COVID-19 infection-to-death lag and lockdown-effect lag in Italy and Spain

Jorge Buescu
Universidade de Lisboa Faculdade de Ciencias
Henrique Manuel Oliveira (henrique.m.oliveira@tecnico.ulisboa.pt)
Technical Institute of Lisbon https://orcid.org/0000-0002-3346-4915
Carlos Alberto Pires
Universidade de Lisboa Faculdade de Ciencias

Research article

Keywords: Covid-19, Logistic approach, Epidemic model, Infection-to-death lag

Posted Date: April 27th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-22130/v1

Title: COVID-19 infection-to-death lag and lockdown-effect lag in Italy and Spain

Authors: Jorge Buescu[1], Henrique M. Oliveira[2,*] and Carlos A. Pires[3]

Affiliations: [1]DM-FCUL and CMAFCIO Universidade de Lisboa, 1749-016 Lisbon. Portugal.

[2]CAMGSD DM-IST Universidade de Lisboa, Av. Rovisco Pais, 1, 1049-001 Lisbon. Portugal.

[3]Instituto Dom Luiz (IDL), Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisbon, Portugal.

[*]Corresponding author. Email: henrique.m.oliveira@math.tecnico.ulisboa.pt

Abstract

Background: The European phase of the COVID (SARS-CoV-2) pandemic started in early February 2020. Lockdown measures have been adopted to flatten the infection curve. Early detection and robust estimation of its inflection point and other phenomenological epidemic parameters are extremely relevant to evaluate effectiveness of containment measures and the epidemic course.

Methods: We use a nonlinear least-squares algorithm to fit logistic functions to the curves of active infected and deceased during the pandemic evolution, studying also the changes and stabilization of the logistic parameters.

Results: We present the computation of the inflection points for Spain and Italy. We use these results to estimate two fundamental disease parameters: the infection-to-death lag and the onset of lockdown effects. These estimates are sharp and robust.

Conclusions: The data-driven logistic-fit approach has proven to be a remarkably powerful tool, allowing us to effectively compute three disease parameters which may prove to be an invaluable help in the control of the epidemic as well as to model response of the health services in other countries in the World. The results and methods can be used to forecast the evolution in countries still in early stages of the epidemic.

Keywords: Covid-19, Logistic approach, Epidemic model, Infection-to-death lag.

Background

The European phase of the COVID pandemic started in February 2020. The initial outbreak in Italy was closely followed in other countries, spreading across Western Europe. Without any immunity, epidemic models predicted an infection curve with a peak value of 70% of the population if no action was taken. These countries had to undergo lockdown – Italy partially on March 8 and totally on March 11 (day 25 from Feb. 15), and Spain on March 16 (day 30). At the end of March Italy had over 100,000 cases and 12,000 dead, and Spain about 85,000 cases and 8,000 dead. It thus became a critical issue to detect the inflection point of the infection curve in order to evaluate effectiveness of containment measures. We perform these computations and use the corresponding results to estimate essential disease parameters as infection-to-death lag and onset of lockdown effects, which may prove crucial for the control of the epidemic elsewhere. The logistic fit of (the increasing section

of) these curves has proven to be an extremely useful tool, allowing us to estimate these parameters quite sharply.

Methods

Detection of the inflection points

In this paragraph we take a completely data-driven approach, treating the system as a black box outputting raw data and ignoring any underlying biological mechanisms. We analyse reliable public data openly available [1,2], using them to detect possible inflection points in the number of active cases and in the number of deaths. We work with the active cases at each moment and not with the accumulated infections since only actives contribute, at each instant, to contagion.

We model the ascent segment of both the active infectious and the deceased curves with a logistic curve [3,4,5,6,7,8]

$$N(t) = \frac{L}{1 + e^{-r(t-T_i)}}$$

performing a least-squares fit (minimization of the sum of squared errors SSE along the fitting period) [9].

We optimize all three independent parameters: L (carrying capacity) , T_i (inflexion time) and r (exponential rate) by solving a nonlinear least-squares problem through the multivariate Newton-Raphson iterative algorithm [10,11,12,13,14]. The initial guess of parameters is obtained from 3 regularly separated values along the fitting period. Moreover, a proper parametric scaling is mandatory in order to ensure convergence of the algorithm. The method details are provided in the appendix.

The goodness-of-fit measure adopted is the relative error $GOF = \sqrt{SSE/SS}$, where *SS* is the sum of squared observations. The sensitivity, both of the optimal parameters and *GOF* to the final day t_e of the fitting period, is studied.

Estimation of the time lag between onset of symptoms and death in COVID-19

The technique presented here provides a simple way to estimate the average time lag between onset of symptoms and death in COVID-19 infected individuals. It can also be applied to other variables which are obtained from the infected cases by integration, as is the case of deceased in the usual ODE epidemiological models [3,15, 16] SIR (Susceptible-Infected-Recovered), SEIR (Susceptible-Exposed-Infected-Recovered) and their variants [17,18,19] dealing specifically with COVID modelling.

The time lag between onset of symptoms and death (TLSD) is an important measure by itself and of vital interest for health care system management. It is particularly relevant in the estimation of under-reporting of infected cases [20] and in the estimation of the severity of the infection [21].

Lag between onset of symptoms and death in Wuhan was assumed by different authors to be 15 to 20 days [*22*], 21 to 22 days [21] or 18 days [23], and was taken, as a general prototype for this disease, to be 14 days in the recent work by Baud et al [24].

Distinct approaches produce different numbers. This lag, naturally, varies from region to region and country to country. Factors like healthcare system and demographic structure, among others, are important variables concerning TLSD.

We present a purely data analytic estimation of TLSD in the two European countries which have suffered to date the most number of fatalities with this pandemic: Spain and Italy.

This measure could also be performed for the onset of symptoms and recovery, but the exactness of the moment of recovery is not clear with current available data, in opposition to the time of death for individuals in Europe, which is usually very precise.

We measured the distance of the graph of the number of active infected cases to the graph of the curve of deceased for Italy and Spain. We shifted the curve of deceased using as time unit the day. We considered the time interval where the behavior of the number of infected $I(t)$ and the number of deceased $D(t)$ with time is approximately exponential in the onset of the local epidemics. Obviously, this assumption is not necessary in a general framework of logistic approximations, or for the general solutions of epidemic models (SIR, SEIR [3,15] and its multiple variants [17,18]).

For the exponential phase of the outbreak we consider the curve of the daily increment of active infected cases and the curve of the daily increment of the fatalities and their corresponding values obtained from the logistic fit of $I(t)$ and $D(t)$ respectively. Then we vary the time lag between these curves, computing the Euclidean distance in function of

each possible time shift. The desired value of TLSD is the minimizer of this distance function. We perform this computation in the period ranging between the time corresponding to a minimum threshold of actives (e.g. 100) and the point of inflection in the curve of actives estimated from the logistic fit.

The technical details of the methods are presented in the appendix.

Results

Inflection Point

Our findings are the following. In the case of Italy, the inflection point of the active curve (Fig. 1) was attained on days 35-36, corresponding to 21-22 March (Fig. 2b), where day 0 is set to 15 February.. This result seems extremely robust: up until day 36 a hypothetical inflection point was placed always in the last day of the series, indicating its effective inexistence. In fact, during that stage, the solution is approximately exponential, and one of the three parameters is redundant, since L/e^{rT_i} becomes approximately constant. From that point onwards the inflection point froze on day 36, as may be seen in Fig. 2b), and parametric degeneracy disappears. Before the inflection the *GOF* measure (relative error), shown in Fig. 2d), increases to a peak, after which it follows a steadily declining curve, which starts exactly at the point where the inflection is included in the fitting, i.e. when the ending day t_e of the fitting period depasses 35. The *GOF* has reached about 1.3% at t_e =49, corresponding to a linear correlation of 99.95% between simulations and observations. The logistic assymptotic value *L* in Fig. 2a) and the exponential rate *r* in Fig. 2c), as a function of t_e, have both stabilized, respectively near 97000 and 0.181 day^{-1}.

