

1 **Near- and forecasting the SARS-COV-2 epidemic requires a global view and multiple methods**

2 Daren Austin PhD ARCS DIC

3 Senior Fellow, Clinical Pharmacology Modelling & Simulation

4 GlaxoSmithKline Research and Development

5 Email: daren.j.austin@gsk.com

6 Phone: +447712 670097

7 09/APR/2020

8 **Abstract**

9 Conventional epidemiological models require estimates of important parameters including
10 incubation time and case fatality rate that may be unavailable in the early stage of an epidemic. For
11 the ongoing SARS-COV-2 epidemic, with no previous population exposure, alternative prediction
12 methods less reliant on assumptions may prove more effective in the near-term. We present three
13 methods used to provide early estimates of likely SARS-COV-2 epidemic progression. During the first
14 stage of the epidemic, growth rate charts revealed the UK, Italy and Spain as outliers, with
15 differentially increasing growth of deaths over cases. A novel data-driven time-series model was
16 then used to near-cast 7-day future cases and deaths with much greater precision. Finally, an
17 epidemio-statistical model was used to bridge from near-casting to forecasting the future course of
18 the global epidemic. By applying *multiple* approaches to global SARS-COV-2 data, coupled with
19 mixed-effects methods, countries further ahead in the epidemic provide valuable information for
20 those behind. Using current daily global data, we note convergence in near-term predictions for Italy
21 signifying an appropriate call on the future course of the global epidemic. For the UK and elsewhere,
22 prediction of peak and eventual time to resolution is now possible.

23

24

25 **Introduction**

26 Prediction of the ongoing global epidemic of SARS-COV-2 infections¹ presents a number of
27 challenges for epidemiology, notably the apparent susceptibility of the population at large, high
28 transmissibility of the pathogen, lack of surveillance to estimate accurate incidence, prevalence and
29 case fatality rate (CFR)^{2,3,4,5}. Classic models of infectious disease transmission describe the processes
30 using the law of mass action for contact transmission⁶. Refinements include mixing patterns, age
31 stratification and spatial distribution. Whilst these models have proven extremely valuable for
32 decision-making⁷, with validation using extended time-course datasets for multiple pathogens, they
33 necessarily make strong parametric assumptions, from which prediction may be exponentially
34 sensitive during the early growth phase of an epidemic. In this paper we present three
35 complementary methods to inform on immediate, near-term (7- and 14-day) and further projection
36 of the global SARS-COV-2 epidemic. Each method uses daily European Centre for Disease Prevention
37 and Control (ECDC) numbers of new cases and deaths (ECDC)⁸, to inform at the individual country-
38 level, the course of the ongoing SARS-COV-2 epidemic.

39 **Early projection using epidemic growth curves and doubling time ratios**

40 During the very earliest stage of a new epidemic, where there is effectively unrestricted transmission
41 into a new susceptible population, with little known about the underlying disease, an exponential
42 model of the numbers of cases and deaths is appropriate. The number of cumulative cases, $C(t)$ at
43 time, t , is well-described using the simple exponential form, $C(t) = C_0 \exp(r(t - t_0))$, with initial
44 number of cases C_0 at time t_0 , growth rate, r , incidence $rC(t)$ new cases per day (also exponential)
45 and doubling time $\log(2)/r$ days. The SARS-COV-2 infection has an incubation of approximately 5
46 days and a typical time from case presentation to death of approximately 14-21 days^{4,5}. For any case
47 fatality rate (CFR) and duration, T , between presenting as a case and eventual death, the number of
48 deaths at a function of time is therefore $D(t) = \text{CFR } C(t - T)$, which in the exponential phase is $D(t) =$

49 CFR $C_0 \exp(r(t - t_0 - T)) = D_0 \exp(rt)$. Hence, we predict that for any (unknown) time delay and CFR,
50 on a log-scale cases and deaths should initially have parallel slopes, with the distance between
51 curves at any timepoint being defined as $\log(CFR)$. Deviations from parallel log-linearity (e.g., deaths
52 out-pacing cases), may be informative as to when healthcare systems are beginning to fail. The
53 reported CFR will vary from country to country, hence any global assessment must adjust for
54 differences in surveillance practice.

