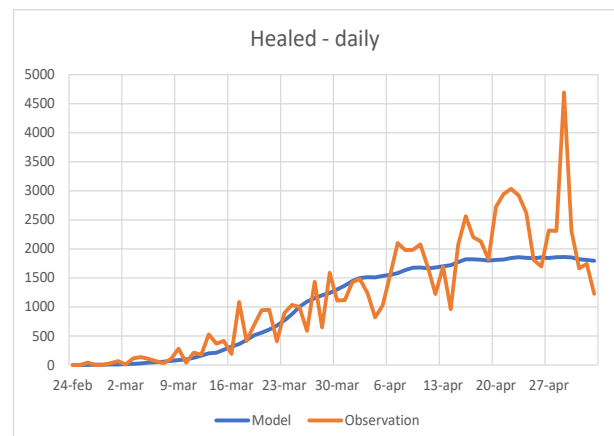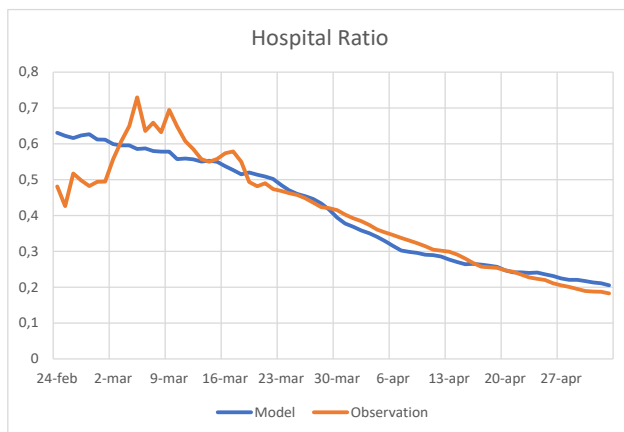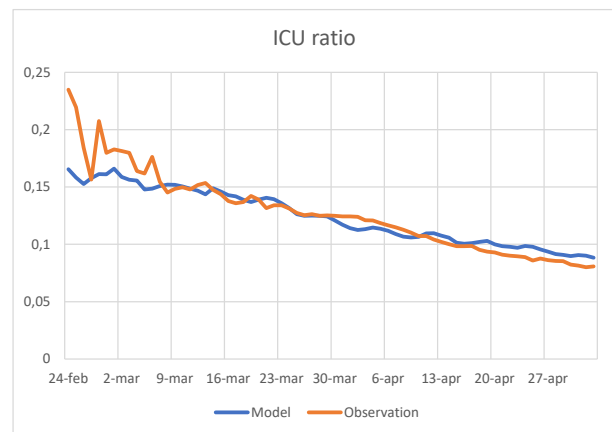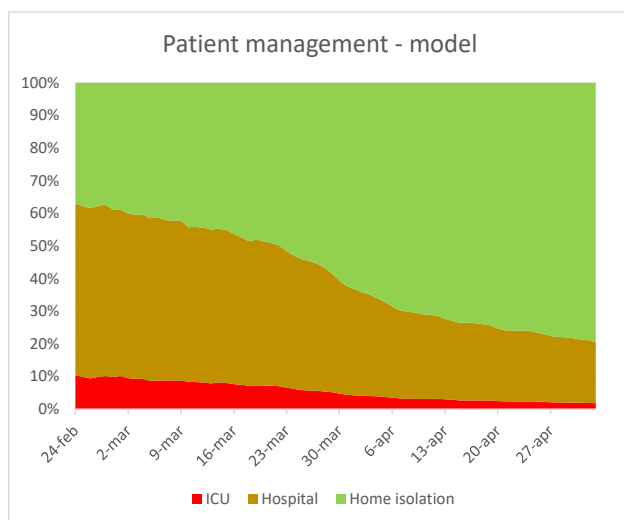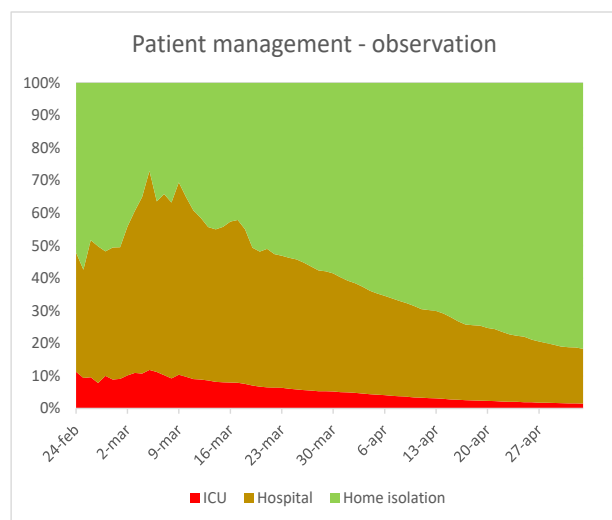