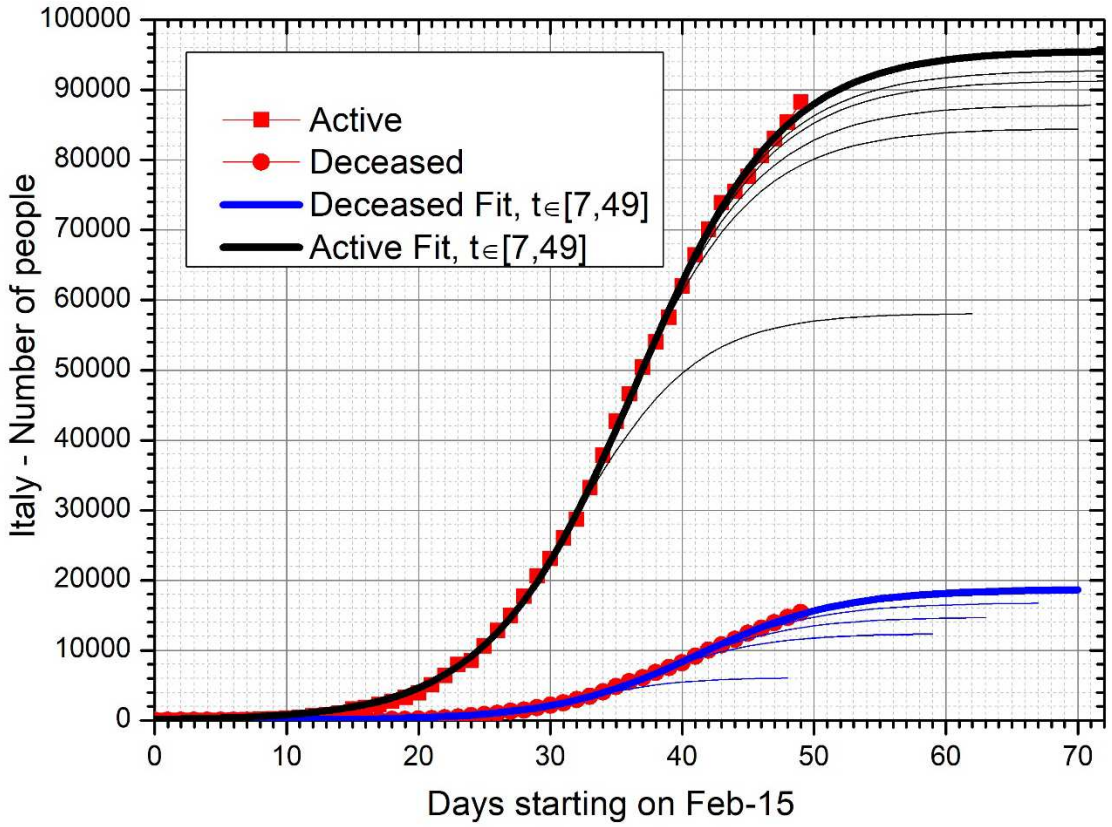


Fig. 1. Fig 1. | Active (red squares) and deceased (red circles) daily sampled declared values in Italy from t=0 (Feb-15) up to t=49 (Apr-04) and logistic fits using a fitting period starting at t=7 (Feb-22, when active = 79) and ending at t_e =33, 36, 39, 42, 45 (thin lines) and t_e=49 (solid lines). Fits stop at L-100 in ordinates. Line colors are black for active and blue for deceased.

With respect to the deceased curve (see Fig. 1), it may also be seen in Fig 2b) that the inflection point has not yet been reached, since the fitted value of T_i is still growing, even for $t_e > 35$. This is to be expected, since the deceased curve has an effective delay with

respect to the active curve. The logistic asymptotic value L has slowly risen up to ~18800 (Fig. 2a) at t_e=49 and r has decreased towards 0.184 day^{-1} (Fig. 2c), with a tendency to converge to the r value of the active curve as new outcomes enter in the fitting. This behaviour is consistent with the slow but steady increase of the inflection point for the deceased curve seen in Fig. 2b).

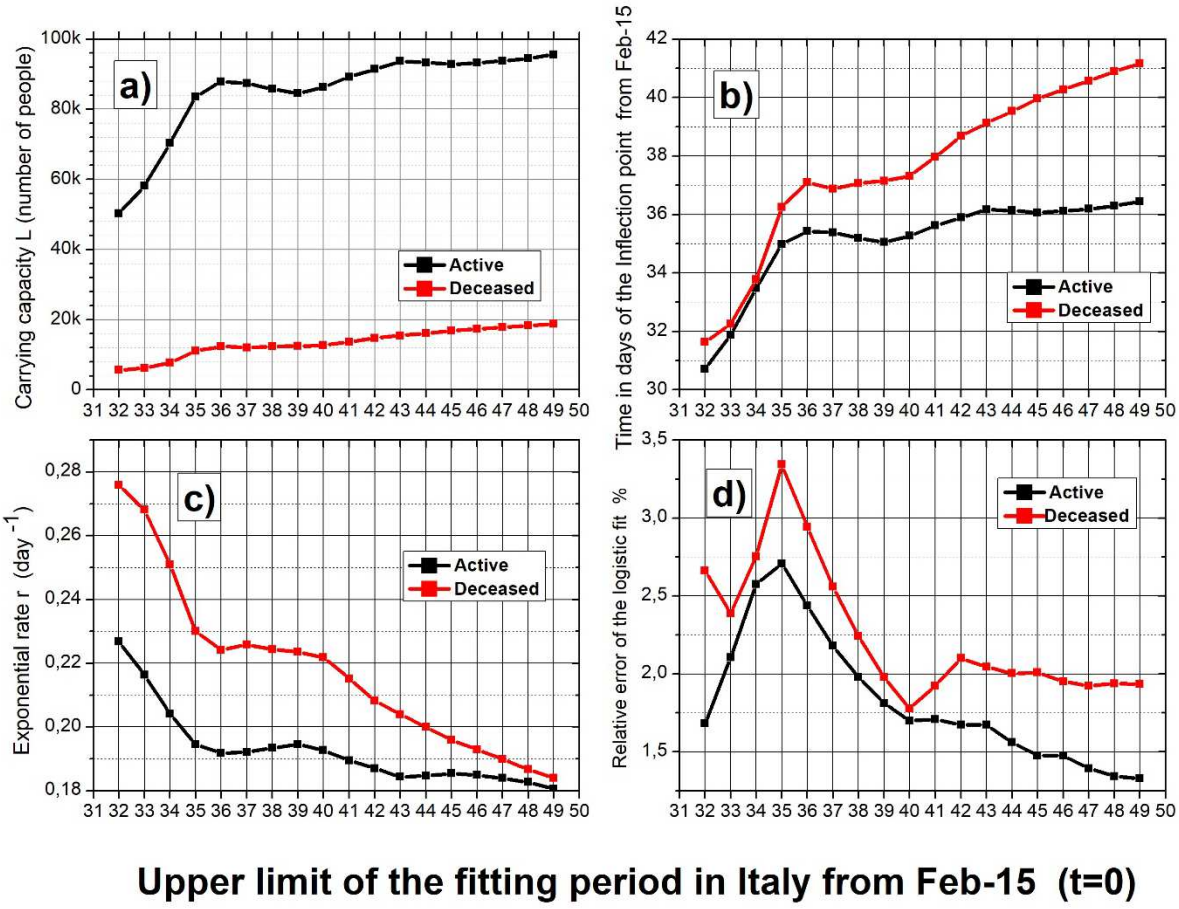


Fig. 2. Logistic parameters for Italy as a function of the fitting period ending day t_e in the following order: L (a) , T_i (b) and r (c) and correspondent relative error (d) (GOF) in % for the active (black) and deceased (red) curves.

In the case of Spain (Fig. 3), these methods reveal that an inflection point in the number of active cases has been reached in days 39-40 that is, 25-26 March, and in a very robust way as seen in Fig. 4b). The *GOF* of the logistic fit in Fig. 4d) decays sharply and monotonously since day 38, reaching 1.75% when the fitting is extended up to $t_e = 49$ (see Fig. 4d).

