# Modelling the epidemiological trend and behavior of COVID-19 in Italy

*Alessandro Rovetta*[1], *Akshaya Srikanth Bhagavathula*[2],

[1]Mensana srls research and disclosure division, Via Moro Aldo 5 - 25124 Brescia, Italy.

[2]Institute of Public Health, College of Medicine and Health Sciences, UAE University, Al Ain, UAE.

**Corresponding authors:**

Alessandro Rovetta

Mensana srls research and disclosure division

Via Moro Aldo 5 - 25124 Brescia, Italy.

Email: rovetta.mresearch@gmail.com

Phone: +39-3927112808

**Abstract**

As of March 16, 2020, over 185,000 across the world, Italy became the red hotspot for the COVID-19 pandemic after China. With over 35,000 cases and 2900 deaths reported in the month of March in Italy, it is necessary to stimulate epidemic trend to understand the behavior of COVID-19 in Italy. By S.E.I.R. simulation, we estimated the most representative epidemic parameters occurred from March 1 to 14, 2020, thus being able to evaluate the consistency of the containment rules and identify possible Sars-Cov-2 local mutations. Our estimations are based on some assumptions and limitations exited.

**Keywords:** *COVID-19, coronavirus, contamination, novel coronavirus, Sars-Cov-2, Europe, Italy.*

## Introduction

The current surge of COVID-19 pandemic is devastating globally, with over 200,000 cases and more than 8600 deaths reported [1]. In Europe, COVID-19 cases are most dramatically started to increase from the first week of March 2020. Of these, Italy is grappling with the worst outbreak, with over 35,713 confirmed cases and around 3000 deaths by March 18, 2020 [1-5]. This exponential increase in COVID-19 positive cases in Italy raised turmoil, and the government decree to a lockdown of the entire country [1]. Given the seriousness of the situation, it is absolutely necessary not only an immediate intervention but also a criterion for assessing its effectiveness.

Article on modeling epidemic transition of COVID-19 have been published [1-8] and based on the previous publication from China, South Korea, Iran, and Japan presented the estimation of epidemic trends and transmission rates [1-8]. Through this research, we evaluated the consistency of the containment rules and identified possible SARS-CoV-2 local mutation using the S.E.I.R mathematical model. We used the most representative epidemic parameters that occurred during the first half of March 2020 to predict the trend of infections. Therefore, to assess the effectiveness of the containment measures, it will signal the presence of a plausible evolutionary mutation. To do this, it will be sufficient to compare the general trend foreseen by the S.E.I.R. with the real Italian population.

## Methods

To carry out this study, we used the most recent data, found in the scientific literature relating to COVID-19 total and active cases, deaths, recoveries, and all epidemic parameters, have been used. Since the Chinese situation has provided the most information, great statistical importance

was given to the Sars-Cov-2 evolution in the Hubei region, with the implicit assumption of "very significant local mutations" absence necessary for generalizing the reports through classical inference. The statistics collected on virus mortality were first divided by age group and then recalibrated on the Italian demography, in order to be as representative as possible. To do this, we considered the most representative population of COVID-19 cases reported between March, 1-14, 2020, the number of infected persons necessary to approach the required mortality rate was added to the correct values.

**S.E.I.R. Modelling**

To obtain a more realistic trend and assuming true the probable "patients non-relapse", and we applied S.E.I.R. model to predict the virus progress in Italy. As a note, it is not possible to estimate the containment measures taken by the government. However, at the same time, assuming that only half of the population was susceptible to the virus precisely due to the above containment measures. Thanks to the comparison between real and theoretical evolution, it is likely to estimate the presence of essential mutations and/or the limitation strategy effectiveness through similarities, anomalies, or substantial deviations from each other. We used S.E.I.R. differential equations and non-linear methods to resolve the gaps analytically. An interactive algorithm was developed using C++ software (version) to find a solution through a finite discretization method. The total population number has been considered constant because of the very low deaths/population ratio.