3g – Model implied infection management strategy

3h – Observed infection management strategy

Chart 3 – Full model (SIR submodel + multi state model) performance vs observations

With regard to the parameters, we point out that the following relationships should hold:

$$p_i^{Hm} < p_{i+1}^{Hm}, p_i^{Hs} > p_{i+1}^{Hs}, p_i^{ICU} = 1 - p_i^{Hm} - p_i^{Hs} > 1 - p_{i+1}^{Hm} - p_{i+1}^{Hs} > p_{i+1}^{ICU}$$

All these relations are consistent with the reasonable hypothesis that patients who survive for a greater number of days have a progressively more favorable prognosis.

The satisfactory agreement with the observed data and further tests carried out by the authors on trends at regional level, using models with the same structure, depose in favor of the reliability of the proposed approach: however, all the usual limitations and cautions in using mathematical models remain applicable in representing reality.

A substantial improvement in the analysis of the model's performance would occur using the comparison with the data relating to the actually observed clinical evolution of the patients: these data are not available while writing this paper.

4. Results

For the purpose of fitting, the patient management model is fed with the observed increments: to integrate the models it is sufficient to feed the model with the simulated increases of the R_1 population taken from the SIR model.

In the following sections, all the simulations are carried out on a total of 430 simulated days.

First of all, we note that the simulated evolution of the epidemic takes place essentially in the first 200 simulated days:
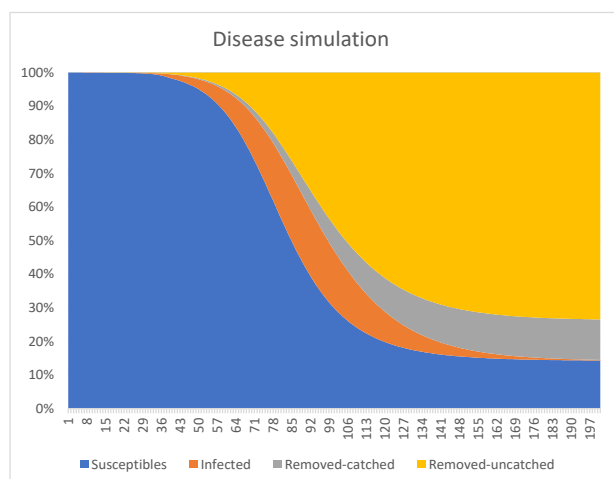
4a – Simulated evolution of epidemiologic quantities
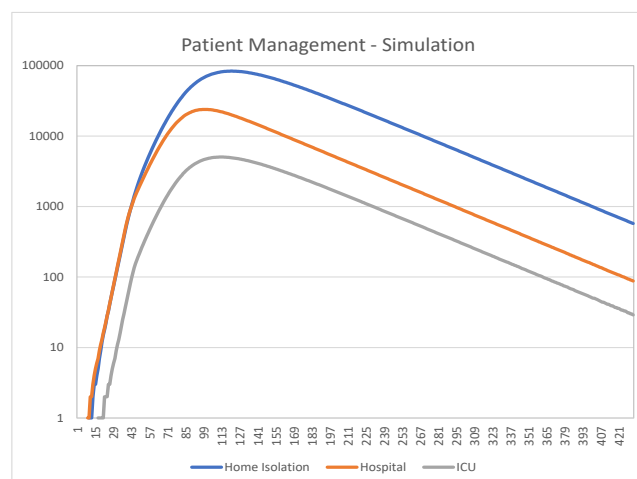
4b – Simulated evolution of the management strategy

Chart 4 – Simulation of the pandemic evolution

By applying the patient management model, the peaks in the number of patients in ICU, hospital and home isolation happen at day 116, 102 and 123 respectively (with a timespan, from the first to the last peak, of 21 days). As of the writing date of this article, the maximum of the classes of hospitalized patients have already been observed (Aprl 3th and 4th), while the maximum of the infected at home isolation has been observed on May 2nd with and observed delay of 29 days from the first peak.

