

# First month of the epidemic caused by COVID-19 in Italy: current status and real-time outbreak development forecast

Rosario Megna\*, PhD

Institute of Biostructure and Bioimaging, National Council of Research, Naples, Italy

## Abstract

**Background:** The first outbreaks of COVID-19 in Italy occurred during the second half of February 2020 in some areas in the North of the country. Due to the high contagiousness of the infection, further spread by asymptomatic people, Italy has become in a few weeks the country with the greatest number of infected people in the world. The large number of severe cases among infected people in Italy led to the hospitalization of thousands of patients, with a heavy burden on the National Health Service.

**Methods:** We analyzed data provided daily by Italian Authorities for the period from 24 February 2020 to 30 March 2020. Considering such information, we developed a forecast model in real-time, based on the cumulative log-logistic distribution.

**Results:** A total of 101,739 infected individuals were confirmed until 30 March 2020, of which 14,620 recovered or discharged, and 11,591 deaths. Until the same date patients quarantined at home were 43,752, whereas hospitalized patients were 31,776, of which 3,981 in intensive care. The active cases (i.e. the number of patients not yet recovered until that date) were 75,528. The forecast model

estimated a number of infected persons for Italy of 234,000 about, and a duration of the epidemic of approximately 4 months.

**Conclusions:** One month after the first outbreaks there seemed to be the first signs of a decrease in the number of infections, showing that we could be now facing the descending phase of the epidemic. The forecast obtained thanks to our model could be used by decision-makers to implement coordinative and collaborative efforts in order to control the epidemic. The pandemic due to novel Coronavirus must be a warning for all countries worldwide, regarding a rapid and complete dissemination of information, surveillance, health organization, and cooperation among the states.

**Keywords:** COVID-19, SARS-CoV-2, outbreak, forecast model, forecast in real-time, epidemiology

\* Correspondence: [rosario.megna@ibb.cnr.it](mailto:rosario.megna@ibb.cnr.it)

## Introduction

The outbreak of the novel Coronavirus called COVID-19 (or SARS-CoV-2), which had the first diagnosed cases at the end of December 2019 in the Hubei Province, China [1], is having a dramatic global evolution, and was recently classified as a pandemic by the World Health Organization (WHO) on 11 March 2020 [2]. The disease, which can be diagnosed through the use of a nasopharyngeal swab, under the most severe forms can lead to bilateral pneumonia [3] which can be lethal especially in elderly patients with comorbidities [4].

The first outbreaks of COVID-19 in Italy occurred during the second half of February 2020 in some areas in the North of the country. First cases were diagnosed in southern Lombardy on February 21, on the border of the Veneto and Emilia Romagna regions [5-7]. On February 23, 11 municipalities were quarantined: nobody could enter and leave those territories (DL n. 6, 23 February 2020) [8]. Quickly other outbreaks occurred in the North of the country, requiring a wider extension of the area of limited human activities to various northern regions including Lombardy, Emilia Romagna, and Veneto (DPCM of 4 March 2020) [9]. Despite the drastic restrictions imposed by the Italian Government in those areas, several other outbreaks began in other areas of northern Italy, forcing the Authorities to extend the previously adopted restrictions to the entire national territory (DL of 9 March 2020, and DPCM of 22 March 2020) [10, 11]. Due to the high contagiousness of the infection, further spread by asymptomatic people [4, 12], in few weeks Italy has become the country with the greatest number of infected people after China (confirmed cases greater than 80,000 from 26 March 2020). The spread of this virus has been going on so fast that for a few days now the primacy of the country with the most infected people is the US [13]. The large number of severe cases among infected people in Italy led to the hospitalization of thousands of patients [14, 15], with a heavy burden on the National Health Service [7]. In particular, the most affected regions are Lombardy and Emilia Romagna, with more than half of the total cases. It is reasonable to assume

that the large spread of the novel Coronavirus in these regions was due to the development of the first outbreaks which caused a high number of people infected before of the social distancing imposed by Government. One month after the beginning of the epidemic in Italy, we report the current status and propose a forecast model in real-time to estimate its evolution in terms of epidemic duration and potential number of infected persons. These informations could be applied in surveillance to inform clinicians and decision-makers to take coordinative and collaborative efforts to control the pandemic.

## Methods

Data on COVID-19 used in our analysis are daily updates from the Italian Ministry of Health managed by the Civil Protection Department [14, 15]. A report is released at 5:00 or 6:00 pm (CET), on the basis of information provided by national and Regional Local Authorities. The most relevant variable is the number of confirmed cases. The other derivate variables to be considered are the number of hospitalized patients (in intensive or non-intensive care), individuals quarantined at home, patients who recovered or were discharged, and the number total deaths.