**Software iterative algorithm**

By entering the initial values for the incubation time $1/\sigma$, recovery time $1/\gamma$, basic reproduction number R0, and number of infected I0 on March 1, 2020, the software prints the S.E.I.R. values

day by day. The best epidemic parameters were estimated through continuous iteration until the "closest values to the real ones" were reached until March 14, 2020. The number of initial incubates was calculated with the formula E0=R0·I0.

## Results

### Epidemic parameters

After the reconstruction of the real data on COVID-19 in Italy in the period March 2020, 1-14, through the above methods, the best estimates obtained for the Italian epidemic parameters are $1/\sigma = (3\pm1)$ days, $1/\gamma = (15\pm3)$ days, R0 = 3.51 ± 5%, I0 = 3350 ± 20%.

### Best statistical forecasts

The model predicts the following values for the next 7 days (March 2020, 15-22) with "*D* " deaths number:

March, 15: $S = 2.99308\cdot10^7$; $E = 21504$; $I = 40651$; $R = 18854$; $D = 1904$;

March, 16: $S = 2.99207\cdot10^7$; $E = 23988$; $I = 45363$; $R = 21719$; $D = 2124$;

March, 17: $S = 2.99095\cdot10^7$; $E = 26759$; $I = 50619$; $R = 24917$; $D = 2417$;

March, 18: $S = 2.98970\cdot10^7$; $E = 29849$; $I = 56481$; $R = 28486$; $D = 2719$;

March, 19: $S = 2.98830\cdot10^7$; $E = 33293$; $I = 63019$; $R = 32467$; $D = 3056$;

March, 20: $S = 2.98674\cdot10^7$; $E = 37132$; $I = 70308$; $R = 36909$; $D = 3431$;

March, 21: $S = 2.98501\cdot10^7$; $E = 41410$; $I = 78437$; $R = 41865$; $D = 3850$;

March, 22: $S = 2.98307\cdot10^7$; $E = 47394$; $I = 87500$; $R = 47394$; $D = 4317$;

These values obviously refer to the population with the constant mortality rate of 3.2%; however, the real number of deaths must be strictly less than the theoretical one expected or we will talk about mutation. Thus, thanks to the results obtained from the application of the S.E.I.R. model we can foresee three possible scenarios: If, in Italy, the real next-days trend of the total infected number will be lower than that shown in figure 1, we can assume the following events set out in order of probability: the containment measures adopted 10 days ago are taking effect; Sars-Cov-2 has undergone a significant anti-evolutionary mutation. If, in Italy, the real next-days trend of the total infected number will be equal than that shown in figure 1. We assume the following events set out in order of probability: the containment measures adopted 10 days ago are not taking effect; the containment measures adopted 10 days ago are taking effect and Sars-Cov-2 has undergone a significant evolutionary mutation. If, in Italy, the real next-days trend of the total infected number will be higher than that shown in figure 1. We assume the following events set out in order of probability: the containment measures adopted 10 days ago are not taking effect and Sars-Cov-2 has undergone a significant evolutionary mutation; the S.E.I.R model is no-more representative of the COVID-19 Italian case and we should utilize the S.E.I.R. model.

**Comparison with real national and per provinces data**

Based primarily on the deaths number i.e. data with the lower margin of error, we compare real and theoretical data [14]:

| | | | | |
|---|---|---|---|---|
| *March Days* | 15 | 16 | 17 | 18 |
| *Real Deaths* | 1809 | 2158 | 2503 | 2978 |
| *Theoretical Deaths* | 1904 | 2124 | 2417 | 2719 |
| *Difference (%)* | -5.3 | +1.6 | +3.4 | +8.7 |

