

1 On the progression of COVID19 in Portugal: A comparative analysis of active cases using non-linear
2 regression

3

4 Ana Milhinhos^{1,2}, Pedro M. Costa^{3,*}

5

6 ¹Green – IT Research Unit, Instituto de Tecnologia Química e Biológica António Xavier, Universidade
7 Nova de Lisboa (ITQB NOVA), Av. da República, 2780-157, Oeiras, Portugal.

8

9 ²Biosystems and Integrative Sciences Institute (BioISI), Faculty of Sciences, Universidade de Lisboa,
10 Campo Grande, 1749-016 Lisbon, Portugal.

11

12 ³UCIBIO – Research Unit on Applied Molecular Biosciences, Departamento de Ciências da Vida,
13 Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

14

15

16 Word count: 799

17

* To whom correspondence should be addressed

pmcosta@fct.unl.pt

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Tel. +351 212 948 300 ext. 11103

18 **Abstract**

19

20 Portugal has been portrayed as a relatively successful case in the control of the COVID-19's March 2020
21 outbreak in Europe due to the timely confinement measures taken. As other European Union member states,
22 Portugal is now preparing the phased loosening of the confinement measures, starting in the beginning of
23 May. Even so, the current data, albeit showing at least a reduction in infection rates, renders difficult to
24 forecast scenarios in the imminent future. Using South Korea data as scaffold, which is becoming a
25 paradigmatic case of recovery following a high number of infected people, we fitted Portuguese data to
26 biphasic models using non-linear regression and compared the two countries. The results, which suggest
27 good fit, show that recovery in Portugal can be much slower than anticipated, with a very high percentage
28 of active cases (over 50%) remaining still active even months after the projected end of mitigation measures.
29 This, together with the unknown number of asymptomatic carriers, may increase the risk of a much slower
30 recovery if not of new outbreaks. Europe and elsewhere must consider this contingency when planning the
31 relief of containment measures.

32

33

34 **Key-words**

35

36 Coronavirus; Modelling; Statistical forecasting; Non-linear estimation; European Union

37

38 The first documented case of COVID-19 infection in Portugal dates March 2, 2020. Motivated by the rapid
39 progression in other countries, especially in neighbouring Spain (see Kinross et al., 2020; Spiteri et al.,
40 2020), the country moved swiftly to control dissemination by shutting down many public services and
41 imposing strict confinement measures (see for instance Mahase, 2020). These date from mid-March and,
42 by the time the present work is prepared, the Portuguese government and competent health authorities are
43 planning the phased cessation of these measures, in alignment with the European Union. It has been
44 consensual that the spreading of COVID-19 in Portugal, with respect to number of infected people, fatalities
45 and ICU internments, is reaching a plateau, making it a potential case of success, as by mid-April fatalities
46 were kept below 1,000 and the healthcare system did not attain saturation. Now is a crucial time to know
47 what to expect from the progression of the disease in the country and how safe it is to begin the relief of
48 confinement measures. However, the available data is still insufficient to draw solid forecasts even on short-
49 term. Indeed, at this stage, epidemiological SIR ('susceptible', 'infected' and 'recovered') models are
50 difficult to produce in Portugal and elsewhere.

51

52 Lessons can be learned from the few countries already clearing the pandemic. The Republic of Korea is a
53 key case study not just due to the overall positive progress but also because the country implemented strict
54 confinement measures, imposed timely limitations to in-bound travelling and closed public services like
55 schools. Also, South Korea endured a high number of total infections (which offers statistical significance),
56 albeit a relatively low mortality rate, estimated at 0.9% by mid-March, when cases totalled almost 8,000,
57 according to the Republic of Korea COVID-19 National Emergency Response Center (2020). Still, the
58 basic reproduction number (R_0) has been estimated at 1.5 ± 0.1 (Shim et al., 2020), therefore within the
59 magnitude of the influenza outbreak in 1918 (Ferguson et al., 2006). Consequently, there are significant
60 similarities between countries albeit likely differences in public behaviour, awareness or susceptibility. We
61 therefore aimed at modelling the progression of active cases in Portugal (up to April 19) by means of non-
62 linear regression using Korean data as scaffold.