We analyzed data used in this study using the R software, version 3.6.3 (R Foundation for Statistical Computing, Vienna, Austria). Continuous variables were expressed as mean  $\pm$  standard deviation and categorical variables as percentages, while differences between groups were evaluated by  $\chi^2$  test for proportions.

We developed a forecast model in real-time, based on the cumulative log-logistic distribution [16, 17]. The equation used is the following:

$$C(t) = \frac{N}{1 + a \cdot \exp(b \cdot \log(t - t_0))} \quad (1)$$

where  $C(t)$  is the cumulative incidence on day  $t$ ,  $N$  is the cumulative incidence at the end of the epidemic,  $a$  and  $b$  are the parameters that govern the symmetry and the growth rate of the curve, and  $t_0$  is the initial day of the epidemic analysis (24 Feb 2020). In order to determine the three parameters of the curve, we developed an algorithm based on the maximum statistical significance of  $a$  and  $b$ , according to their P-values (smaller values of P indicate greater significance), and varying  $N$ . Nonlinear least squares (nls) function of R was used, normalizing  $C(t) \rightarrow C(t, N = 1) \equiv C$ . The steps of the algorithm are the following:

```

for (  $n$  from  $min\_n$  to  $max\_n$ , step  $delta\_n$  )
    computes  $nls( C( a( n ), b( n ) ) )$ 
    if (  $p\text{-value}( a( n ) ) < p\text{-value}( a( n-1 ) )$  ) and (  $p\text{-value}( b( n ) ) < p\text{-value}( b( n-1 ) )$  )
    then
        continue
    else
         $a = a( n-1 ); b = b( n-1 ); N = n-1$ 

```

with  $min\_n = 30,000$ ,  $max\_n = 300,000$ , and  $delta\_n = 1,000$  for the National evaluation and  $min\_n = 10,000$ ,  $max\_n = 100,000$ , and  $delta\_n = 500$  for the Regional evaluations (data used for the national model and computational R code are reported in the supplementary materials). At the end of the last cycle, the algorithm provides parameters for best-fit. In order to assess the 95% confidence interval (CI) of the fit values, 1,000 bootstrap resamplings were computed, through the IPEC package of R. The date of the epidemic peak was computed by the maximum of the first derivative of  $C(t)$ . Due to the asymptotic pattern of  $C(t)$ , the date of the epidemic end was computed as the day with

less than 100 new cases for the National evaluation, and less than 50 new cases for the Regional evaluations. The graphics were obtained using the ggplot2 package of R.

## Results

A total of 101,739 infected individuals were confirmed until 30 March 2020, of which 14,620 recovered or discharged, and 11,591 deaths. Until the same date patients quarantined at home were 43,752, whereas hospitalized patients were 31,776, of which 3,981 in intensive care. The active cases (i.e. the number of patients not yet recovered to the date) at the time were 75,528. Figure 1 shows the daily distribution of performed swabs and confirmed cases. The numbers over the bars are related to the ratio, in percentage, between the two variables ( $8.7\% \pm 7.7\%$ ). Until 30 March 2020, the total number of performed swabs is 477,359. In Figure 2 we report the cumulative distributions of the confirmed cases and patient categories. During the first month of the epidemic, the confirmed cases increased of almost 3 orders of magnitude, the trends of patients quarantined at home and in non-intensive care remained very similar among them, whereas from mid-March the deceased persons (greater than 1,000 per day) exceed the patients in intensive care. In Table 1 is reported the number of confirmed cases and the patient categories for at a National and Regional (Lombardy and Emilia Romagna) level. The confirmed cases to date were greater than 100,000, the hospitalized patients are greater than 40,000 as well the persons quarantined at home, while the number of deceased patients is significantly greater than the Chinese [13, 18] ( $P < 0.001$ ).

In Table 2 we summarize the forecasted method results for at a National and Regional (Lombardy and Emilia Romagna) level. The model predicts a number of the infected persons of 230,000 in Italy, of 90,500 in Lombardy, and of 44,000 in Emilia Romagna. The duration of the epidemic is estimated

to be approximately of 4 months. In figure 3 we depict the cumulative log-logistic curve obtained by the forecast model in real-time for the National overview.