It is also possible to calculate some relevant quantities, such as the number of deaths, the number of days in the various management classes, as well as the duration of the epidemic until reaching predetermined levels of the infected population (remember that infected people spread the disease).

The following values are obtained:

	430 days (simulation)	As at May 4th	%
Total deaths	46,011	29,079	63%
ICU days occupied	563,497	157,238	28%
Hospital days occupied	2,325,604	1,232,943	53%
Home isolation days occupied	9,676,022	2,978,830	31%
Time to 10,000 infected		181	days
Time to 1,000 infected		228	days
Time to 100 infected		274	days
Time to 10 infected		320	days

Table 4 – Simulation of the pandemic evolution implied by the model, key figures

The previous table also shows the progress compared to May 4th (which would be the 108th simulated day). Assuming a standard cost of 500 euros for a day of normal hospitalization and 1,500 euros for a day of hospitalization in intensive care, the direct costs of the epidemic arising from hospitalization should be around 2 billion euros:

Cost for 1 day and 1 person		
Hospital (no ICU)	500	euro
Hospital (ICU)	1,500	euro
Total cost		
Hospital (no ICU)	1,162,802,000	euro
Hospital (ICU)	845,245,500	euro
Hospital (tot)	2,008,047,500	euro

Table 5 – Simulated direct costs of the pandemic

It is important to notice that an estimation of the complete impact of the pandemic (in terms not only of costs, but also of GDP loss) can be achieved evaluating the average loss of a lost workday multiplied by the total number of days of persons in home isolation (in our simulation total quarantined days persons is 9,676,022).

Within the average loss it should keep into account the ability of different economic sectors, distinguished by Ateco classification, to leverage smart working, substitution of the workers and flexibility of production. In order to obtain the total impact of the pandemic, indirect costs such as the impact on consumption should also be minded, but their estimation is significantly harder requiring a more holistic approach to pandemic modeling.

A useful exercise is then to apply the model to different sets of initialization parameters, varying the "size" of the population N and the delay in activating the restrictive measures.

The following values are obtained:

Dimension	Delay	Deaths	Hospital days	ICU days	Isolation days	Duration to			
						10000 infected	1000 infected	100 infected	10 infected
10.000	15	198	2.988	12.307	52.724		86	137	183
100.000	15	2.267	28.303	116.672	485.957	111	162	209	256
1.000.000	15	23.004	281.531	1.162.254	4.833.101	188	234	280	327
5.000.000	15	114.997	1.405.601	5.805.454	24.126.683	237	283	331	377
10.000	30	208	3.094	12.728	54.556		71	114	157
100.000	30	2.279	28.449	117.275	488.585	93	143	189	235
1.000.000	30	23.034	282.056	1.164.088	4.843.180	168	215	263	309
5.000.000	30	115.084	1.407.892	5.812.470	24.171.071	220	265	313	359
10.000	45	227	3.290	13.523	57.965		67	103	139
100.000	45	2.391	29.877	123.203	513.544	80	121	162	202
1.000.000	45	23.281	285.298	1.177.216	4.899.544	148	194	239	285
5.000.000	45	115.431	1.412.816	5.830.999	24.259.517	200	246	292	339

Table 6 – Scenario simulations: pandemic key figures implied by
different dimension and lockdown delay hypothesis

The underlying idea of the exercise is to simulate the reopening of an epidemic event in our country (Italy), due to the emergence of an outbreak during a hypothetical phase of release of the restrictive measures. Assuming as the duration of the lockdown the time from the delayed detection of the outbreak to the return to the threshold of 1,000 infected people, the following relationship could be empirically found:

$$T_{lockdown} = T_{I=1000} - Delay = a\log(N) + b - Delay$$

with $a = 30.64$ and $b = -208.82$. It is reasonable to assume the existence of an exponential relationship between the infected population size N and the lockdown delay[7]: observing that N of the adopted model is equal to about 2 million people reached with a detection delay of 38 days, we can find the following relation:

$$N = \exp(k \times Delay)$$

where $k = 0.38$. By replacing this formula in the expression for the duration of the lockdown, we are able to find that each day of delay in the detection costs about 11 days of lengthening the lockdown and that, neglecting the effects of saturation, a detection delay of 47 days would be sufficient to expand the infected people until the total national population amount.