Even considering an admissible "theoretical-estimates statistical error" of ± 5% these values could not be accepted since the trend indicates a faster growth than that predicted by the model

with a breakeven point between 15 and 16 March and a percentage difference in net growth. Deepening the discussion and going to analyze the cases region by region, a very significant result turns out: with the exception of Lombardy, the trend in the number of infected is very good [18]. Unfortunately, this figure is extremely negative since it indicates the possible presence of a much more aggressive and/or contagious evolutionary mutation in Lombardy: in fact, the weight of Lombardy in the national statistics of total dead and infected is very high and, moreover, the slope of the growth curve of both it is much greater even than in the second region with the most infected. Quantifying, Lombardy went from 5791 cases on March 10 to 17713 on March 18, i.e. a slope of 1490 cases/day; Emilia Romagna, second region by growth and number of cases, has an average slope over the same period of time equal to 374 cases/day (1533 to 4525 cases). Going further, we can see how three cities in particular are causing alarm: Brescia, Bergamo and Milan, with similar angular coefficients. As if that were not enough, Lombardy holds 66% of national deaths despite representing "only" 49% of national cases; in addition, the number of Deaths and Infections grow at a rate much higher than the S.E.I.R. model, despite all government restrictions.

## 1. Choice of model

Given that

- It seems unlikely that the virus, without significantly changing, could infect a patient again [1,2] [12],
- the mortality rate in Italy is too low to be significant in the short term [8],
- young children appear to have a secondary role in the spread of the infection [8][17],

We will investigate the possible scenarios through S.E.I.R. model.

### 1.1 S.E.I.R. model: Italy's COVID-19 evolution

Using the discretization carried out in the previous study "Mathematical- statistical modeling of COVID- 19 on the restricted population" [8], a new C++ software was created to print the values of the S.E.I.R. day by day until a specific date and/or situation(2). However, unlike the aforementioned study, the prediction period will be extended to a certain number of days in order to verify the effectiveness of the containment measures in Italy.

## 2. Data analysis

Since no European or international standard for data collection has been established, one of the crucial points of this analysis is the interpretation of the information available. On February 27, in Italy, a real "communicative turnaround" was announced so that only the most important symptomatic cases would be publicly counted from then on [6]. Since there do not appear to have been significant large-scale mutations and that a good amount of data is now available, we expect the WHO mortality data to be valid in Italy as well [13] [Table 1].

The more recent study about Sars-Cov-2 mortality on symptomatic and asymptomatic COVID-19 patients in the Hubei region shows the following Sars-Cov-2 mortality per age-groups [3] [Table 2].

Since the Italian population is distributed according to the following age-percentages [15][3] [Table 3]

$\Sigma i \; P\, i \cdot DR\, i \,/\, 100 = 3.2\%$

For this reason, it makes sense to conclude that the Italian COVID-19 death rate is exaggeratedly overestimated if we rely on Ministry of Health data. On the other side, the true-infect number (as the aforementioned "communicative turn" announced) must necessarily be higher. Trying to respect the percentage just calculated we obtain the following table 4.

## 3 Application of the S.E.I.R. predictive model

Citing the adapted system of non-linear differential equations and their discretization [8] and and using the WHO estimated values for incubation and recovery times respectively equal to $1/\sigma \in [3;6]$ days range[9] and a generous $1/\gamma \in [15;35]$ days range (since you can infect even after being healed [5]), with a $R0 \in [2;4]$ [7] [10] [4] we have the following plausible results from our C++ software:

$$
\begin{aligned}
S'(t) &= \beta \cdot I \cdot S/N \\
E'(t) &= \beta \cdot I \cdot S/N - \sigma \cdot E \\
I'(t) &= \sigma \cdot E - \gamma \cdot I \\
R'(t) &= \gamma \cdot I \\
N'(t) &= 0 \\
Si+1 &= \beta \cdot Ii \cdot Si/Ni \cdot \delta t + Si \\
Ei+1 &= \beta \cdot Ii \cdot Si \cdot \delta t + Ei \\
Ii+1 &= (\sigma \cdot \mathrm{E}i - \gamma \cdot \mathrm{I}i) \cdot \delta t + Ii \\
Ri+1 &= \gamma \cdot Ii \cdot \delta t + Ri \\
Ni &= Si + Ei + Ii + Ri
\end{aligned}
$$

Legend:

- $1/\gamma$ is the asymptomatic incubation time;
- $1/\sigma$ is the recovery time;
- R0 is the basic reproduction number;

- $S$ is the number of susceptible people;
- $E$ is the number of active exposed people (people in incubation);
- $I$ is the number of active infected people;
- $R$ is the number of recovered people (no longer infectable).
- $X_n$ is the number of day-n "X-category" people;
- $X_t$ is the number of table 4 "X-category" people;

<u>Significant Test-Collection 1</u>

Assumptions: $1/\sigma \in [3;6]$ days, $1/\gamma = 35$ days, $R0 \in [3;4]$

Significant results: $I0 \in [5;8] \cdot 10^3 >> I0t = 2202$, $R14 \in [8;9] \cdot 10^3 >> R14t = 3327$.

