Review Article ISSN 2639-944X

Journal of Medical - Clinical Research & Reviews

Coronavirus (COVID-19): The Status in Italy Taken as an Example of the Virus Spreading in The World. Part II: The Models

Martino Recchia[1*], Umberto Cornelli[2] and Giovanni Belcaro[3]

[1]*University of Lugano, Switzerland.*

[2]*Loyola University School of Medicine, Chicago.*

[3]*IRVINE, University of Chieti-Pescara, Italy.*

[*]Correspondence:
Recchia Martino, Via Salaino 7, 20144 Milan, Italy.
Umberto Cornelli, Piazza Novelli 5 20129 Milan, Italy.
Giovanni Belcaro, 94 Strada Statale 16 Bis 65110 Spoltore (PE) Italy.

Received: May 2020; **Accepted:** June 2020

Citation: Martino Recchia, Umberto Cornelli, Giovanni Belcaro. Coronavirus (COVID-19): The Status in Italy Taken as an Example of the Virus Spreading in The World. Part II: The Models. J Med - Clin Res & Rev. 2020; 4(6): 1-5.

ABSTRACT

Background: *Model of the COVID-19 development are needed to have indication about its growth and decline over time.*

Objective: *To generate a model in relation to the Italian situation in the period between March 1st and April 22nd.*

Methods: *Five non-linear mathematical equations (Logistic 4P Rodbard; Weibull Growth; Logistic 4P; Logistic 3P; Exponential 3P) were used to describe the growth curves models relating to positive cases, deaths and their respective acceleration patterns. The analysed period is from 3rd March to 21st April.*

Each model has been adapted separately to the observations using JMP14 software of SAS Institute. Next, all models were tested for goodness of Fit (quality of prediction) using Akaike's information criterion (AICc), AICc Weight and Bayesian information criterion (BIC). The data were released by the Italian Heath Authorities and consistent for each of the 20 Italian Regions.

Results: *The Rodbard 4PL Logistics equation was the best model for describing the trend of positive, death and differential cases, because it presented the lowest AICc, and BIC values and the highest AICc weight, compared tothe other models.*

Conclusion: *The data show that the COVID-19 reached the pick in the month of March, and in April was starting the decay of both positive cases and deaths. A large difference was shown among the Regions both in terms of number of cases and deaths. The shape of the curves show that the decay seems to be asymptotic.*

Keywords
COVID-19, Non-Linear models, Italy, Italian Regions.

Introduction
The infection due to COVID-19 in Italy was one of the most severe in the world, accounting for > 200,000 cases and >24,000 deaths in a period of < 2 months.

However, differences among the Italian Regions, either in term of positive cases and death has been shown, giving some indication about the causes determining these differences [1]. Six of the most prosperous and industrial Regions (Lombardia, Veneto, Piemonte, Liguria, Toscana, and Marche) where affected, while in other Regions of the Centre, South and Island the infection was not very aggressive.

The correlations with the COVID-19 either in terms of positive cases and deaths were shown for the population density and the territory, in that people living in flat land and mountains were much more affected, as was also for the gross domestic product (GPD). Significant correlations with the number of workers and companies was shown, while temperature and humidity seemed not to be relevant.

The aim of the present study was to determine the best mathematical model describing the trend of the COVID-19 infection in Italy, in terms of positive cases and death.

Material and Methods

The values presented by the COVID-19 positive cases in Italy were recorded between March 1st and April 22nd as already presented in the previous paper [1]. Similarly, were considered for the same period the differentials (D) consisting of the values X_n-X_{n-1} (where N are the value of the day N^{-1}).

To describe the curves of positive and dead cases, 5 non-linear models were tested (Logistic 4PL Rodbard; Weibull Growth; Logistic 4P; Logistic 3P; Exponential 3P). The nonlinear procedure provides weighted estimates of the least squares of the parameters of the nonlinear model. The choice of the best fit was made by choosing the curve with the minimum AICc and BIC and the maximum AICc Weight. The 4PL model was applied for data analysis.

The same procedure was applied for differentials (D). The trends for the 20 different Italian regions were compared with the same model.