5. Conclusions

This paper focused on defining and applying a model for the management of the CoViD-19 epidemic. The phenomenon has been studied based on the Italian experience, with the purpose of identifying a preliminary estimate for the economic impacts of the disease. As previously described, the methodology we propose consists of two different independent models to be jointly used: the epidemic dynamics model (SIR) and a multi-state model for patient management.

The first model (SIR) seems to intercept pretty well the dynamics that characterize the contagion, especially in terms of basic reproduction number (r_0) and the number of people in whom the epidemic develops (N). As highlighted from the empirical evidence, captured by the model, it appears that the basic reproductive number of the epidemic r_0 for uncatched cases remains above 1 even during the lockdown. The authors believe that this phenomenon is essentially linked to the interaction between the fitting procedure and the nature of the lockdown: on one side the model is free to identify the dimension of the population within the disease spreads, and obviously concentrates on the fraction of population within active diffusion is detected, on the other the lockdown limited the size of the potentially exposed population through which the disease spread but was not effective in slowing down the diffusion within the confined population clusters (best known cases are the outbreaks in retirement homes RSA throughout Italy). It should also be noted that, in the model considered, asymptomatic infected patients, whose number seems to be high compared to symptomatic cases, took a relevant role in the spread dynamics.

The second model (multi-state) focused instead on the relationships that take into account the transition of patients through the possible duration of the different stages of the disease, represented respectively by patients recovered, deceased, in home isolation, hospitalized or in intensive care. By considering this information together with the estimates of the SIR model, it is possible to obtain a forecast of the duration in terms of days of lockdown to reach a residual threshold (e.g. 100 infected patients).

---

[7] The dynamic is consistent with the hypothesis that each simulated person has a constant number of contacts with new people for each simulated day. This hypothesis is reasonable in estimating a population only as long as saturation effects are not achieved.

We also proposed a scenario analysis related to the spread of the disease in smaller confined populations, and a preliminary cost measure of the "delay" of activation of the lockdown measures in terms of the duration of the lockdown itself: these analyzes, resulting from an extrapolation of the calibrated model (and therefore subject to the limits of precision and reliability of the same), could constitute a basis for the modeling of future outbreaks that may reoccur during the subsequent phase of relaxation of the current lockdown constraints.

However, an estimate of the direct costs of the epidemic is a partial, and in some ways secondary, aspect of the total economic impact: the loss of Italy's GDP caused by the lockdown, to be assessed with the necessary sector specificities, is to be considered the main effect of the epidemic. The authors believe that the proposed model, allowing to obtain indications on the duration of the lockdown, could also be useful for estimating indirect impacts, to which we will eventually dedicate future analyses.

Appendix – Technical details on the models

Fitting of the epidemiological model

The fitting of the epidemiological model is obtained starting from the data made available by the Civil Protection through the GitHub repository [23] that covers, as of the writing date of this article, the period from February 24th to May 4th 2020[8]. A first consideration is that the SIR model considered has a diffusion parameter for the pre-lockdown period and a different parameter for the post-lockdown period: the post-lockdown period, however, being made up of dates after March 8th 2020 included, has a data intensity much higher than the previous one (40 points versus 13), and therefore the estimates made for this second period are to be considered more stable. The SIR parameter estimation process is made of the following steps:

I. we define an auxiliary SIR model based on the equations $dS/dt = -\beta^* SI/N$; $dI/dt = +\beta^* SI/N - p_{catch}\gamma_1 I - (1 - p_{catch})\gamma_2 I$; $dR_1/dt = p_{catch}\gamma_1 I$; modeling of component R_2 at this level is ignored; the analysis compares the model with historical data for the lockdown period, or for data subsequent to the instant t^* assumed equal to March 8th. Furthermore, the parameters γ_1 and γ_2 are locked at 5.2 days and 15 days (in line with the literature indications for the time to onset of symptoms and for the evolution of the disease in shorter cases; time to onset of symptoms is used in the first case since the fraction of the population routed to the first class of the removed is intercepted and isolated at the onset of symptoms, while in the second case the infected remain infectious and asymptomatic for the entire duration of the disease), the initial state is: $S(0) = S_0; I(0) = I_0; R_1(0) = 7,535$, where I_0^* and S_0^* are both identified by the model;