**Discussion**

In order for the supposed data to be representative of reality, we should admit the existence of many more cases than expected. All this is consistent both with the idea the infection had been present on the territory for some time, in the form of a less aggressive strain, than with the time course of the disease [1]. The number of people healed is also in line with international percentages [16].

Validity grade: plausible and supported by empirical evidences.

Warning level: high.

<u>Significant Test-Collection 2</u>

Assumptions: $1/\sigma = 3$ days, $1/\gamma = 35$ days, $R0 = 6.05$.

Significant results: $I0 = I0t = 2202$, $R14 = 7169 >> R14t = 3327$.

Another possible scenario that describes the table 4 data. In this case the value of R0 would mean an extremely more aggressive/infective genetic mutation.

Validity grade: not excludable, but improbable.

Warning level: extremely high.

<u>Significant Test-Collection 3</u>

Assumptions: $1/\sigma = 3$ days, $1/\gamma = 15$ days, $R0 = 3.51$.

Significant results: $I0 = 3350 > I0t = 2202$, $R14 = 16219 >> R14t = 3327$.

Discussion: a third notable possible scenario that describes the table 4 infections. This latter case fits the estimated theoretical data almost perfectly and ranks first as a representative model.

Validity grade: very plausible and supported by empirical evidences.

Warning level: high.

## 3.1 Best Test forecast and graphs: Test 3

Since test 3 turned out to be the most likely, we focus on it.

Peak-day: 80.

Peak value: $Imax = 8.86694 \cdot 10^5$ people (almost 9 million people);

$I < 10$ – day: 333 (almost one year);

$I = 0$ – day: 386.

We can also fit the curve through a convenient Gaussian $f(x) = y0 + Gauss(A,x0,\sigma)(x)$ function, useful for modeling the initial trend of the virus.

**Conclusions**

The model is approximate as a direct consequence of the uncertainty of the data it was supposed to fit but it can act as a valid comparison for the identification of any Sars-Cov-2 genetic mutations as well as for the evaluation of the containment measures effectiveness. Observed the comparison between the predictions of the S.E.I.R. and the actual trends highlighted above, as well as between the trends of Lombardy and the other Italian regions, there are valid reasons to assert that in Lombardy, in particular in the cities of Brescia, Bergamo and Milan, a very significant evolutionary mutation of the virus may be underway, which would require immediate containment.

**Acknowledgment:** None.

**Authors contribution**

AR and ASB designed the study, collected the data, analyzed the data, prepared the manuscript. All the author read and approved the final manuscript.

**Funding:** No source of funding

**Available data and materials:** All the materials are attached as supplementary and information related to the study are in the manuscript.

**Consent for publication:** Not applicable.

**Competing interests:** The authors declare that they have no competing interests.

**References**

[1] Alessia Lai, Annalisa Bergna, Carla Acciarri, Massimo Galli, Gianguglielmo Zehender, Early Phylogenetic Estimate Of The Effective Reproduction Number Of 2019-nCoV, February 23, 2020. DOI: 10.1101/2020.02.19.20024851. https://www.medrxiv.org/content/10.1101/2020.02.19.20024851v1

[2] Centers for Disease Control and Prevention. https://www.cdc.gov/coronavirus/2019-ncov/hcp/faq.html

[3] Julien Riou, Anthony Hauser, Michel J Counotte, Christian L Althaus, Adjusted age-specific casefatality ratio during the COVID-19 epidemic in Hubei, China, January and February 2020, p.5. DOI: https://doi.org/10.1101/2020.03.04.20031104.