Results

Italy COVID-19 positive cases

The best fitting was obtained with the logistic equation characterized by 4 parameters (Rodbard 4PL or APL) according to the following formula:

$$Y=D+ A-D/ 1 + (x/C) \hat{} B$$

There are a number of ways to parameterize the 4PL model [3,4]. One parameterization, is shown in the previous equation, where Y is the observed response (positive cases, deaths, differentials) and A, B, C, and D are the curve parameters [2]. The values of A, B, C, and D are estimated using the data, and are not known a priori. The parameters A and D correspond to the asymptotes; B is called growth rater and C inflection point.

The different non-linear equations applied to the data suggested that the 4PL model well described (AICc Weight, p=1) the phenomenon of the infected whose number was 899 cases at 1st March and 69092 at 22nd April. The results of the selection models are shown in Table 1.

Positive cases

Model Comparison Report

Model	AICc	AICc Weight	BIC
Logistic 4P Rodbard	899,82	1	908,27
Weibull Growth	937,07	8,14E-09	944,03
Logistic 4P	974,19	7,09E-17	982,64
Logistic 3P	1018,76	1,49E-26	1025,71
Exponential 3P	1080,60	5,56E-40	1087,55

Table 1: Statistical indicators used for model's selection.

AICc

Gives a measure of the goodness of fit of an estimated statistical model that can be used to compare two or more models. AICc is a modification of the AIC adjusted for small samples.

AICc Weight

Gives normalized AICc values that sum to one. The AICc weight can be interpreted as the probability that a particular model is the true model given that one of the fitted models is the truth.

BIC

Gives a measure based on the likelihood function of model fit that is helpful when comparing different models. The model with the lower BIC value is the better fit.

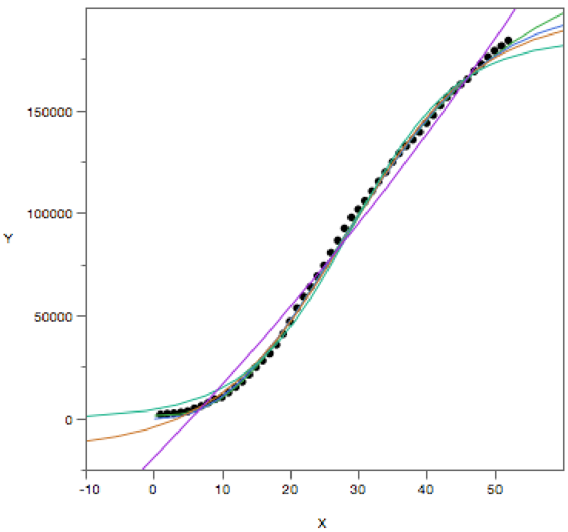

Figure 1: Positive cases of the five models.

All the parameters of the equation (Table 2), are statistically different from 0 to the Chi-Square test (p<0.001). By inserting in the equation, the number of positive cases of any day between March 1th and April 22th (Figure 1), the theoretical value calculated by the equation is obtained. Most of these values fall within the corresponding confidence intervals (CI 95%) to demonstrate the goodness of the fitting obtained.

Parameter Estimates

Parameter	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare
Growth Rate	-2,651	0,045	3459,505	<,0001
Inflection Point	34,468	0,422	6684,187	<,0001
Lower Asymptote	1601,429	486,007	10,858	0,001
Upper Asymptote	242877,500	3644,303	4441,663	<,0001

Table 2: Statistical tests on curve parameters.

Differentials between the daily data of positive cases (D)

The parameter values in the 4PL equation and the corresponding significance to Wald's test are shown in Table 3. The data are reported in Figure 2, and indicate the reduction of D during time. It can be observed one peak March 7th and a subsequent rapid reduction in the phenomenon. The decrease of D was evident, reaching the value of 1.015 at April 22nd (-32 % from March 1th).

Parameter Estimates				
Parameter	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare
Growth Rate	3,193	0,569	31,535	<,0001
Inflection Point	20,264	1,112	332,061	<,0001
Lower Asymptote	1,004	0,013	6317,000	<,0001
Upper Asymptote	1,232	0,009	17794,646	<,0001

Table 3: Statistical tests on curve parameters.