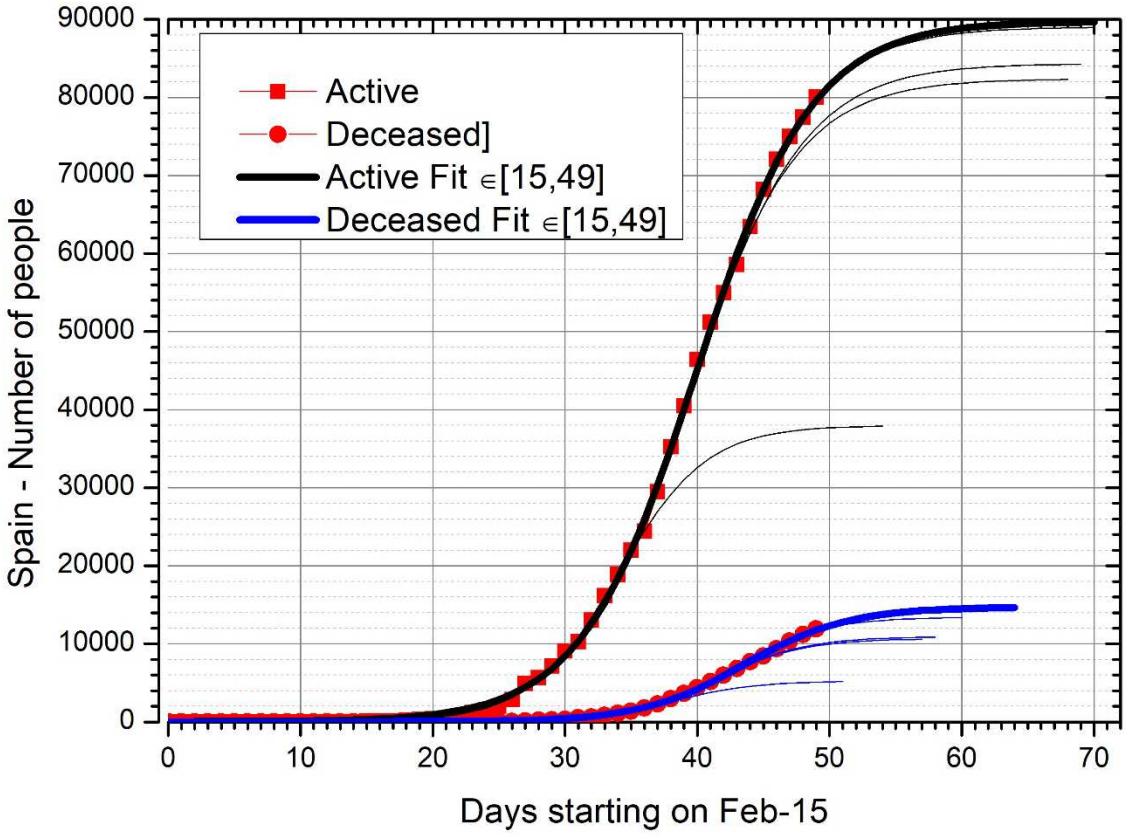
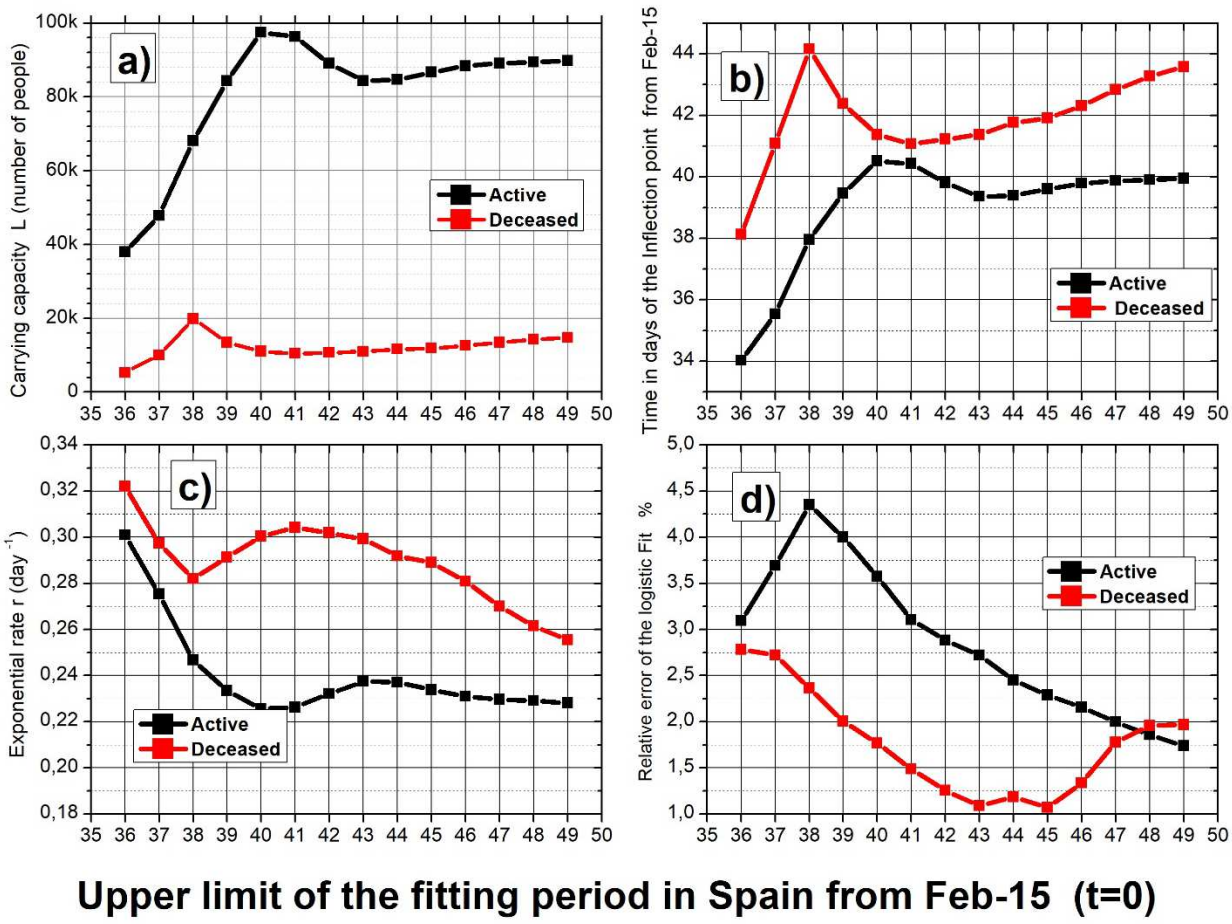


Fig. 3. Active (red squares) and deceased (red circles) daily sampled declared values in Spain from t=0 (Feb-15) up to t=49 (Apr-04) and logistic fits using a fitting period starting at t=15 (Mar-01, when active = 82) and ending at t_e =36,39,42,45 (thin lines) and t_e=49 (solid lines). Fits stop at *L*-100 in ordinates. Line colors are black for active and blue for deceased.

157

158 The deceased curve seems to show the same tendency, but in a less robust way: the *GOF*
159 has a relative minimum (Fig. 4d) when the fit ends at t_e =41, growing afterwards. This
160 indicates that it is still too early to confirm the existence of an inflection point in the
161 deceased curve.

162 For the active and deceased curves, the values of the asymptotic value *L* are converging to
163 89800 and 14800 respectively (see Fig. 4a) whereas the *r* values (see Fig. 4c) are stabilizing
164 at 0.23 day^{-1} for the active curve but still decreasing for the deceased curve, with values
165 around 0.26 day^{-1} respectively. This is another indication that the estimated inflection point
166 is still not very stable.



167

168 **Fig. 4.** Logistic parameters for Spain as a function of the fitting period ending day t_e in the

169 following order: L (a) , T_i (b) and r (c) and correspondent relative error (d) (GOF) in % for

170 the active (black) and deceased (red) curves.