[4] Kamalich Muniz-Rodriguez, Isaac Chun-Hai Fung, Shayesteh R. Ferdosi, Sylvia K. Ofori, Yiseul Lee, Amna Tariq, Gerardo Chowell, Transmission potential of COVID-19 in Iran, March 2020, 10,Abstract. DOI: https://doi.org/10.1101/2020.03.08.20030643. https://www.medrxiv.org/content/medrxiv/early/2020/03/10/2020.03.08.20030643.full.pdf?

%253fcollection=

[5] Lan Lan, Dan Xu, Guangming Ye, Positive RT-PCR Test Results in Patients Recovered From COVID-19, February 27, 2020. DOI: 10.1001/jama.2020.2783, JAMA. https://jamanetwork.com/journals/jama/fullarticle/2762452

[6] Nicola Barone e Marzio Bartoloni, ilsole24ore, La giravolta comunicativa sul coronavirus, menotamponi e contare solo i casi gravi, February 2020, 27. https://www.ilsole24ore.com/art/la-giravolta-comunicativa-coronavirus-meno-tamponi-e-contare-solo-casi-gravi-ACQYXQMB

[7] Novel coronavirus disease 2019 (COVID-19) pandemic: increased transmission in the EU/EEA and the UK - sixth update. https://www.ecdc.europa.eu/sites/default/files/documents/RRA-sixth-update-Outbreak-of-novel-

coronavirus-disease-2019-COVID-19.pdf

[8] Rovetta Alessandro, Mathematical-statistical modeling of COVID-19 on the restricted population, February 2020, 29, p.2 and p.3 f. 2.1.2. Mensana srls research and disclosure division, Via Moro Aldo 5 - 25124 Brescia, Italy. https://www.researchgate.net/publication/339566420_Mathematical-statistical_modeling_of_COVID-19_on_the_restricted_population

[9] WHO database, Global research on coronavirus disease (COVID-19). https://www.who.int/emergencies/diseases/novel-coronavirus-2019/global-research-on-novel-coronavirus-2019-ncov

[10] Zhidong Cao, Qingpeng Zhang, Xin Lu, Dirk Pfei#er, Zhongwei Jia, Hongbing Song, Daniel Dajun Zeng, Estimating the effective reproduction number of the 2019-nCoV in China January 29, 2020. DOI: 10.1101/2020.01.27.20018952. https://www.medrxiv.org/content/10.1101/2020.01.27.20018952v1

[11] https://www.aljazeera.com/news/2020/01/timeline-china-coronavirus-spread-200126061554884.html

[12] https://www.newscientist.com/article/2236544-coronavirus-are-there-two-strains-and-is-one-more-deadly/

[13] https://nextstrain.org/ncov

[14]http://www.salute.gov.it/portale/news/p3_2_1_1.jsplingua=italiano&menu=notizie&p=dalministero&dataa=2020/12/31&datada=2017/01/01

[15] https://www.tuttitalia.it/statistiche/popolazione-eta-sesso-stato-civile-2019/

[16] https://www.worldometers.info/coronavirus/

[17] https://www.worldometers.info/coronavirus/coronavirus-age-sex-demographics/

**Table 1:** Italy Ministry of Health March 2020, 1-14 official data [14]**.**

| March Day | Total Cases | Deaths | Recoveries | Active Cases | Recovery (%) | Death (%) |
|---|---|---|---|---|---|---|
| 1 | 1101 | 41 | 83 | 977 | 7,5 | 3,7 |
| 2 | 1436 | 52 | 149 | 1235 | 10,4 | 3,6 |
| 3 | 1902 | 79 | 160 | 1663 | 8,4 | 4,2 |
| 4 | 2489 | 107 | 276 | 2106 | 11,1 | 4,3 |
| 5 | 3258 | 148 | 414 | 2696 | 12,7 | 4,5 |
| 6 | 4036 | 197 | 523 | 3316 | 13,0 | 4,9 |
| 7 | 5283 | 233 | 589 | 4461 | 11,1 | 4,4 |
| 8 | 6775 | 366 | 622 | 5787 | 9,2 | 5,4 |
| 9 | 8572 | 463 | 724 | 7385 | 8,4 | 5,4 |
| 10 | 10749 | 631 | 1004 | 9114 | 9,3 | 5,9 |
| 11 | 12462 | 827 | 1045 | 10590 | 8,4 | 6,6 |
| 12 | 15113 | 1016 | 1258 | 12839 | 8,3 | 6,7 |
| 13 | 17660 | 1266 | 1439 | 14955 | 8,1 | 7,2 |
| 14 | 21157 | 1441 | 1996 | 17720 | 9,4 | 6,8 |