55 The simplest prediction method is a direct comparison of any one country with the population of all
56 other countries rebased to the same number of cases and deaths. Figure 1a shows this prediction for
57 the United Kingdom from all data to 21/MAR/2020. Although prediction interval is wide, when every
58 country's cases and deaths were rebased to 64 and 4, respectively, early prediction of the UK
59 epidemic was possible based on the same percentile growth charts familiar to any paediatrician⁹.
60 The effects of global social distancing might be expected to move the trajectory of a country across
61 percentiles, and this is suggested in more recent analyses (not shown).

62 A more sensitive analysis is provided using the Doubling Time Ratio (DTR), defined as the ratio of
63 doubling times of cases compared to deaths. Figure 1b shows the correlation between exponential
64 growth rates for cumulative cases and deaths estimated from log-linear mixed-effects models of
65 global epidemic data (Date of analysis 21/MAR/2020), with their ratio shown (i.e., the DTR) plotted
66 by descending DTR. The global mean exponential rate of growth of cases for the 33 countries and
67 dependencies having reported 64 or more cases and 4 deaths, for the following 14-day period, was
68 0.20/day (95% CI 0.16, 0.23), doubling time 3.51 days (94% CI 3.01, 4.33), between-country
69 variability 36%. The corresponding growth rate for deaths was 0.22/day (95% CI 0.17, 0.27), doubling
70 time 3.15 days (94% CI 2.57, 4.08), between-country variability 52%. Increased variability in deaths
71 reflects the early stage of the epidemic. Although the estimated global DTR was 1.02 (95% CI 0.97,
72 1.08), it is noteworthy that DTR for Italy and Spain, two EU countries particularly affected by COV-
73 SARS-2, and the United Kingdom (DTR: 1.12, 95% CI 1.02, 1.23) were all significantly greater from

74 unity. A DTR of greater than unity implies that deaths are outstripping cases (for any CFR),
75 suggesting intervention measures are necessary to prevent further unrestricted differential
76 exponential epidemic growth. An up-to-date analysis (Date 08/APR/2020) shows movement in UK
77 DTR (1.07, 95% CI 1.04, 1.10) relative to the Global average (1.05, 95% CI 1.03, 1.06), suggestive of
78 the effects of social distancing.

79 **Near-casting using time-series projection**

80 Forward projection time-series methods have been used for many years in multiple disciplines¹⁰.
81 Conventional time-series projections make assumptions regarding correlations in previous data to
82 inform on future observations. Propagation of uncertainty, however, often limits precision
83 (predictions are based on predictions not observations). The time-delay between cases and deaths
84 inherent in the SARS-COV-2 epidemic lends itself to an alternative novel time-series methodology
85 (Methods). Using past log-case data lagged to 7- and 14-days, and linear mixed-effects models we
86 predict future cases based on past observations, and further predict deaths using the same model.
87 All data is analysed in the log domain, with CFR described by a random effect to account for country-
88 to-country variability in reporting practices. The use of dual lag times incorporates a (weekly) rate of
89 change into the model. The minimum time delay of 7-days means that 7-day predictions are possible
90 based on past *observations* (14-day predictions require one observation and one prediction,
91 predictions beyond that use future predictions). This methodology opens a near-casting window on
92 future events not available to conventional models, with the mixed-effects approach¹¹ allowing
93 prediction in countries that are still early in their epidemic.

94 Figure 2 shows sample predictions for selected countries (cut-off date 08/APR/2020, rebased to 64
95 cases and 16 deaths, 50 countries included in the analysis dataset) of cumulative cases, deaths and
96 incidence. The model has a residual error of 11% (Extended Data) and is useful for near-casting of
97 the SARS-COV-2 epidemic in all countries included, with sufficient precision and sensitivity (to 7- 14-
98 day time delays in actions) for helpful decision-making. The strength of this approach is the

99 population-method used. Information from the global epidemic informs on individual country
100 behaviour, and outlier behaviour can be identified early from individual country posterior model
101 parameters (not shown).