171

172 **Time lag between onset of symptoms and death in COVID-19**

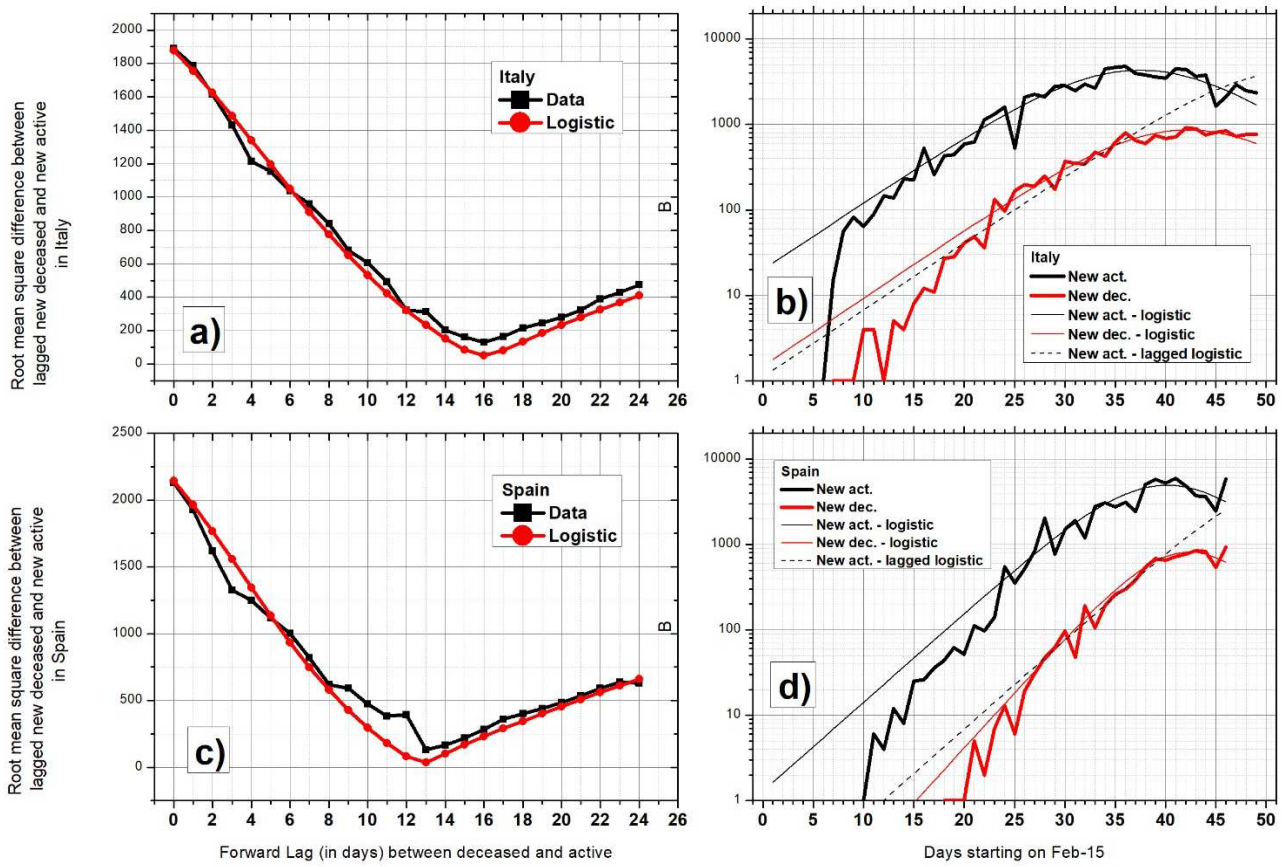
173 The graph of distance vs. time shift is shown for Italy (Fig. 5a) and Spain (Fig. 5c), both

174 using observed data and the logistic fit with the longest fitting period (t_e =49). Both curves

have a remarkable agreement, with minima lying in the same specific TLSD for each country.

In the case of exponential approximation τ gives the lag that makes the exponents of $I(t)$ and $D(t+\tau)$ equal, making the interpretation straightforward: in these conditions, the shifting time that produces the minimal distance between the two loglinear graphs of the daily new cases (Fig. 5b for Italy and Fig. 5d for Spain) is precisely the desired time lag τ between the onset of symptoms and death.

In Fig. 5b and Fig. 5d, we present the results of our computations using the available and public data of Johns Hopkins repository [1].



184

185 **Fig. 5.** Root mean square difference between the daily new active and lagged daily new

186 deceased curves for Italy (a) and Spain (c) with black symbols. The estimation obtained

187 from the daily increments of the logistic fit of active and deceased is given by red symbols.

188 b), for Italy) and d) for Spain, show the curves of daily new active (Solid black), and its

189 logistic approximation (Thin black), daily new deceased (Solid red), and its logistic

190 approximation (Thin red). The lagged daily new active by the logistic approximation, with

191 the lag of minimum curve distance is also plotted (Dashed thin black).

192

Our findings show that the distance between the curves is minimal when τ =16 days for Italy and when τ =13 days for Spain, with errors of about 1 day. These numbers agree remarkably well with the bounds mentioned in the literature [*4,20,21,22*]. The method is fairly straightforward, and in the presence of the appropriate data may be very efficiently adapted to obtain TLSD for any country.

Discussion

This work was performed during the evolution of the COVID-19 viral outbreak in Europe in 2020; thus, some of the predictions we perform, namely the asymptotic values of deceased people, will only be validated in the near future. Therefore, the predictive tools we present can only be taken as an accompanying diagnostic to help Healthcare policy- and decision-makers during the epidemic progress. Moreover, the results we obtain are only valid as long as the hypotheses of validity of the logistic approach remain applicable. For instance, if a second pandemic wave occurs, the logistic fit must be reset. Another limitation is the representativeness of the declared values of the actively infected.

In fact, the total of actives for many countries may be aggregate values from different heterogeneous regions corresponding to distinct local outbreaks, in different stages of the pandemic, with different contagion factors (dependent on e.g. geographical or sociological variables) and where eventually different lockdown measures were applied. This has happened very specifically in Italy, with an earlier outbreak in Lombardy and later regional outbreaks, leading to the Lombardy lockdown in March 8 and full lockdown in March 11.

The sum of different epidemiological variables corresponding to sources of infection in different stages of development may result in the heteroscedasticity (differing variance) of the summands totalizing the number of infected, which may prove to be a source of noise for global logistic modeling of the epidemic. However, for large enough populations like those studied in this paper, the heterogeneity factors will in fact become sufficiently diluted as not to perturb too much the logistic fit, as is evidenced by the stability of our computations for the cases of Italy and Spain. This fact allows thus for a robust estimation of the inflection point of the epidemic curve, as well as the corresponding estimations of TLSD, making our conclusions sound and reliable.

Unfortunately, using only the curves of infected and deceased, we have no information to estimate standard deviations of the time lag from onset of symptoms to death using only the logistic fit.

Conclusions

The evolution of the logistic fitting parameters, as well as the goodness-of-fit as far new recent observations are included during the pandemic, is a good diagnostic and early warning of the inflection point both in active and deceased epidemiologic curves. The estimation of the inflection point of epidemic curves through a smooth logistic curve is certainly more robust than that obtained from difference-based estimations of time derivatives, which are contaminated by noise.

Our results show that Italy has reached the inflection point of the active curve on days 35-36 (21-22 March) and Spain on days 39-40 (25-26 March). We note that full lockdown was

declared, respectively, on days 25 and 30. This is extremely good evidence that the period between full lockdown measures and reaching the inflection point is 10 days.

Our results may also be used to estimate the lag between onset of symptoms and death for both countries, which we find to be 16 days for the case of Italy and 13 days for the case of Spain. In fact, these lags may be immediately read off from Figs. 5b) and 5d), which are just the loglin plot of the daily step of observed data (new daily cases): in both cases the deceased curve is seen to be an almost perfect time-translation of the active curve by the respective lag, also plotted in Fig.5b and Fig. 5d. Both these data may be extremely useful as guidelines for controlling the epidemic in other countries in the course of the current pandemic.