**Table 1** Confirmed cases and patient categories updated to 30 March 2020

|                                    | Italy         | Lombardy      | Emilia Romagna |
|------------------------------------|---------------|---------------|----------------|
| Confirmed cases                    | 101,739       | 42,161        | 13,531         |
| Quarantined at home                | 43,752 (43.0) | 11,861 (28.1) | 6,636 (49.0)   |
| Hospitalized in non-intensive care | 27,795 (27.3) | 11,815 (28.0) | 3,779 (27.9)   |
| Recovered or discharged            | 14,620 (14.4) | 10,337 (24.5) | 1,227 (9.1)    |
| Deceased                           | 11,591 (11.4) | 6,818 (16.2)  | 1,538 (11.4)   |
| Hospitalized in intensive care     | 3,981 (3.9)   | 1,330 (3.2)   | 351 (2.6)      |

Note: Between brackets we report the ratio, in percentage, between each patient category and the confirmed cases

**Table 2** Best-fit parameters and relevant dates obtained by the forecasted model for the epidemic caused by COVID-19 in Italy

|                             |      | Italy                        | Lombardy                  | Emilia Romagna            |
|-----------------------------|------|------------------------------|---------------------------|---------------------------|
| Best-fit parameters         | A    | $3.67 \cdot 10^5$ *          | $7.29 \cdot 10^4$ *       | $2.71 \cdot 10^5$ *       |
|                             | B    | -3.51 *                      | -3.10 *                   | -3.28 *                   |
|                             | N #  | 234,000<br>(214,000-251,000) | 90,500<br>(82,500-97,500) | 44,000<br>(38,500-51,000) |
| Relevant dates <sup>§</sup> | Peak | 27 March (33)                | 24 March (30)             | 1 April (38)              |
|                             | End  | 28 June (126)                | 27 June (125)             | 19 June (117)             |

Notes:

\*  $P < 10^{-5}$

# Between brackets we report 95% CI

§ Between brackets we report the number of days since the outbreak started

## Discussion

One month after the outbreak in Italy the situation remains complicated. Despite the high number of performed swabs as compared to the confirmed cases, the epidemic has been growing with a very high rate. COVID-19 is proving to have a high capacity for infection, probably reinforced by asymptomatic people, who represent a real danger for elderly and fragile individuals. In particular, the disease is showing to be lethal for the elderly (95.2% in patients aged  $\geq 60$ ) and men (70.8%) [18]. On the date we finalized this article (30 March 2020), the trend of daily distribution of confirmed cases seems to show an initial decline of the growth of the epidemic. However, the total number of confirmed cases already exceeds those that occurred in China. In addition to this, many individuals died without the possibility of checking if they were actually infected and therefore not recorded as such [19, 20]. The data related to the patient categories give us an estimate of the epidemic in terms of cases that can be treated at home, those who need hospitalization, and the mortality. The implications in terms of public health, workload in hospitals, and economic damage were worrisome. The hospital beds, in particular in intensive care, were saturated in several areas of Northern Italy [21-23]. While the lockdown imposed from the Government to limit the infections has stopped almost all commercial activities in the whole country. In general, through the Italian epidemiological findings, countries with similar characteristics to those present in Italy (demographic characteristics of the population, health structures, etc.), should take earlier restrictive measures and arrange the necessary treatments for potential patients.

The forecast model in real-time indicates a total number of national cases greater than 230,000 patients, with a figure of approximately 90,000 in Lombardy only. In addition, the model estimates the duration of the epidemic in approximately 4 months. Since the theoretical cumulative curve has an asymptotic pattern (i.e. the maximum value is achieved for the  $t$  time towards infinite), considering 100 new cases in a day as the end of the outbreak is a convention. Therefore, if instead

of 100 new cases we considered 50 new cases, the overall number of days estimated for the epidemic to come to an end would increase of approximately 21 days, whereas considering 200 new cases the duration of the outbreak would decrease of about 19 days. Moreover, several factors could affect the total number of cases and the duration of the epidemic. For example, a contribution to the spread of outbreaks in southern Italy was caused by the movement of students and workers from Northern to Southern Italy following the first governmental restrictions. On the other hand, more stringent restrictions imposed later on by the Government could lower the expected number of total cases and reduce the number of days towards the end of the epidemic. On this specific topic, a previous study on SARS-CoV-2 in China found a nonlinear and chaotic behavior of the virus, which emerged gradually but was highly responsive to massive interventions [24]. Another important factor is related to possible mutations of the novel Coronavirus [25], which could have a positive or negative outcome on the trend of the pandemic.

Aware of what is happening in Italy, the other countries of the European Union should adopt agreed measures regarding health and economic aids, and also regulate uniformly the movement of people among the member States, to avoid a new spread of the SARS-CoV-2. In fact, the pandemic due to novel Coronavirus is the most widespread in the globalization era, and the lesson about what is happening must be a warning for all countries worldwide. A rapid and thorough dissemination of information, surveillance, optimization of health systems, and cooperation among states is needed in order to reduce contagion and economic damages.