102 **Long-term prediction using parametric epidemiological models**

103 For long-term prediction, parametric epidemiological models will be necessary to control for
104 exponential time-series prediction error. However, caution is advised on model selection and
105 parameter estimation, most notably in the early epidemic phase. Conventional SEIR-type models
106 assume a law of mass action to estimate case reproductive number, R_0 , which measures the
107 transmissibility of the disease. Although intuitive, the notion of mass action, particularly in the
108 presence of significant population-level social distancing (currently in place globally), may be
109 limiting. Instead, we use the Gompertz model of epidemic growth, a limiting form of the logistic
110 model, in which log-cumulative cases is described by the differential equation $d\log C(t)/dt = r -$
111 $s\log C(t)$; r again being the unrestricted exponential growth rate, and r/s the eventual log-carrying
112 capacity. The cumulative number of log-deaths is $\log D(t) = \log CFR + \log C(t - T)$, where the $\log CFR$ is a
113 country-level case fatality rate, again modelled as a random effect to account for case reporting
114 practices. The principal deviations of this model from conventional mass-action are the brake on
115 epidemic growth (proportional to $\log C(t)$) and log-carrying capacity. Transformation to the log-
116 domain makes solution and model-fitting straightforward, controls for variability in highly non-linear
117 processes and permits log-normally distributed random effects. The use of random effects describes
118 country-to-country variability, particularly in surveillance and reporting methods, provided practices
119 remain relatively consistent within a country during the epidemic.

120 Solution of the model for cumulative $\log(\text{Cases})$ and $\log(\text{Deaths})$ for country, i , with addition of 2018
121 population as a possible covariate in the model gives:

$$122 \quad \log C_i(t) = (r_i/s)(1-\exp(-st)) + a_{0i} \exp(-st) + a_{1i} \log N_i$$

123
$$\log D_i(t) = \log \text{CFR}_i + (r_i/s)(1 - \exp(-s(t - T))) + a_{0i} \exp(-s(t - T)) + a_{1i} \log N_i,$$

124 where (r_i/s) is the log-final epidemic size (r_i being a random effect), a_{0i} the intercept (epidemic
125 initiation, also a random effect), $\log N_i$ the reported 2018 World Bank population, and $\log \text{CFR}_i$ the
126 individual country log-case fatality rate. The three random effects presume log-normal global
127 distributions for epidemic initiation, eventual epidemic size and CFR.

128 Figure 3 shows results of the individual country predictions selected countries from analysis of global
129 ECDC data (08/APR/2020 rebased to 128 cases and 32 deaths, 50 countries included in the analysis
130 dataset). Overlaid is the time-series near-casting prediction for comparison and quality control.
131 Where the epidemic shows significant curvature, estimation of peak incidence of death and overall
132 epidemic size is possible. At the global level, population parameters are all well-estimated (Extended
133 Data), without significant correlation for projection. The estimated global unrestricted epidemic
134 growth rate is 0.69/day (95% CI 0.65, 0.73) – in a Gompertzian model, observed exponential growth
135 is always less than r , between-country variability 73%, with CFR 6.3% (95% CI 4.9, 7.7), between-
136 country variability 69%. Model residual for cases and deaths (modelled simultaneously with equal
137 log-residual) is only 12%. Caution is necessary, however, regarding model interpretation –
138 particularly eventual epidemic size, because model *selection* (as opposed to parameter estimation)
139 may be invalid if unqualified by the near-casting model.

140 Discussion

141 In the ongoing SARS-COV-2 epidemic, multiple methods are available for projection of numbers of
142 cases and deaths. Without a means of projection, evaluation of the effectiveness (or otherwise) of
143 intervention methods will prove challenging. Traditional epidemiology models based on contact and
144 transmission patterns are important for planning intervention policies (e.g., mass vaccination and
145 treatment), but they may struggle with calibration in a new setting when key parameters for their
146 evaluation are poorly known. SARS-COV-2 presents such an example; an emergent global epidemic
147 where previous immunity is absent. Robust intervention decisions will, however, be insensitive to