List of abbreviations

COVID-19: Virus SARS-CoV-2

GOF: Goodness of Fitting

SIR: Epidemiologic model Susceptible-Infected-Recovered

SEIR: Epidemiologic model Susceptible-Exposed-Infected-Recovered

TLSD: Time lag between onset of symptoms and death

Appendix

Fitting of the Logistic Model

The active, deceased and recovered curves during the rising stage of an epidemic outbreak can be well modeled by a logistic or Richards growth model [3,4], obeying the differential equation

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{L}\right)$$

where r, L are fixed positive parameters for each of those groups. The solution is

$$N(t, \boldsymbol{\theta}) = \frac{L}{1 + a\, e^{-rt}}$$

where $\boldsymbol{\theta} = (L, a, r)'$ is the column vector of the solution parameters. The time evolving graph of $N(t, \boldsymbol{\theta})$ is a sigmoid curve reaching the asymptotic upper bound L. The constant r is the exponential growth rate and $a = \frac{L - N(0,\boldsymbol{\theta})}{N(0,\boldsymbol{\theta})} = e^{rT_i}$ is positive, depending on the initial condition $N(0, \boldsymbol{\theta})$. It also relates to the mid-point time $N(T_i, \boldsymbol{\theta}) = L/2$ where the sigmoid passes through an inflection point (null second time derivative).

We now provide the details on how to fit a logistic curve and its three parameters from a set of regularly sampled observed or estimated available numbers z_k at times $t_k = k\Delta\,, (k = 0, \dots, m)$, where Δ is the sampling time, taken as unit w.l.g. (e.g. 1 day). Those values may be affected by operational errors of systematic or random nature. Let us stress that if z_k is obtained from a representative finite sample, then it is reasonable to think of z_k as a constant fraction of the real number which satisfies a proportional logistic function with a different carrying capacity L. In order to get the misfit solution, we look for the absolute minimum of the cost function:

$$J(\boldsymbol{\theta}, m_i, m_f) = \sum_{k=m_i}^{m_f} (N(t_k, \boldsymbol{\theta}) - z_k)^2$$

where m_i is the initial fitting instant, which in the case of epidemiological curves is set to the time after which z_k is above a certain threshold $z_{min,}$ i.e. $z_k \geq z_{min}\,, k \geq m_i$ (e.g. $z_{min} = 100$).The cost function is therefore the sum of squared errors (SSE) or squared deviations between observations z_k and a given control solution $N(t_k, \boldsymbol{\theta})$. The absolute minimum of J in parameter space yields the optimal or misfit logistic. Given the nonlinear dependence on parameters, the optimal solution can only be reached by Monte-Carlo, genetic or gradient-descent-based methods.

Here we will use the Newton-Raphson iterative algorithm (). It starts from a first guess $\boldsymbol{\theta}_{fg} = \left(L_{fg}, a_{fg}, r_{fg}\right)^T$ of the parameters, where superscript T stands for transpose. They can be easily obtained from three equally spaced values z_n, z_{2n}, z_{3n}, considering that they are well fitted by a logistic. Its parameters are obtained as follows. Let us take $c_1 \equiv \frac{z_{3n}}{z_{2n}}, c_2 \equiv \frac{z_{2n}}{z_n}$ and $b_1 \equiv \frac{c_1 c_2 - 1}{c_1 (c_2 - 1)}, b_2 \equiv \frac{c_2 - 1}{c_1 (c_2 - 1)}$. Then $x \equiv e^{-n r_{fg}}$ satisfy a second-order equation: $x^2 - b_1 x + b_2 = 0$ with solution $x = \frac{1}{2}\left(b_1 - \sqrt{b_1^2 - 4b_2}\right)$. If x is complex, then choose the largest n possible: $3n \leq m$. Thus, we easily get $r_{fg} = -\log(x)/n$, $a_{fg} = \frac{c_1 - 1}{x^2 - c_1 x^3}$, $L_{fg} = \left(1 + x a_{fg}\right) z_n$.

Then, parameter steps in each iteration depend both on the cost-function gradient and the Hessian matrix. In order to keep its well-conditioning, both the parameters and time must be rescaled such that new parameters are O(1). The vector of rescaled parameters is $\boldsymbol{\eta} = (L', a', r')^T$ with $L' = L/L_{fg}$, $a' = a/a_{fg}$, $r' = r/r_{fg}$ and all the correspondent first-guess values are set to 1. The scaled time is $t' = t\, r_{fg}$ and the rescaled cost-function comes as:

$$\hat{J}\left(\boldsymbol{\eta}, m_i, m_f\right) = \sum_{k=m_i}^{m_f} \left(\widehat{N}(t_k', \boldsymbol{\eta}) - \frac{z_k}{L_{fg}} \right)^2$$

with

$$\widehat{N}(t', \boldsymbol{\eta}) = \frac{L'}{1 + a_{fg} a'\, e^{-r' t'}} = \frac{N(t, \boldsymbol{\theta})}{L_{fg}}$$

Let $\boldsymbol{\eta}_l$ the l-th parameter iteration with $\boldsymbol{\eta}_1$ set to a vector of ones. The iteration process is written as

$$\boldsymbol{\eta}_{l+1} = \boldsymbol{\eta}_l - \varepsilon \boldsymbol{H}\left(\hat{J}(\boldsymbol{\eta}_l)\right)^{-1} \boldsymbol{grad}\left(\hat{J}(\boldsymbol{\eta}_l)\right),$$

where $\boldsymbol{grad}(\hat{J}(\boldsymbol{\eta}_l))$ and $\boldsymbol{H}\left(\hat{J}(\boldsymbol{\eta}_l)\right)$ are the gradient and Hessian matrix of $\hat{J}$ with respect to the parameters evaluated at $\boldsymbol{\eta}_l$ and $\varepsilon \in]0,1]$, preferentially small (e.g. $\varepsilon = 0.005$). Expressions for the derivatives may be automatically obtained in https://www.wolframalpha.com/.

If the first guess lies in the attraction basin of a minimum, then $\hat{J}(\boldsymbol{\eta}_l) - \hat{J}(\boldsymbol{\eta}_{l+1}) > 0$ and the process is iterated, stopping at $\boldsymbol{\eta}_{final}$ when consecutive cost-function differences or the gradient norm is less than a certain tolerance ε_J (e.g. the rounding error) or the maximum number of iterations is reached. At the optimal solution, its sensitivity to small changes in the parameters may be obtained by diagonal terms of $\boldsymbol{H}\left(\hat{J}(\boldsymbol{\eta}_{final})\right)^{-1}$. Some first guesses must be tried, choosing the solution with the least cost function.

Goodness-of-fitting

The goodness-of-fitting (*GOF*) is obtained as the root square of the ratio between the SSE at the optimal solution and the sum of squares of the solution.

$$GOF = \left(\frac{J(\boldsymbol{\theta}_{final}, m_i, m_f)}{\sum_{k=m_i}^{m_f} z_k^2} \right)^{1/2}$$

Delay between active and deceased new cases

We provide a measure of similarity between the curve of new active cases and the curve of new deceased cases. Let $I(t)$ and $D(t)$ the active and deceased curve at time t and correspondent daily variations denoted as $I'(t) \equiv I(t) - I(t-1)$ and $D'(t) \equiv D(t) - D(t-1)$ respectively. The logistic fit of $I(t)$ is $I_l(t)$ and of $D(t)$ is $D_l(t)$. The daily variations by the logistic approach are then approximated by ${I_l}'(t) \equiv I_l(t) - I_l(t-1)$ and ${D_l}'(t) \equiv D_l(t) - D_l(t-1)$ respectively.

We perform this computation in the period ranging between the time correspondent to a minimum threshold of actives (e.g. 100) and the point of inflection in the curve of actives.

We then compute the root mean average squared difference (or Euclidean distance) between the curve $I'(t)$ and the lagged curve $D'(t+\tau)$ for a positive τ ($Dist(\tau)$):

$$Dist(\tau) = \sqrt{\frac{1}{T_i - m_i + 1} \sum_{t=m_i}^{T_i} \left(I'(t) - D'(t+\tau) \right)^2}$$

Equally, we compute the same distance using the logistic fit ($Dist_l(\tau)$), where we simply replace $I'(t)$ by ${I_l}'(t)$ and $D'(t+\tau)$ by $D_l'(t+\tau)$.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

All data is available in the main text or the supplementary materials and on public repositories [*1,2*].

Competing interests

Authors declare no competing interests.

Funding

C.P. was supported by FCT through project UIDB/50019/2020 – IDL.