63

64 We used a four-parameter log-logistic model to estimate the maximum number of infections in Portugal,
65 which, according to the current data, will surpass 25,500. The same model yielded a maximum of about
66 1,000 daily fatalities by day 116 (June 24) since the first cases were reported (March 2), a date that
67 landmarks the beginning of the outbreak in Portugal (Fig. 1A). The current mortality rate is 3.5 % (as per
68 April 19), well below Italy, with over 10% (Rubino et al., 2020). One of the most positive signs for COVID-
69 19 control in Portugal has been considered the reduced percentage of daily cases (Fig. 1B). This information
70 must, nonetheless, be interpreted with caution as reported new cases are highly variable, in part due to
71 increased testing. Active cases were then fitted to a five-parameter log-Gaussian distribution, as described
72 by Martin-Betancor et al. (2015), a biphasic asymmetric model. In fact, the recovery rates in Portugal are
73 low (only 610 cases by April 19), seemingly accordant to reports from elsewhere. At this stage, Portugal as
74 yet to reach the peak of active cases, which means that data cannot be fitted to the descending phase of the
75 curve. Still, the fit was nearly perfect to the ascending phase (Fig. 1C). Moreover, Korean data also fitted
76 the same model perfectly, yielding, as expected, slower recovery than infection rates (Fig. 1D). By
77 juxtaposing the two models and expanding them to a full-year timeframe, the differences between the two

78 countries become evident (Fig. 1E). At day 50, South Korea reported 3591 active cases whereas the model
79 estimated 3653 cases (half of the 7307 projected maximum), with the real maximum being 7293, which,
80 again, shows the good fit of the model. Portugal may only reach 50% of recoveries after 140 days. The
81 different shapes of curves, reflected in differential parametrization of models (Table 1) should reflect not
82 only the number of cases but also different rates of recovery.

83

84 Even though Korean data validates the model, caution is mandatory when interpreting the Portuguese
85 model, as data are incomplete and model parameters are sure to change in time, either accelerating or
86 slowing recovery, depending on the success of mitigation measures and on how the loosening of confinement
87 policies, projected to begin in May, will proceed. It is clear, though, that recovery will be long. With 50%
88 cases still active by July, the risks of new peaks are high, furthermore considering the high percentage of
89 untraced asymptomatic carriers of COVID-19 (Yu et al., 2020). To this we must add the fact that persistence
90 of the virus increases the odds of mutation.

91

92

93 **Acknowledgements**

94

95 This work was supported by the Applied Molecular Biosciences Unit (UCIBIO) which is financed by
96 national funds from 'Fundação para a Ciência e a Tecnologia', FCT (UIDB/04378/2020). AM is supported
97 by CEEC/IND/00175/2017 contract, BioISI (UIDB/04046/2020 and UIDP/04046/2020) and Green-IT
98 (UID/Multi/04551/2013) grants from FCT.

99

100

101 **Conflicts of interest**

102

103 The authors declare no conflicts of interest

104

105

106 **References**

107

108 COVID-19 National Emergency Response Center, Epidemiology and Case Management Team, Korea
109 Centers for Disease Control and Prevention. (2020). Coronavirus disease-19: The first 7,755 cases in the
110 Republic of Korea. *Osong Public Health Res. Perspect.* 11, 85–90 (doi: 10.24171/j.phrp.2020.11.2.05).

111

112 Ferguson, N.M., Cummings, D.A., Fraser, C., Cajka, J.C., Cooley, P.C., Burke, D.S. (2006). Strategies for
113 mitigating an influenza pandemic. *Nature* 442, 448–452 (doi: 10.1038/nature04795).

114

115 Ihaka, R., Gentleman, R. (1996). R: A language for data analysis and graphics. *J. Comput. Graph. Stat.* 5,
116 299–314 (doi: 10.1080/10618600.1996.10474713).