II. we fit the SIR model based on equations $dS/dt = -\beta(1 - \delta H(t - t^*))(SI/N)$; $dI/dt = +\beta(1 - \delta H(t - t^*))SI/N - p_{catch}\gamma_1 I - (1 - p_{catch})\gamma_2 I$; $dR_1/dt = p_{catch}\gamma_1 I$; $dR_2/dt = (1 - p_{catch})\gamma_1 I$. Both the parameters γ_1 and γ_2 are locked as in the previous step, and the further constraint $\beta\delta = \beta^*$ is introduced in order to benefit from the greater amount of information of the second period. The parameter t^* is also left free so as to allow the model to infer the initial instant of evolution of the disease, conventionally placed at $S(0) = N - 10, I(0) = 10, R_1(0) = R_2(0) = 0$.

In both cases the dynamic used is a gradient descent. The fitting parameters of step I are used as the starting point for step II. The fitting of both steps is performed through the minimization of the L_1 norm, calculated on the total number of cases detected versus the time series of the simulated population $R_1(t)$. Defined $\widehat{R_1}(t_i)$ the time series of the cases observed by the Civil Protection, the fitting procedure seeks the minimum value of the norm $\sum_i |R_1(t_i) - \widehat{R_1}(t_i)|$. A commonly used approach is the maximization of the log-likelihood of the increases: it is assumed that the quantities $\Delta R_1(t_i) = R_1(t_{i+1}) - R_1(t_i)$ represent the expected values λ_i of a random Poisson variable whose corresponding observation is $\Delta \widehat{R_1}(t_i) = \widehat{R_1}(t_{i+1}) - \widehat{R_1}(t_i) = p_i$ and then the fitting procedure look for the minimum of the norm $\sum_i [\lambda_i - p_i \log(\lambda_i) + \log(p_i!)]$. In principle, different combinations of the two approaches are possible by performing steps I and II for one norm or the other:

Parameter	Norm L_1	Norm $LLk_{Poisson}$
S_0	1,015,782	86,725,604
I_0	134,116	11,656,148
β^*	0,1663	0,1921
p_{catch}	9,81%	0,00%

Table Appedix.1 – Fitted parameters of auxiliary
SIR submodels obtained with different norms

---

[8] An early version of this paper appeared on AIFIRM website, https://www.aifirm.it/news/da-leggere/, in which models were fitted using data from February 24th to April 16th.

Using the obtained parameters as:

Parameter	I. Norm L_1 II. Norm L_1	I. Norm L_1 II. Norm $LLk_{Poisson}$	I. Norm $LLk_{Poisson}$ II. Norm L_1	I. Norm $LLk_{Poisson}$ II. Norm $LLk_{Poisson}$
t^*	38	51	33	20
N	1,995,898	60,000,000 (active limit)	24,802,511	56,694,585
δ	60,0%	68,0%	47,5%	39,6%
p_{catch}	5,38%	0,15%	0,38%	0,15%

Table Appendix.2 – Fitted parameters of SIR submodel obtained with different combinations of norms

The authors opted for the strategy Norm L_1+ Norm L_1 both for consistency between the two steps and for the reasonability on the dimension of the population in which the disease spread, that, looking at the death toll, should not be in the tens of millions range.

Fitting the multi-state model

For the state model, the strategy is to minimize the norm L_1:

$$\sum_i \left|\Delta D_i - \Delta\widehat{D}_i\right| + \left|\Delta G_i - \Delta\widehat{G}_i\right| + w\left(\left|\Delta hr_i - \Delta\widehat{hr}_i\right| + \left|\Delta icr_i - \Delta\widehat{icr}_i\right|\right)$$