J.B. *was supported by National Funding from FCT - Fundação para a Ciência e a Tecnologia, under the project: UIDB/04561/2020*. H.M.O. was partially supported by FCT/Portugal through projects UID/MAT/04459/2019 and UIDB/04459/2020;

Authors' contributions

All authors have contributed for the manuscript text, analyses and revisions. Jorge Buescu has conceived the experiments and main guidelines. Carlos Pires has implemented the fitting and preparation of figures. Henrique Oliveira has computed and discussed the delay between active and deceased;

Acknowledgements

We thank the invaluable and fruitful discussions with Altamiro da Costa Pereira, director of the Faculty of Medicine of Oporto University in the preparation of this article.

Authors' information

Not applicable

References

1. Coronavirus Resource Center. *John Hopkins University Medicine* (2020). Available at https://coronavirus.jhu.edu/. (Accessed: April 2 2020).

2. Worldometer - Confirmed Cases and Deaths by Country, Territory, or Conveyance (2020). Available at https://www.worldometers.info/coronavirus/#countries. (Accessed: April 2 2020).

3. M. Martcheva. Introduction To Mathematical Epidemiology. Texts in applied mathematics, vol. 61. (Springer, New York 2015)

4. F. J. Richards "A flexible growth function for empirical use." *Journal of experimental Botany* 10.2 (1959): 290-301.

5. K. Wu, D. Darcet, Q. Wang & D. Sornette. Generalized logistic growth modeling of the COVID-19 outbreak in 29 provinces in China and in the rest of the world. Preprint at https://www.medrxiv.org/content/10.1101/2020.03.11.20034363v1, (2020) doi: https://doi.org/10.1101/2020.03.11.20034363

6. D. G. Kleinbaum, L. K. Lawrence & L. E. Chambless. Logistic regression analysis of epidemiologic data: theory and practice. *Communications in Statistics-Theory and Methods* 11.5 (1982): 485-547.

7. S. Sperandei. Understanding logistic regression analysis. *Biochemia medica: Biochemia medica* 24.1 (2014): 12-18. doi: 10.11613/BM.2014.003

8. J. Bouyer. Logistics regression in epidemiology. II. *Revue d'epidemiologie et de sante publique* 39.2 (1991): 183-196.

9. G. Chowell. Fitting dynamic models to epidemic outbreaks with quantified uncertainty: a primer for parameter uncertainty, identifiability, and forecasts. *Infectious Disease Modelling* 2.3 (2017): 379-398.

10. S. Endre & D. F. Mayers. *An introduction to numerical analysis*. (Cambridge university press, 2003). ISBN 0-521-00794-1.

11. B. Yard. Nonlinear Parameter Estimation, (Academic Press, New York, 1974.)

12. D. M. Bates & D. G. Watts. Nonlinear Regression Analysis and its Applications, (Wiley, New York, 1988).

13. J. E. Dennis, D. M. Gay &, R. E. Welsch. An adaptative nonlinear least squares, in The State of the Art in Numerical Analysis, D. Jacobs, editor, (Academic Press, New York, 1977), pp. 269-312.

14. H. T. Banks, S. Hu & W. C. Thompson. **Modelling and inverse problems in the presence of uncertainty (**CRC Press, 2014).

15. L. J. Allen, F. Brauer, P. Van den Driessche. & J. Wu. *Mathematical epidemiology* (Vol. 1945). (Springer Berlin 2008).

16. Z-Q. Xia, J. Zhang, Y-K. Xue, G-Q. Sun & Z. Jin. Modeling the Transmission of Middle East Respirator Syndrome Corona Virus in the Republic of Korea. PLoS ONE 10(12): e0144778 (2015). doi:10.1371/

journal.pone.0144778

17. L. Peng, W. Yang, D. Zhang, C. Zhuge & L. Hong. Epidemic analysis of COVID-19 in China by dynamical modeling. *arXiv preprint arXiv:2002.06563*. (2020).

18. 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. 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* (2020). ISSN 1473-3099, https://doi.org/10.1016/S1473-3099(20)30144-4.

19. Y-C. Chen, P-E. Lu & C-S. Chang. A Time-dependent SIR model for COVID-19

arXiv:2003.00122v1 (28 Feb 2020)

20. T. W. Russell, J. Hellewell, S. Abbott, C. I. Jarvis, K. van Zandvoort, CMMID nCov working group, S. Flasche, R. M. Eggo, W. J. Edmunds & A. J. Kucharski. Using a delay-adjusted case fatality ratio to estimate under-reporting. Center for Mathematical Modelling and Infectious Diseases. *London School of Hygiene & Tropical Medicine. Available at https://cmmid.github.io/topics/covid19/severity/global_cfr_estimates.html. (Accessed: April 2 2020)*

21. I. Dorigatti, L. Okell, A. Cori, N. Imai, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunubá, G. Cuomo-Dannenburg, R. FitzJohn, H. Fu, K. Gaythorpe, A. Hamlet, W. Hinsley, N. Hong, M. Kwun, D. Laydon, G. Nedjati-Gilani, S. Riley, S. van Elsland, E. Volz, H. Wang, Y.(R.) Wang, C. Walters, X. Xi, C. Donnelly, A. Ghani & N. Ferguson. Report 4: Severity of 2019-novel coronavirus (nCoV), WHO Collaborating Centre for Infectious Disease Modelling , RC Centre for Global Infectious Disease Analysis, Abdul Latif Jameel

Institute for Disease and Emergency Analytics (J-IDEA), Imperial College London. Available at https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-4-severity-of-covid-19/. (Accessed: April 2 2020).

22. N. M. Linton, T. Kobayashi, Y. Yang, K. Hayashi, A. R. Akhmetzhanov, S. Jung, B. Yuan, R. Kinoshita & H. Nishiura. Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data. *Journal of clinical medicine* 9, no. 2 (2020): 538.

23. X. Yang, Y. Yu, J. Xu, H. Shu, J. Xia, H. Liu, Y. Wu, L. Zhang, Z. Yu, M. Fang, T. Yu, Y. Wang, S. Pan, X. Zou, S. Yuan, & Y. Shang. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *The Lancet Respiratory Medicine* (2020). doi: 10.1016/S2213-2600(20)30079-5

24. D. Baud, X. Qi, K. Nielsen-Saines, D. Musso, L. Pomar, & G. Favre. Real estimates of mortality following COVID-19 infection. *The Lancet Infectious Diseases* (2020).