117

- 118 Kinross, P., Suetens, C., Gomes Dias, J., et al. (2020). Rapidly increasing cumulative incidence of
119 coronavirus disease (COVID-19) in the European Union/European Economic Area and the United
120 Kingdom, 1 January to 15 March 2020. *Euro Surveill.* 25, 2000285 (doi: 10.2807/1560-
121 7917.ES.2020.25.11.2000285).
- 122
- 123 Mahase E. (2020). Covid-19: Portugal closes all medical schools after 31 cases confirmed in the country.
124 *BMJ.* 368, m986 (doi: 10.1136/bmj.m986).
- 125
- 126 Martin-Betancor, K., Ritz, C., Fernández-Piñas, F., Leganés, F., Rodea-Palomares, I. (2015). Defining an
127 additivity framework for mixture research in inducible whole-cell biosensors. *Sci. Rep.* 5, 17200 (doi:
128 10.1038/srep17200).
- 129
- 130 Rubino, S., Kelvin, N., Bermejo-Martin, J.F., Kelvin, D. (2020) As COVID-19 cases, deaths and fatality
131 rates surge in Italy, underlying causes require investigation. *J. Infect. Dev. Ctries.* 14, 265–267 (doi:
132 10.3855/jidc.12734).
- 133
- 134 Shim, E., Tariq, A., Choi, W., Lee, Y., Chowell, G. (2020). Transmission potential and severity of COVID-
135 19 in South Korea. *Int. J. Infect. Dis.* 93, 339–344 (doi: 10.1016/j.ijid.2020.03.031).
- 136
- 137 Spiteri, G., Fielding, J., Diercke, M., et al. (2020). First cases of coronavirus disease 2019 (COVID-19) in
138 the WHO European Region, 24 January to 21 February 2020. *Euro. Surveill.* 25, 2000178 (doi:
139 10.2807/1560-7917.ES.2020.25.9.2000178)
- 140
- 141 Yu, X., Yang, R. (2020). COVID-19 transmission through asymptomatic carriers is a challenge to
142 containment. *Influenza Other Respir. Viruses* (doi: 10.1111/irv.12743. doi:10.1111/irv.12743).
- 143

144

Tables

145

146 **Table 1.** Summary of parameter estimates for the fitting of active COVID-19 cases in Portugal and the
 147 Republic of Korea by April 19, 2020. The fitted model consisted of a five-parameter log-Gaussian
 148 distribution, a biphasic response model. Parameters were obtained through least squares estimation, using
 149 the package ‘drc’. All statistics were performed using R 3.5 (Ihaka and Gentleman, 1996). The Portuguese
 150 data (March 2 – April 19) were compiled from the official daily reports on COVID-19 provided by the
 151 General Directorate for Health, available at [https://covid19.min-saude.pt/ponto-de-situacao-atual-em-](https://covid19.min-saude.pt/ponto-de-situacao-atual-em-portugal/)
 152 [portugal/](https://covid19.min-saude.pt/ponto-de-situacao-atual-em-portugal/). Data from South Korea (February 15 – April 19) were retrieved from Worldometer
 153 (<https://www.worldometers.info/coronavirus/>). Data is provided in Supplementary information.

	Parameter	Estimate	Standard error	t-value	p-value
Portugal					
	<i>b</i>	0.60	1.53E-01	3.9315	0.0002885
	<i>c</i>	51.52	4.76E+01	1.0822	0.2849116
	<i>d</i> ¹	21,559.00	1.45E+03	14.9038	< 2.2E-16
	<i>e</i>	69.50	1.30E+01	5.3548	2.799E-06
	<i>f</i>	2.40	3.98E-01	6.0222	2.894E-07
Republic of Korea					
	<i>b</i>	0.48	1.49E-02	32.185	<2E-16
	<i>c</i>	-66.75	1.33E+02	-0.502	0.6176
	<i>d</i> ¹	7,306.07	1.25E+02	58.253	<2E-16
	<i>e</i>	27.69	1.92E-01	144.09	<2E-16
	<i>f</i>	1.71	9.64E-02	17.706	<2E-16

154 ¹ parameter *d* is the predicted maximum

155

156

Figure captions

157

158 **Fig. 1.** An overview of the evolution of COVID-19 in Portugal from March 1 (day 0) to April 19 (day 49)
159 2020. A) Cumulative number of deaths fitted to a log-logistic model. Scale was extended to highlight the
160 quality of fit and predicted asymptotic limit ($\approx 1,000$ deaths). B) A simple log-linear regression for the
161 percentage of daily new cases (infected subjects) relatively to cumulative new cases. C) Total active cases
162 (i.e. total cases excluding deaths and recoveries) fitted to a log-Gaussian (asymmetric) model with an
163 estimated maximum at $\approx 21,500$ cases, highlighting the near perfect fit to the growth phase of the model.
164 D) Active cases reported in the Republic of Korea between February 15 and April 19 fitted to log-Gaussian
165 Model, as previous. The South Korean scenario already has sufficient data to fit both growth and decrease
166 phases, again yielding a near perfect fit. E) Juxtaposition of predicted models (scaled to a full year from the
167 first day of reported cases) for Portuguese and Korean data (log-Gaussian non-linear regression). The
168 models highlighting maxima and half-maximal estimates (50% of cases recovered). Whereas south Korea
169 already surpassed the estimate (as day 50 corresponds to April 4), in Portugal, day 140 means July 17. The
170 shaded areas indicate 95% confidence intervals around the predicted model. Actual observations are
171 juxtaposed to the models (\bullet). The R^2 goodness-of-fit statistic means quadratic Spearman's *Rho*.