where ΔD_i and $\Delta\widehat{D}_i$ respectively represent the expected and observed daily deaths, the values ΔG_i and $\Delta\widehat{G}_i$ the expected and observed daily healings, the values Δhr_i and $\Delta\widehat{hr}_i$ the hospital ratio expected and observed, Δicr_i and $\Delta\widehat{icr}_i$ the planned and observed intensive care ratio, and w a weight that has the purpose of connecting the scales of the variables (assumed equal to 500). The hospital ratio is defined as the ratio between the patients hospitalized against the total number of infected people (for each moment of time) and the intensive care ratio is defined as the ratio between the patients admitted to intensive care against the patients hospitalized (both at internal and outside of intensive care). The model is powered by the observed inputs $\widehat{I}_i = \widehat{R_1}(t_{i+1}) - \widehat{R_1}(t_i) - \left(\widehat{G}(t_{i+1}) - \widehat{G}(t_i)\right) - \left(\widehat{D}(t_{i+1}) - \widehat{D}(t_i)\right)$ and initialized by dividing the 229 cases detected at February 24th into 5 classes (4 of 46 cases and one of 45 cases) distributed over the first 5 days of population development. The norm is calculated on the 71 increments observed from February 24th to May 4th, 2020.

Multi-state model implementation code

This section shows the pseudocode that implements the multi-state model: the objective is an operational description of the model, which allows its possible replication. It is assumed that the outputs of the SIR epidemiological model are accumulated in the vectors R1[t], R2[t]. The population state at every moment of time is accumulated in the array of vectors ActiveCatched[t][i], where the vectors ActiveCatched[t] are supposed to be of variable length.

```
ActiveCatched[1]=[0]
For any t in 2…length(R1[t])

      EnteringR=R1[t]-R1[t-1]

      ActiveCatched[2..N+1]=ActiveCatched[1..N]

      AcriveCatched[1]=EnteringR

      Base1=sum(ActiveCatched[t-1][1…T1])

      Base2=sum(ActiveCatched[t-1][T1…T2])

      Base3=sum(ActiveCatched[t-1][T2…])

      ActiveCatched[t]=FIFOOut(ActiveCatched[t],(p_{D,M3}+p_{H,M3})*Base3)

      ActiveCatched[t]=FIFOOut(ActiveCatched[t],(p_{D,M2}+p_{H,M2})**Base2,T2)

      ActiveCatched[t]=FIFOOut(ActiveCatched[t], p_{D,M1} *Base1,T1)

      DeathEvents[t]=(Base1*p_{D,M1}+Base2*p_{D,M2}+Base3*p_{D,M3})

      HealEvents[t]= (Base2*p_{H,M2}+Base3*p_{H,M3})
```

The FIFO out function eliminates people from the active case vector by subtracting them from those of older age:

```
FIFOOut(vector,events,index)
        If index not specified
                Index=length(vector)
        Aux=events
        For i=index to 1
                If vector[i]<aux
                        Aux=aux-vector[i]
                        Vector[i]=0
                Else
                        Vector[i]=vector[i]
                        Aux=0
        If Aux>0 throw exception("Too many events!")
```

The execution of the previous code segment defines the vectors `DeathEvents` and `HealEvents`, representing deaths and daily healings respectively. To obtain the daily consistency of the populations in home isolation, hospitalized or in intensive care, the following pseudocode is performed:

```
for any t in 1…length(ActiveCatched)
        Base1=sum(ActiveCatched(1…T1))
        Base2=sum(ActiveCatched(T1…T2))
        Base3=sum(ActiveCatched(T2…))
        PopulationVector=[Base1,Base2,Base3]
        HomeIsolated[t]=PopulationVector•Home_V
        Hospital[t]=PopulationVector•Hosp_V
        ICU[t]=PopulationVector•ICU_V
```

where `Home_V`=$\begin{pmatrix} p_1^{Hm} \\ p_2^{Hm} \\ p_3^{Hm} \end{pmatrix}$, `Hosp_V`=$\begin{pmatrix} p_1^{Hs} \\ p_2^{Hs} \\ p_3^{Hs} \end{pmatrix}$, `ICU_V`=$\begin{pmatrix} 1 - p_1^{Hm} - p_1^{Hs} \\ 1 - p_2^{Hm} - p_2^{Hs} \\ 1 - p_3^{Hm} - p_3^{Hs} \end{pmatrix} = \begin{pmatrix} p_1^{ICU} \\ p_2^{ICU} \\ p_3^{ICU} \end{pmatrix}$.