Figures

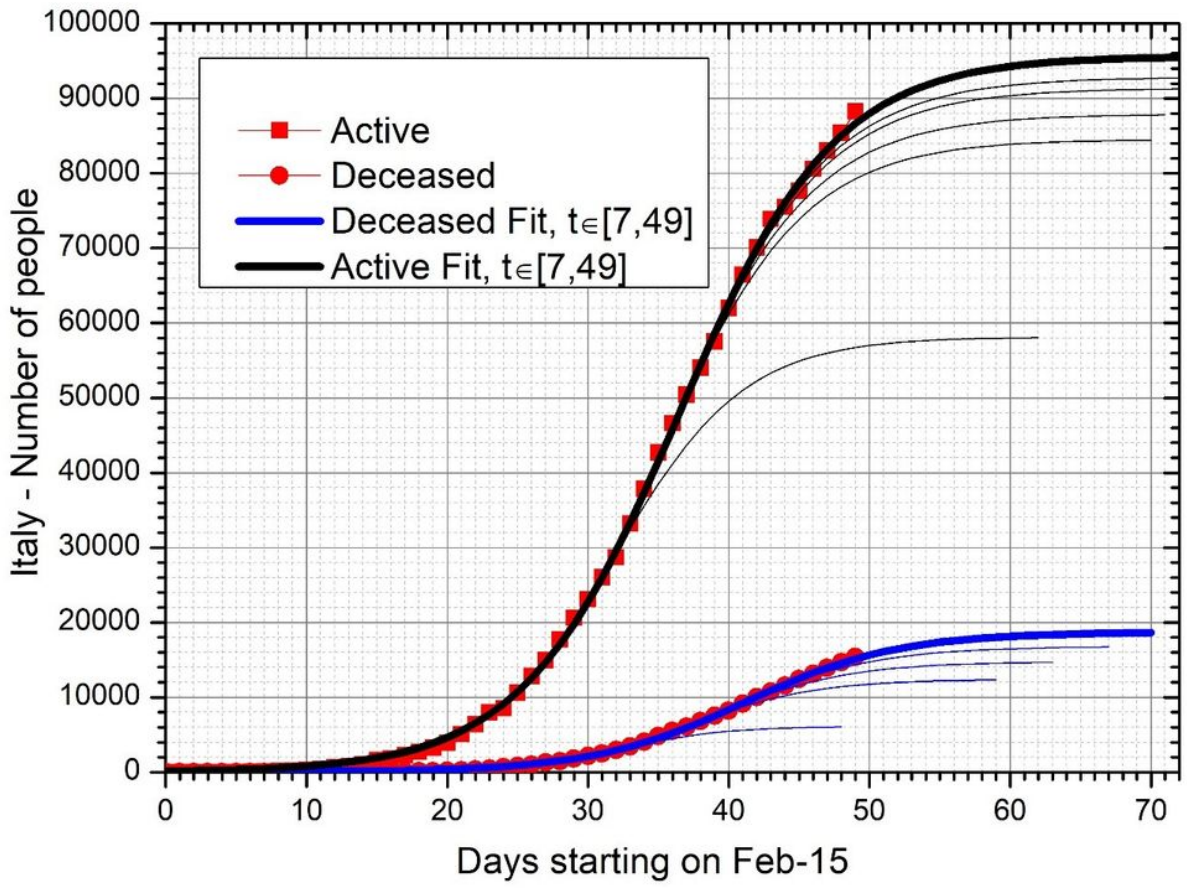


Figure 1

Active (red squares) and deceased (red circles) daily sampled declared values in Italy from t=0 (Feb-15) up to t=49 (Apr-04) and logistic fits using a fitting period starting at t=7 (Feb-22, when active = 79) and ending at t_e =33, 36, 39, 42, 45 (thin lines) and t_e=49 (solid lines). Fits stop at L-100 in ordinates. Line colors are black for active and blue for deceased.

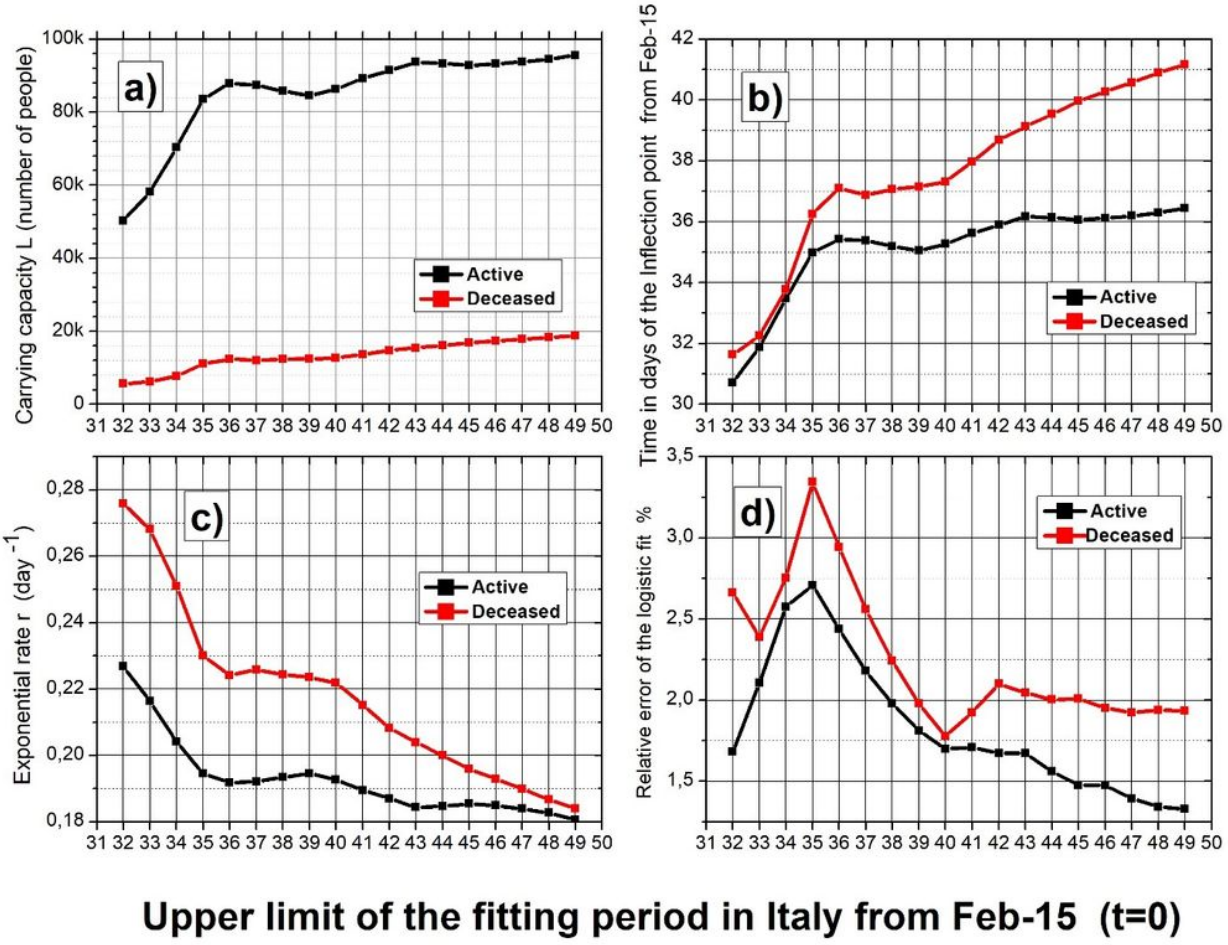


Figure 2

Logistic parameters for Italy as a function of the fitting period ending day te in the following order: L (a) , Ti (b) and r (c) and correspondent relative error (d) (GOF) in % for the active (black) and deceased (red) curves.

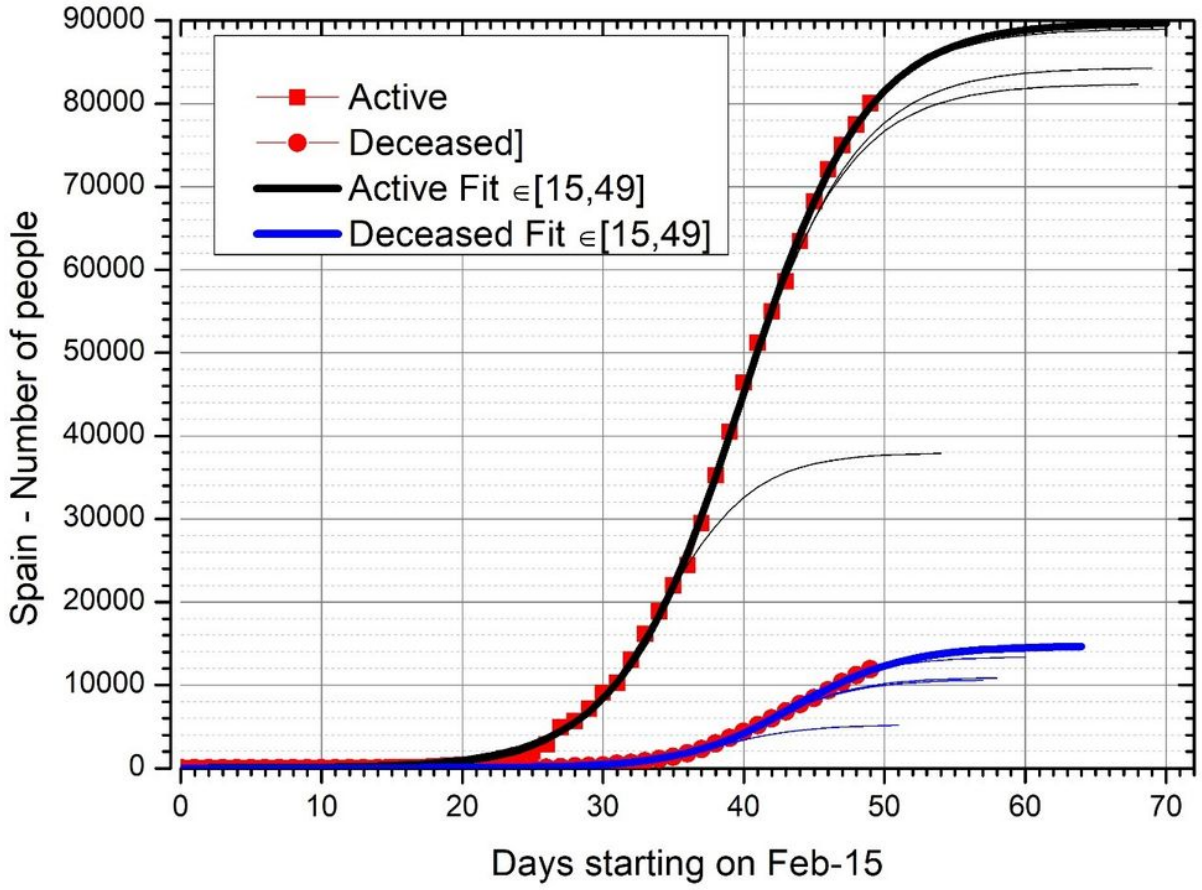


Figure 3

Active (red squares) and deceased (red circles) daily sampled declared values in Spain from t=0 (Feb-15) up to t=49 (Apr-04) and logistic fits using a fitting period starting at t=15 (Mar-01, when active = 82) and ending at te =36,39,42,45 (thin lines) and t_e=49 (solid lines). Fits stop at L-100 in ordinates. Line colors are black for active and blue for deceased.