It is also necessary to consider the intrinsic limitations of this study. First of all, data was not always updated on a daily basis by each Regional Authority (an extract of the warnings list provided from Civil Protection is reported in the supplementary materials, Table SM 1). This limitation can have effects on the trend of the epidemiological curve, therefore on the fit of the data. Another limitation is represented of the reported cumulative counts, that are known be under-reported, especially at

the beginning of the pandemic due to public awareness. If the counts are under-reported in the beginning of pandemic, all reported accumulated counts would be all under-estimated. We also have to consider that the number of infected people is underestimated, since there are many undetected asymptomatic individuals. These individuals can unknowingly infect several other persons contributing to the spread of the epidemic. More specifically, in Italy 5.9% of individuals who had a check through a swab were diagnosed as asymptomatic and 12.9% were considered people with non-specific symptoms [18]. These percentages could to be underestimated, because the majority of the population did not take a swab. Finally, the factors that determine the trend of the epidemic curve could change without respecting the pattern of the forecasted model. In fact, the populations' growth is exponential-like at the beginning (as also verified in [26]), but toward the end it flattens due to saturation. Likewise, the tail-end of the log-logistic curve will be governed by the quarantined population and the consequent social distancing. As a matter of fact, the end of an epidemic can happen either thanks to the immunization of individuals affected (herd immunity) or thanks to the extinction of the virus. Since the overall number of infected citizens did not reach the majority of the population (approximately 60 million), the end of the epidemic is expected to occur with the extinction of the virus.

## **Conclusions**

The epidemic caused by COVID-19 in Italy is having a dramatic evolution in terms of confirmed cases, hospitalized and deceased patients. After one month since the first outbreaks, first signs of a decrease in the number of infections became apparent, showing the possibility of a descending phase of the epidemic. The model presented in this article fits well with the data, therefore it is expected to be reliable in predicting the evolution of the epidemic within the limits discussed. The forecast could be applied by decision-makers to take coordinative and collaborative efforts to

control the epidemic. The pandemic due to novel Coronavirus is characterized by a fast spread worldwide, with dramatic repercussions on the health of the population and the economy.

## List of the figures

**Figure 1** Daily distribution of the performed swabs and confirmed cases. The numbers are related to the ratio, in percentage, between confirmed cases and performed swabs

**Figure 2** Cumulative distribution of the confirmed cases and patient categories

**Figure 3** Cumulative curve obtained by the forecasted model of the epidemic in Italy in real-time. Predictions are represented by the red line, with the gray area to indicate 95% CI. The blue points represent the confirmed cases

## **Declarations**

### **Ethics approval and consent to participate**

Not applicable.

### **Consent for publication**

Not applicable.

### **Availability of data and material**

The data that support the findings of this study are available from

<http://opendatadpc.maps.arcgis.com/apps/opsdashboard/index.html#/b0c68bce2cce478eaac82fe38d4138b1>

### **Competing interests**

The author declare that he has no competing interests.

### **Funding**

Not applicable.

### **Authors' contributions**

The article has been realized by RM.

### **Acknowledgements**

Not applicable.

### **Authors' information (optional)**

Rosario Megna has a Master's Degree and a PhD in Physics. He works in data analysis and modelling, machine learning, and imaging.

## References

1. Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., & Feng, Z. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *New England Journal of Medicine*. <https://doi.org/10.1056/NEJMoa2001316>
2. World Health Organization. <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>
3. Lee EYP, Ng MY, Khong PL. COVID-19 pneumonia: what has CT taught us? *Lancet Infect Dis*. 2020 Feb 24. pii: S1473-3099(20)30134-1. doi:10.1016/S1473-3099(20)30134-1.
4. Leung C. Clinical features of deaths in the novel coronavirus epidemic in China. *Rev Med Virol*. 2020 Mar 16:e2103. doi: 10.1002/rmv.2103.
5. Porcheddu R, Serra C, Kelvin D, Kelvin N, Rubino S. Similarity in Case Fatality Rates (CFR) of COVID-19/SARS-COV-2 in Italy and China. *J Infect Dev Ctries*. 2020 Feb 29;14(2):125-128. doi: 10.3855/jidc.1260.
6. La Marca A, Niederberger C, Pellicer A, Nelson SM. COVID-19: lessons from the Italian reproductive medical experience. *Fertil Steril*. 2020 Mar 25. pii:S0015-0282(20)30297-1. doi: 10.1016/j.fertnstert.2020.03.021. [Epub ahead of print].
7. Rosenbaum L. Facing Covid-19 in Italy - Ethics, Logistics, and Therapeutics on the Epidemic's Front Line. *N Engl J Med*. 2020 Mar 18. doi: 10.1056/NEJMp2005492.
8. Law Decree n. 6, issued on 23 February 2020. Available from the official source of knowledge of the rules in force in Italy. <https://www.gazzettaufficiale.it/eli/id/2020/02/23/20G00020/sg>
9. Decree of the President of the Council of Ministers issued on 4 March 2020. <http://www.governo.it/it/articolo/coronavirus-firmato-il-dpcm-4-marzo-2020/14241>
10. Law Decree n. 9, issued on 9 March 2020. Available from the official source of knowledge of the rules in force in Italy. <https://www.gazzettaufficiale.it/eli/id/2020/03/09/20G00030/sg>