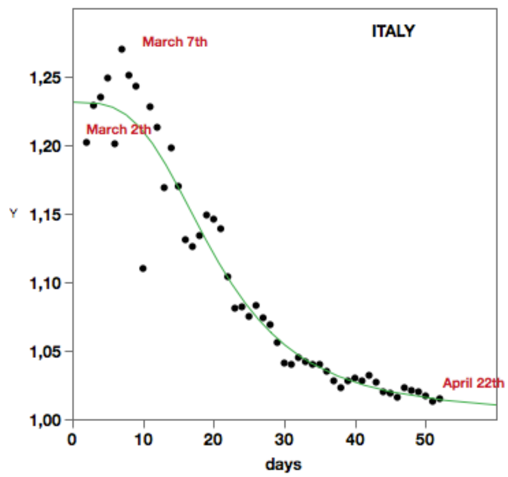

Figure 2: Curve of the D values from March 1st up to April 22nd.

Deaths from COVID-19

Similarly, to the positive cases the best fitting was obtained with the logistic equation characterized by 4 parameters (Rodbard 4PL).

Data are reported in Table 4, 5, and Figures 3 and 4.

Model Comparison Report			
Model	AICc	AICc Weight	BIC
Logistic 4P Rodbard	691,766	1	700,218
Weibull Growth	735,1155	3,86E-10	742,069
Logistic 4P	765,0702	1,21E-16	773,522
Logistic 3P	800,1625	2,90E-24	807,116

Table 4: Statistical indicators used for model's selection.

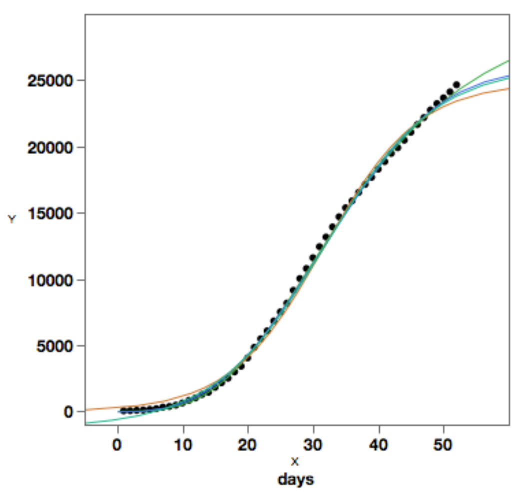

Figure 3: deaths from COVID-19 with the four models.

The Growth Rate for the curve is statistically different from 0 ($p<0.01$). The negative sign for Growth Rate specifies that expected mortality tends to increase over time.

Parameter Estimates				
Parameter	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare
Growth Rate	-3,192	0,052	3718,994	<,0001
Inflection Point	36,056	0,358	10123,217	<,0001
Lower Asymptote	44,198	56,606	0,610	0,4349
Upper Asymptote	31760,388	457,714	4814,851	<,0001

Table 5: Statistical tests on curve parameters.

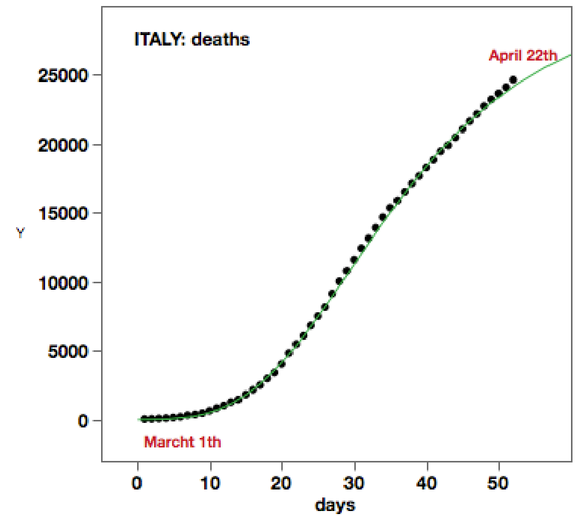

Figure 4: Trend for deaths during the considered period.

Differentials between the daily data on deaths (D)

The D is decreasing very slowly with a limited inversion at April 16th as reported in Figure 5.