`HomeIsolated`, `Hospital` and `ICU` are vectors that contain, at any time, the population subjected to home isolation, hospitalization and intensive care treatment respectively.

Fabio Verachi, Luca Trussoni e Luciano Lanzi

Bibliography

[1] A J Kucharski, T W Russell, C Diamond, Y Liu, J Edmunds, S Funk, R M Eggo, F Sun, M Jit, J D Munday, N Davies, A Gimma, K van Zandvoort, H Gibbs, J Hellewell, C I Jarvis, S Clifford, B J Quilty, N I Bosse, S Abbott, P Klepac, S Flasche. Early dynamics of transmission and control of COVID-19: a mathematical modelling study, The Lancet Infectious Diseases, Vol 20 (5) 2020, 553-558, doi: 10.1016/S1473-3099(20)30144-4.
[2] Qun Li, Xuhua Guan, Peng Wu, Xiaoye Wang, Lei Zhou, Yeqing Tong, Ruiqi Ren, Kathy S M Leung, Eric H Y Lau, Jessica Y Wong, Xuesen Xing, Nijuan Xiang, Yang Wu, Chao Li, Qi Chen, Dan Li, Tian Liu, Jing Zhao, Man Liu, Wenxiao Tu, Chuding Chen, Lianmei Jin, Rui Yang, Qi Wang, Suhua Zhou, Rui Wang, Hui Liu, Yinbo Luo, Yuan Liu, Ge Shao, Huan Li, Zhongfa Tao, Yang Yang, Zhiqiang Deng, Boxi Liu, Zhitao Ma, Yanping Zhang, Guoqing Shi, Tommy T Y Lam, Joseph T Wu, George F Gao, Benjamin J Cowling, Bo Yang, Gabriel M Leung, Zijian Feng. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia, New England Journal of Medicine 2020 Mar 26;382(13):1199-1207. doi: 10.1056/NEJMoa2001316. [3] J T Wu, K Leung, G M Leung, Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study, The Lancet, Vol 395 (10225) 2020, 689-697, doi:10.1016/S0140-6736(20)30260-9.

[4] World Health Organization, Statement on the meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV), https://www.who.int/news-room/detail/23-01-2020-statement-on-the-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov).

[5] World Health Organization, Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19), Feb 2020, https://www.who.int/publications/i/item/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-(covid-19).

[6] Tao Liu, Jianxiong Hu, Jianpeng Xiao, Guanhao He, Min Kang, Zuhua Rong, Lifeng Lin, Haojie Zhong, Qiong Huang, Aiping Deng, Weilin Zeng, Xiaohua Tan, Siqing Zeng, Zhihua Zhu, Jiansen Li, Dexin Gong, Donghua Wan, Shaowei Chen, Lingchuan Guo, Yan Li, Limei Sun, Wenjia Liang, Tie Song, Jianfeng He, Wenjun Ma. Time-varying transmission dynamics of Novel Coronavirus Pneumonia in China, bioXriv, Feb 2020, https://doi.org/10.1101/2020.01.25.919787.

[7] Rocklöv J, Sjödin H, Wilder-Smith A. COVID-19 outbreak on the Diamond Princess cruise ship: estimating the epidemic potential and effectiveness of public health countermeasures. J Travel Med. 2020;27(3):taaa030. doi:10.1093/jtm/taaa030.

[8] Backer Jantien A, Klinkenberg Don, Wallinga Jacco. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. Euro Surveill. 2020;25(5):pii=2000062. https://doi.org/10.2807/1560-7917.ES.2020.25.5.2000062.

[9] Jonathan M Read, Jessica RE Bridgen, Derek AT Cummings, Antonia Ho, Chris P Jewell. Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions, medXriv Jan 2020, https://doi.org/10.1101/2020.01.23.20018549.

[10] Qifang Bi, Yongsheng Wu, Shujiang Mei, Chenfei Ye, Xuan Zou, Zhen Zhang, Xiaojian Liu, Lan Wei, Shaun A Truelove, Tong Zhang, Wei Gao, Cong Cheng, Xiujuan Tang, Xiaoliang Wu, Yu Wu, Binbin Sun, Suli Huang, Yu Sun, Juncen Zhang, Ting Ma, Justin Lessler, Teijian Feng. Epidemiology and Transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts, medXriv Mar 2020, https://doi.org/10.1101/2020.03.03.20028423.