11. Decree of the President of the Council of Ministers issued on 22 March 2020.  
<http://www.governo.it/it/articolo/coronavirus-firmato-il-dpcm-22-marzo-2020/14363>
12. Mizumoto K, Kagaya K, Zarebski A, Chowell G. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. Euro Surveill. 2020 Mar; 25(10). doi: 10.2807/1560-7917.ES.2020.25.10.2000180.
13. Johns Hopkins Center for System Science and Engineering.  
<https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6>
14. Italian Ministry of Health (Ministero della Salute). <http://www.salute.gov.it/portale/home.html>
15. Italian Department of Civil Protection (Protezione Civile).  
<http://opendatadpc.maps.arcgis.com/apps/opsdashboard/index.html#/b0c68bce2cce478eaac82fe38d4138b1>
16. Turner MEJ, Bradley ELJ, Kirk KA, Pruitt MK. A theory of growth. Mathematical Biosciences, 29 (3–4) (1976), pp. 367-373.
17. Tadikamalla PR. A look at the Burr and related distributions. Inter. Statist. Rev., 48 (1980), pp. 337-344.
18. Higher Institute of Health (Istituto Superiore di Sanità). Press releases, 20 and 27 March 2020.  
<https://www.iss.it/comunicati-stampa>
19. “Il sole 24 ore”, newspaper of national importance.  
<https://www.infodata.ilsole24ore.com/2020/03/27/43630/>
20. “Il fatto quotidiano”, newspaper of national importance.  
<https://www.ilfattoquotidiano.it/2020/03/19/coronavirus-perche-cosi-tanti-morti-in-italia-ce-una-risposta-ottimistica-e-altre-meno/5742749/>

21. "La Repubblica", newspaper of national importance.  
[https://www.repubblica.it/economia/2020/03/15/news/straordinari\\_maggiorati\\_per\\_il\\_personale\\_ospedaliero\\_il\\_governo\\_supporta\\_medici\\_e\\_infermieri-251363824/](https://www.repubblica.it/economia/2020/03/15/news/straordinari_maggiorati_per_il_personale_ospedaliero_il_governo_supporta_medici_e_infermieri-251363824/)
22. "La Repubblica", newspaper of national importance.  
[https://www.repubblica.it/cronaca/2020/03/02/news/coronavirus\\_la\\_sanita\\_nazionale\\_chiamata\\_a\\_raccolta\\_gli\\_ospedali\\_privati-249989748/](https://www.repubblica.it/cronaca/2020/03/02/news/coronavirus_la_sanita_nazionale_chiamata_a_raccolta_gli_ospedali_privati-249989748/)
23. "La Stampa", newspaper of national importance.  
<https://www.lastampa.it/torino/2020/03/29/news/coronavirus-piemonte-in-emergenza-le-ogr-diventano-ospedale-da-campo-e-si-raddoppiano-i-test-1.38652109>
24. Chen X, Yu B. First two months of the 2019 Coronavirus Disease (COVID-19) epidemic in China: real-time surveillance and evaluation with a second derivative model. Glob Health Res Policy. 2020 Mar 2;5:7. doi: 10.1186/s41256-020-00137-4.
25. Tang X, Wu C, Li X, Song Y, Yao X et al. On the origin and continuing evolution of SARS-CoV-2. National Science Review, 03 March 2020. <https://doi.org/10.1093/nsr/nwaa036>
26. Remuzzi A, Remuzzi G. COVID-19 and Italy: what next? Lancet. 2020 Apr 11;395(10231):1225-1228. doi: 10.1016/S0140-6736(20)30627-9. Epub 2020 Mar 13.