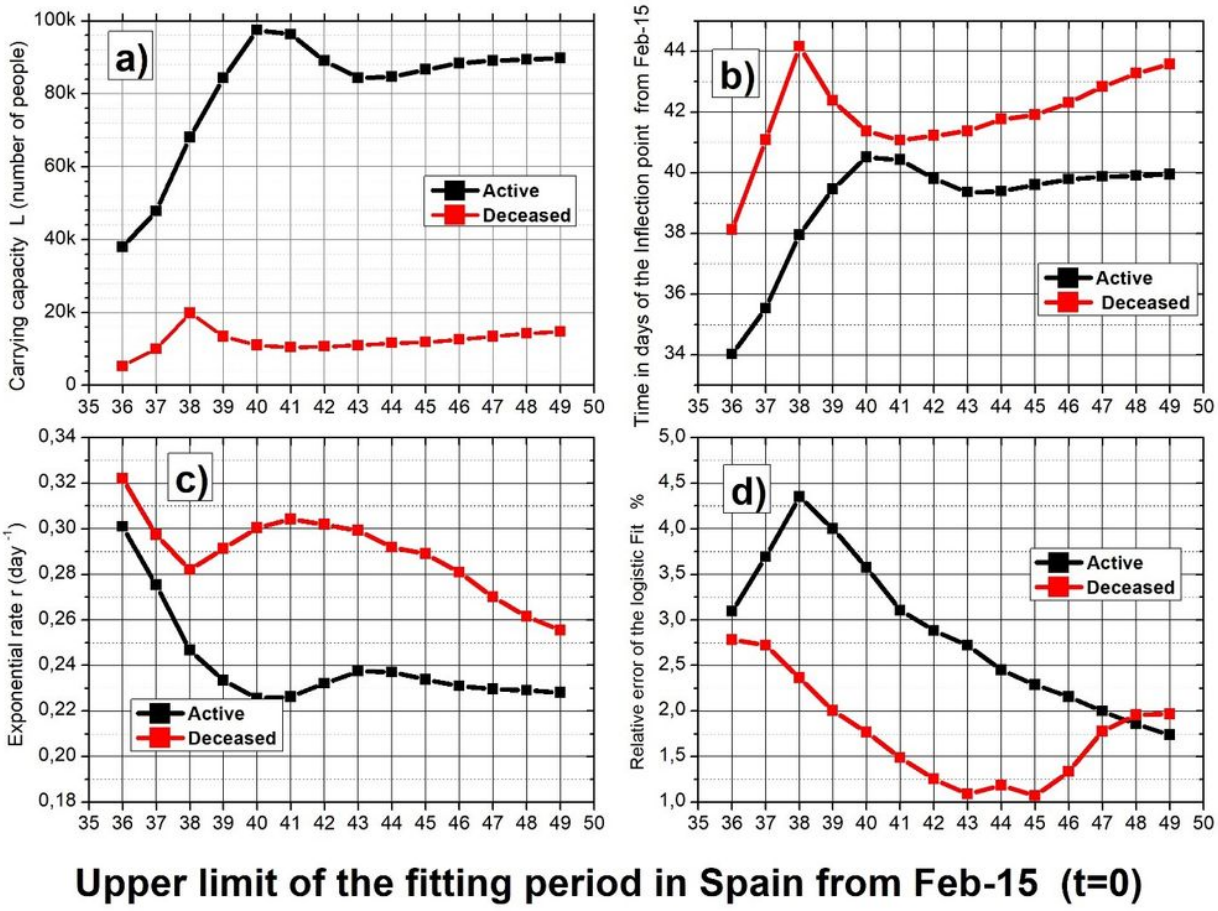


Figure 4

Logistic parameters for Spain as a function of the fitting period ending day t_e in the following order: L (a) , Ti (b) and r (c) and correspondent relative error (d) (GOF) in % for the active (black) and deceased (red) curves.

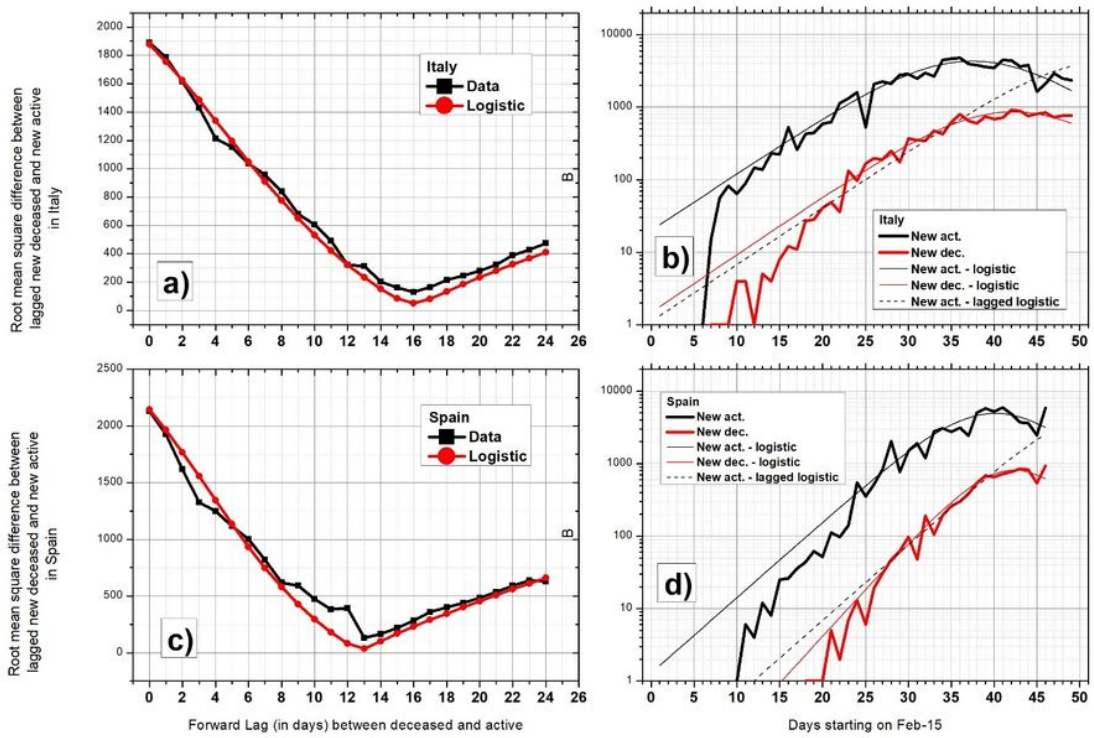


Figure 5

Root mean square difference between the daily new active and lagged daily new deceased curves for Italy (a) and Spain (c) with black symbols. The estimation obtained from the daily increments of the logistic fit of active and deceased is given by red symbols. b), for Italy) and d) for Spain, show the curves of daily new active (Solid black), and its logistic approximation (Thin black), daily new deceased (Solid red), and its logistic approximation (Thin red). The lagged daily new active by the logistic approximation, with the lag of minimum curve distance is also plotted (Dashed thin black).