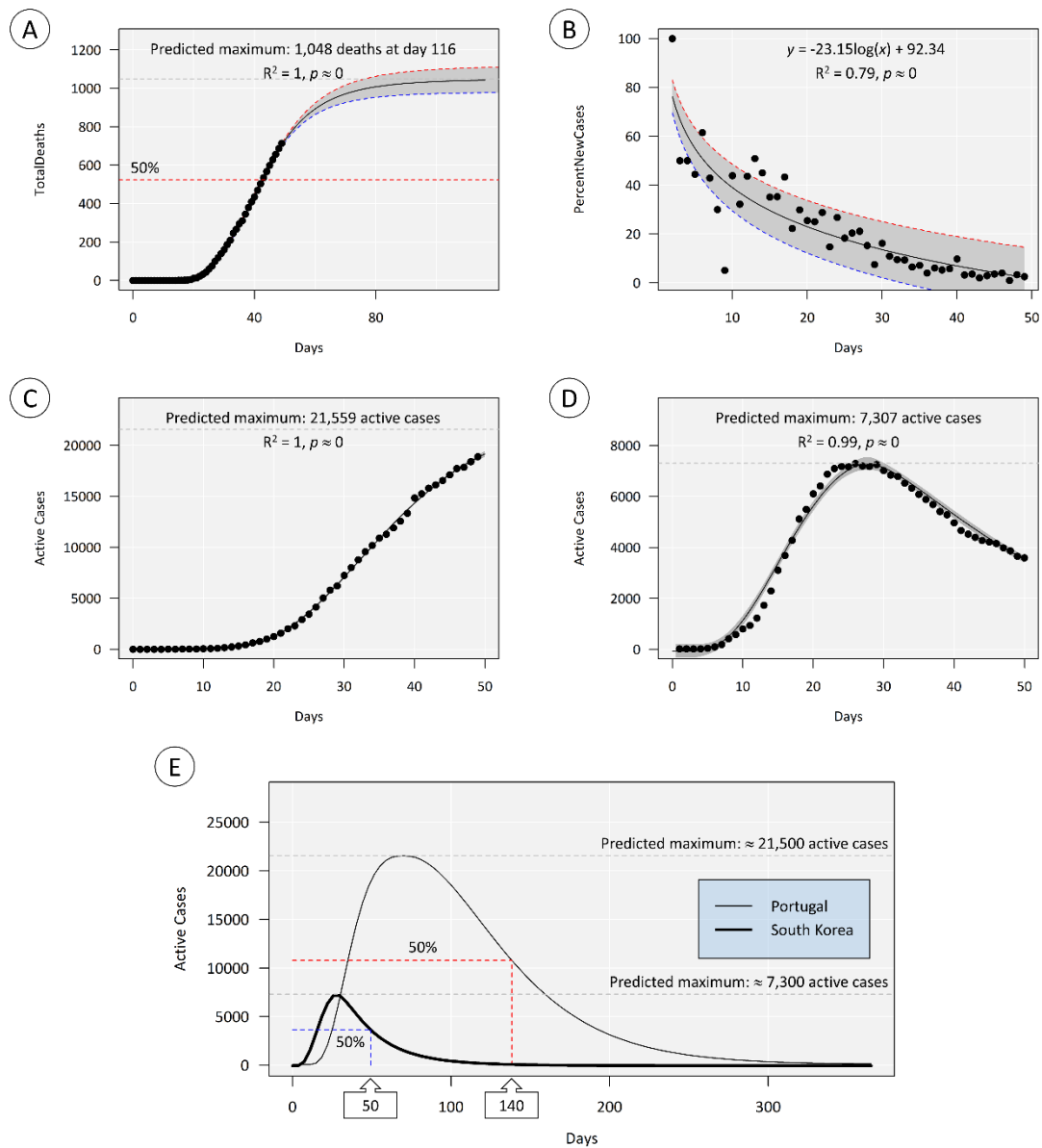
172

173

174

Fig. 1

175



176

177