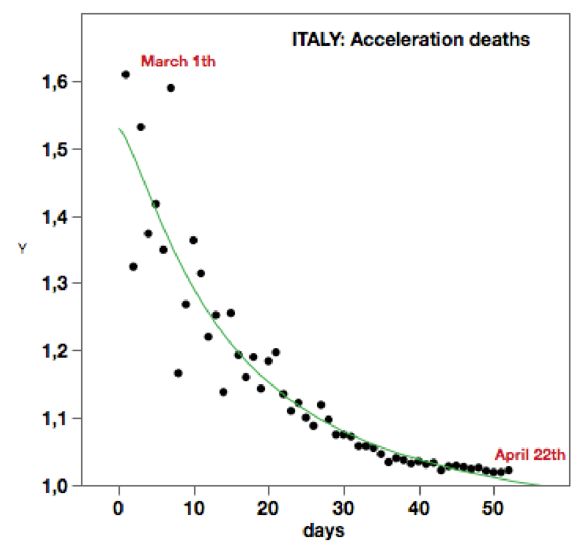

Figure 5: Curve of the D values from March 1st up to April 22nd.

The Growth Rate for the curve is statistically different from 0 ($p=0.0042$). The positive sign for Growth Rate specifies that expected mortality tends to decrease over time (Table 6).

Parameter Estimates				
Parameter	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare
Growth Rate	1,332	0,466	8,181	0,0042
Inflection Point	14,031	2,986	22,084	<,0001
Lower Asymptote	0,916	0,094	95,961	<,0001
Upper Asymptote	1,530	0,063	589,335	<,0001

Table 6: Statistical tests on curve parameters.

Predictor screening (PS)

The following topic deals with the research of the day with more responsibility than others in determining the positivity of cases on March 31th. The PS procedure indicates that this day corresponds to March 9th, as sown in Figure 6.

Predictor Screening

	March 31th		
Predictor	Contribution	Portion	Rank
March 9th	101115573	0,1215	1
March 2th	76243561	0,0916	2
March 1th	75176654	0,0904	3
March 7th	71554550	0,0860	4
March 4th	71228881	0,0856	5
March 8th	56148031	0,0675	6
March 11th	53635965	0,0645	7
March 6th	50242262	0,0604	8
March 10th	49463820	0,0595	9
March 14th	46272231	0,0556	10
March 13th	39868442	0,0479	11
March 12th	28125059	0,0338	12
March 5th	22523829	0,0271	13
March 3th	20299612	0,0244	14
March 18th	20078237	0,0241	15
March 17th	18188280	0,0219	16
March 19th	9039342,8	0,0109	17
March 21th	3675389,2	0,0044	18
March 27th	3153212,8	0,0038	19
March 15th	2605695	0,0031	20
March 24th	2541770,1	0,0031	21
March 29th	2245267,5	0,0027	22
March 16th	1943495	0,0023	23
March 30th	1682638,5	0,0020	24
March 25th	1618271,1	0,0019	25
March 28th	1198068,3	0,0014	26
March 22th	1045292,3	0,0013	27
March 26th	573791,64	0,0007	28
March 20th	392909,09	0,0005	29
March 23th	83525,949	0,0001	30

Figure 6: PS from the day March 1st to March 31st.

Positive cases at March 9th vs. days were correlated to define the best fitting curve (Table 7 and Figure 7).

Model Comparison Report			
Model	AICc	AICc Weight	BIC
Logistic 4P Rodbard	350,675	0,887	351,368
Logistic 4P	355,672	0,073	356,365
Logistic 4P Hill	356,898	0,040	357,591
Logistic 3P	366,420	0,000	367,736
Exponential 2P	383,559	0,000	385,046

Table 7: Statistical measures used for model's selection.