**Table 2:** The more recent Hubei death-rate per age.

| Age-groups (years) | | Death Rate (%) = DR i |
|---|---|---|
| 80 ++ | | 18 |
| 70 | – 79 | 9,8 |
| 60 | – 69 | 4,6 |
| 50 | – 59 | 1,3 |
| 40 | – 49 | 0,4 |
| 30 | – 39 | 0,18 |
| 20 | – 29 | 0,09 |
| 10 | – 19 | 0,02 |
| 0 | – 9 | 0 |

**Table 3:** Italian population per age**.**

| Age-groups (years) | | Population (%) = P i |
|---|---|---|
| 80++ | | 7,2 |
| 70 | 79 | 9,9 |
| 60 | 69 | 12,2 |
| 50 | 59 | 15,5 |
| 40 | 49 | 15,3 |
| 30 | 39 | 11,7 |
| 20 | 29 | 10,3 |
| 10 | 19 | 9,6 |
| 0 | 9 | 8,3 |

**Table 4:** Recalculated Italian data of COVID-19 cases in March, 2020

| March Day | Total Cases | Deaths | Recoveries | Active Cases | % Recovery | % Death |
|---|---|---|---|---|---|---|
| 1 | 2202 | 41 | 166 | 1995 | 7,5 | 1,9 |
| 2 | 2872 | 52 | 298 | 2522 | 10,4 | 1,8 |
| 3 | 3804 | 79 | 320 | 3405 | 8,4 | 2,1 |
| 4 | 4978 | 107 | 552 | 4319 | 11,1 | 2,1 |
| 5 | 6516 | 148 | 828 | 5540 | 12,7 | 2,3 |
| 6 | 8072 | 197 | 1046 | 6829 | 13,0 | 2,4 |
| 7 | 10566 | 233 | 1178 | 9155 | 11,1 | 2,2 |
| 8 | 13550 | 366 | 1244 | 11940 | 9,2 | 2,7 |
| 9 | 17144 | 463 | 1448 | 15233 | 8,4 | 2,7 |
| 10 | 21498 | 631 | 2008 | 18859 | 9,3 | 2,9 |
| 11 | 24924 | 827 | 2090 | 22007 | 8,4 | 3,3 |
| 12 | 30226 | 1016 | 2516 | 26694 | 8,3 | 3,4 |
| 13 | 35320 | 1266 | 3166 | 30888 | 9,0 | 3,6 |
| 14 | 42314 | 1441 | 3992 | 36881 | 9,4 | 3,4 |