[11] Tang B, Wang X, Li Q, Bragazzi N.L, Tang S, Xiao Y, Wu J. Estimation of the Transmission Risk of the 2019-nCoV and Its Implication for Public Health Interventions. J. Clin. Med. 2020, 9, 462.

[12] O. Diekmann, H. Heesterbeek and T. Britton. Mathematical Tools for Understanding Infectious Disease Dynamics. Princeton Series in Theoretical and Computational Biology. Princeton University Press, Princeton, (2013).

[13] Hethcote H W. The mathematics of infectious diseases. SIAM Rev, 2000, 42: 599–653.

[14] Keeling, M., & Rohani, P. (2008). Modeling Infectious Diseases in Humans and Animals. PRINCETON; OXFORD: Princeton University Press. doi:10.2307/j.ctvcm4gk0.

[15] Milan Batista. Estimation of the final size of the coronavirus epidemic by the logistic model, medXriv Feb 2020, https://doi.org/10.1101/2020.02.16.20023606.

[16] Zhou T, Liu Q, Yang Z, et al. Preliminary prediction of the basic reproduction number of the Wuhan novel coronavirus 2019-nCoV. J Evid Based Med. 2020; 13: 3– 7. https://doi.org/10.1111/jebm.12376.

[17] Zhao S, Lin Q, Ran J, et al. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. Int J Infect Dis. 2020;92:214-217. doi:10.1016/j.ijid.2020.01.050.

[18] A Agosto, P Giudici, A Poisson Autoregressive Model to Understand COVID-19 Contagion Dynamics (March 9, 2020). http://dx.doi.org/10.2139/ssrn.3551626.

[19] Hansen E, Day T. Optimal control of epidemics with limited resources. J. Math. Biol. 62, 423–451 (2011). https://doi.org/10.1007/s00285-010-0341-0.

[20] Yang, Zifeng, Zhiqi Zeng, Ke Wang, Sook-San Wong, Wenhua Liang, Mark Zanin, Peng Liu, Xudong Cao, Zhongqiang Gao, Zhitong Mai, Jingyi Liang, Xiaoqing Liu, Shiyue Li, Yimin Li, Feng Ye, Weijie Guan, Yifan Yang, Fei Li, Shengmei Luo, Yuqi Xie, Bin Liu, Zhoulang Wang, Shaobo Zhang, Yaonan Wang, Nanshan Zhong, & Jianxing He. "Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions." Journal of Thoracic Disease [Online], 12.3 (2020): 165-174. Web. 9 Jul. 2020.

[21] Milan Batista. Estimation of the final size of the coronavirus by the SIR model, Feb 2020, https://www.researchgate.net/publication/339240777_Estimation_of_the_final_size_of_coronavirus_epidemic_by_the_logistic_model.

[22] I. Nesteruk, "Statistics-based predictions of coronavirus epidemic spreading in Mainland China," Innovative Biosystems and Bioengineering, vol. 4, no. 1, pp. 13–18, 2020.

[23] https://github.com/pcm-dpc/COVID-19/blob/master/dati-andamento-nazionale/dpc-covid19-ita-andamento-nazionale.csv.

[24] Day Michael. Covid-19: four fifths of cases are asymptomatic, China figures indicate BMJ 2020; 369 :m1375.

[25] Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science. 2020;368(6490):489-493. doi:10.1126/science.abb3221.

[26] https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-COVID-19_23-aprile-2020.pdf.

[27] Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention [published online ahead of print, 2020 Feb 24]. JAMA. 2020;10.1001/jama.2020.2648. doi:10.1001/jama.2020.2648.

[28] L.J.S. Allen, P. van den Driessche. Relations between deterministic and stochastic thresholds for disease extinction in continuous- and discrete-time infectious disease models, Mathematical Biosciences, Vol 243(1) 2013, 99-108, doi: 10.1016/j.mbs.2013.02.006.

RISK MANAGEMENT MAGAZINE

AIFIRM

In collaborazione con REPLY AVANTAGE

IN QUESTO NUMERO

Rivista scientifica riconosciuta da ANVUR e AIDEA

anvur — ACCADEMIA ITALIANA DI ECONOMIA AZIENDALE