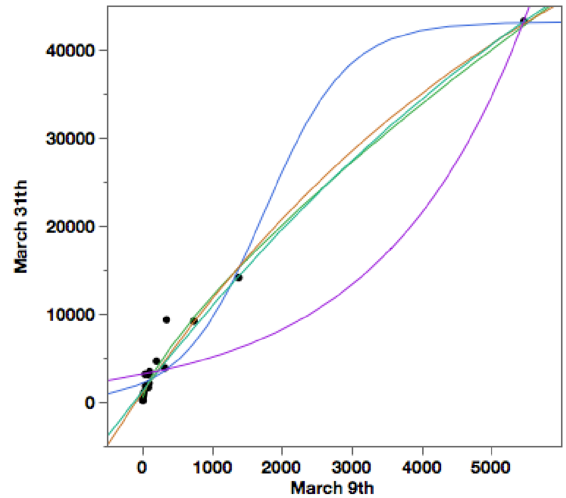

Figure 7: Each curve represents a model of the Table 7

The best fitting was obtained again with 4PL equation. According to these data the 4PL model were applied and results are reported in Figures 8.

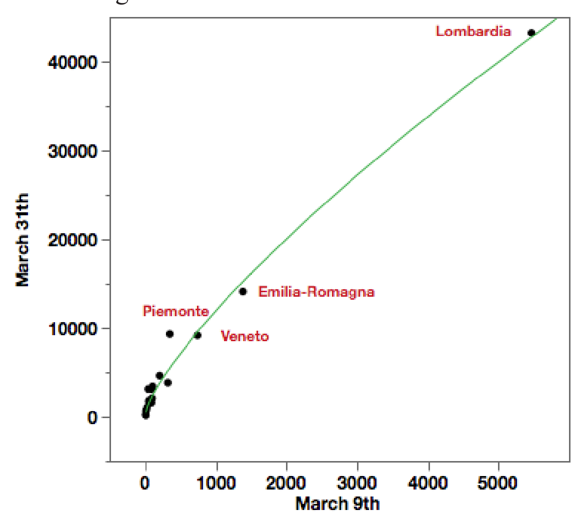

Figure 8: Positive case curve March 9th vs March 31th.

The most affected Italian Regions (Lombardia, Emilia Romagna, Veneto, Piemonte) are indicated.

Discussion

The limitation of the presents analysis is due to the lack of the knowledge of the real data, both in terms of positive cases and deaths.

In the case of positive swabs, the Regions were undertaking different methods: some of the Regions were taking swabs only for subjects presenting symptoms, while other were also considering apparently heathy subjects.

In reaction to deaths, only those within the hospitals were considered, but patients who died at home may have not been completely reported. A highly significant correlation was found between swabs, positive cases and deaths, meaning that the present data, despite they represent the "pick iceberg", can give a sufficient reliable picture of the real curves during time. Considering that the positive cases are supposed to represent about 8-10% of the real cases figures around 2 million of people can be credible. The present ratio between positive cases/death account for and with the present ratio positive cases/death accounts for about 10% of the cases.

However, the clinical protocols are improving day by day, and hopefully one may expect that this ratio will improve. A particular attention should be given to the aerol use of antiviral/anti-inflammatory agents to bypass the organ toxicity and focus on the lung with less aggressive dosages.

The simplest curves, of both positive cases and deaths in terms of D, seems to be asymptotic, which indicates that it will be hardly possible to define a time "zero", but most realistically numbers that for a long period of time will stay much over this value, hopefully with affordable figures.

One may not rule out the event that COVID-19 is already part of the human metaorganisms, which means we may expect periodical recrudescence. As for every infection/pandemia, the real matter is the prevention which can be done also in the case of COVID-19 as it was for AIDS or viral hepatitis.

Prevention cannot be based only on the lockdown, or segregating elderly people at home. There are many other possibilities, and mask use with a limitation of large gatherings can be the first affordable step.

Conclusion

Despite some minor bias, the trend of COVID-19 infection in terms of positive cases and death can be explained at the best using the Logistic 4PL Rodbard.

References

1. Cornelli U, Belcaro G, Cerarone MR, et al. Coronavirus COVID-19 the status in Italy taken as an example of the virus spreading in the world. J MCRR. 2020; 4: 1.
2. JMP14 Pro SAS Insitute
3. Prentice, Finney, Rodbard, et al. 1976.
4. Ratkowsky DA, Reedy TJ. Choosing near-linear parameters in the four-parameter logistic model for radioligand and related assays. Biometrics. 1986; 42: 575-582.