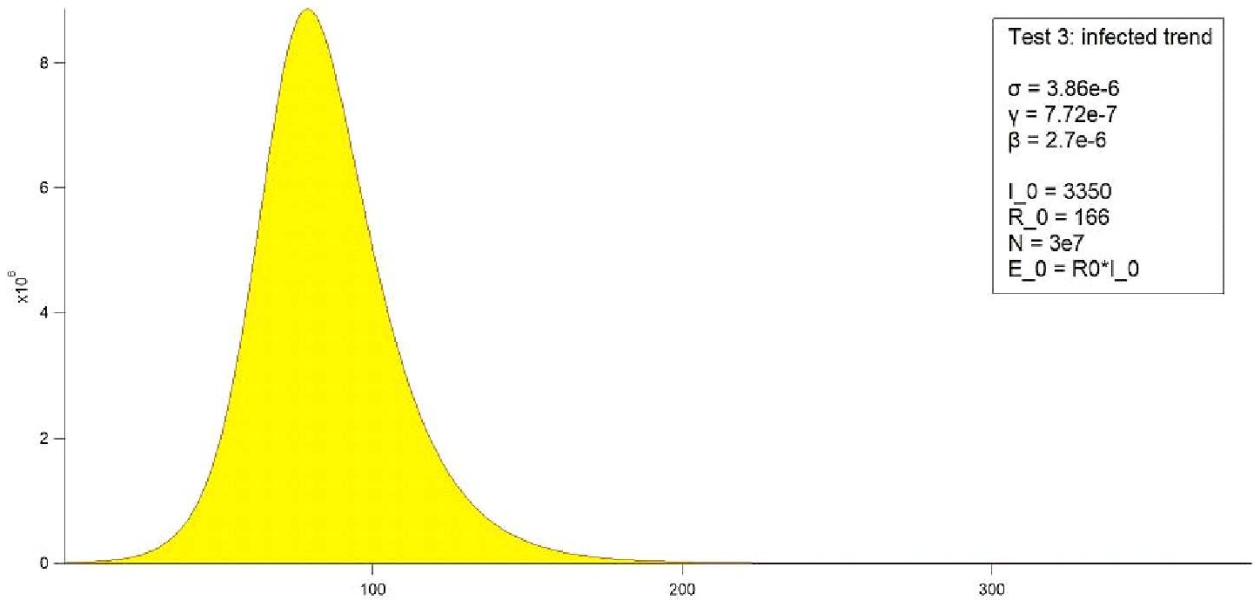


**Figure 1: Test-3 – Italy infected trend.**

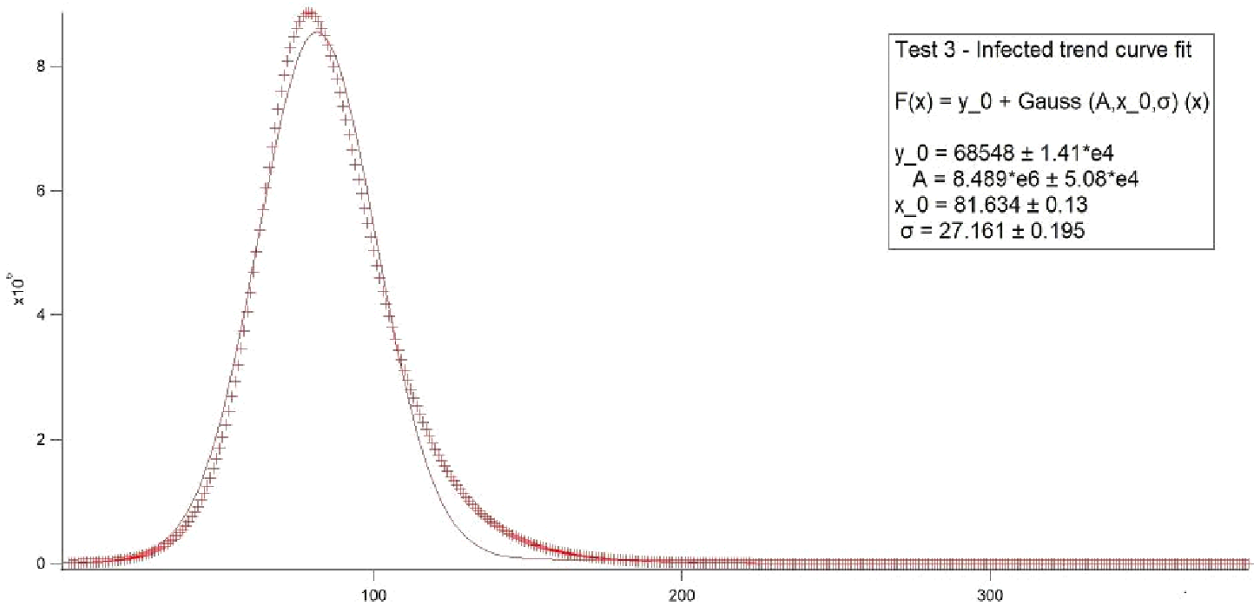


**Figure 2: Test 3 – Italy infected trend, gaussian curve fit.**