148 parameter assumptions, and the predictions of Ferguson et al⁷, that there are no good outcomes
149 without population intervention are therefore robust. In this analysis, we show that even in the
150 absence of (key) epidemiological parameters (accurately known incubation time, underlying case
151 fatality rate and transmissibility), by taking a global approach to the epidemic, it is possible to
152 provide an early accurate assessment. We have shown that a simple “growth chart”, calculated from
153 rebased case and death data from all countries other than the one of interest, was able to make an
154 early prediction of the course of the UK SARS-COV-2 epidemic. This analysis, coupled with estimation
155 of global log-linear growth rates for cases and deaths showed that the UK was likely to be an outlier,
156 and that intervention was likely necessary. The Doubling Time Ratio for the UK showed that deaths
157 were doubling 12% faster than cases, and faster than the global average, during the initial
158 exponential phase ($P < 0.05$ compared with a global estimate of equality), possibly implicating strain
159 on healthcare services from an early stage of the epidemic. Similar findings were noted in both Italy
160 and Spain.

161 Such a simple model is useful for immediate prediction but makes a strong assumption regarding
162 future time course and is obviously unsuited to the first outbreak (in this case in Wuhan). It only
163 assumes that progression (on a log scale) will follow other countries when rebased to the same point
164 in the epidemic. Prediction intervals are necessarily very wide but still useful for the immediate (7-
165 to 14-day time-frame) near-casting. Refinements such as scaling for population size (which was not
166 available in the original ECDC dataset) may be unnecessary when an early epidemic is confined to
167 the regional scale with population size of order 10^6 (i.e., Wuhan Province, Northern Italy, London).

168 Once in the exponential growth phase, data-driven prediction is necessary. The characteristics of the
169 epidemic (pseudo-exponential growth and an incubation from case presentation to eventual death)
170 lend themselves to novel time-series analysis. We find that rebasing global data to a common
171 number of cases, using 7- and 14-day lagged log-cumulative cases, and application of linear mixed-
172 effects models (accounting for country-to-country variability), can be used to near-cast epidemic

173 direction. The advantage over traditional time-series approaches is that the first 7-days of
174 predictions are founded on past observations (not future predictions), thereby reducing
175 (exponential) error. Case fatality rate differences were well-described at the country level by a log-
176 linear random effect.

177 For eventual long-term forecasting, a parametric epidemiological model is necessary. Without such a
178 model, a call on the location and size of the peak, and eventual size and duration of the epidemic, is
179 not possible. These parametric models suffer from uncertainties and parameter correlations that
180 make initial estimation and future very challenging from single country data. However, we find that
181 using global, rather than individual country, data, a non-linear mixed-effects Gompertz epidemio-
182 statistical model can describe the epidemic well with excellent parameter estimation and low
183 residual. Providing a prediction of the time course for any single country from the model, however,
184 requires some form of quality control. That quality control is provided by the near-casting model.
185 Once near-term prediction intervals for the two modelling approaches coalesce, it is possible to
186 move from near-casting to forecasting with confidence. Based on our most up-to-date analysis, this
187 coalescence has happened in Italy, and it is now possible to provide more confident forward
188 projections for that country and others.

189 The power of the epidemio-statistical global mixed-effects analysis is that, like the two previous
190 approaches, countries further advanced in their epidemic inform on those following behind. We find
191 concordance between the two approaches in Italy, with support for an eventual epidemic size of
192 approximately 31.8k (95% CI 28.8k, 35.2k) deaths, with resolution by the beginning of July 2020.
193 Using the model, we also find support for a peak in the UK epidemic somewhere from 7/APR/2020
194 to 21/APR/2020, approximately 22.6k (95% CI 12.3, 26.5k) deaths, and resolution also by July 2020.
195 It is still too early to call the eventual global epidemic size, although values are predicted for every
196 country in the analysis dataset. All predictions are founded on a continuation of the current global
197 social distancing intervention program, which supports the use of a Gompertz epidemiological

198 model. A relaxation of restrictions would release the brake on contacts, increase transmission,
199 invalidate future model assumptions, and most-likely lead to an extension of the epidemic duration.
200 Such predictions are out of scope for this analysis.

201 **Methods**

202 Daily global data on new cases and deaths was downloaded from EDCCD. Cumulative cases and
203 deaths were log-transformed for analysis. Data was selected based on number of rebased cases, and
204 time rebased from this point. Growth rate charts were calculated from geometric mean and
205 percentiles assuming log-normal distribution, with doubling time estimated by log-linear regression.
206 Doubling Time Ratio was estimated by separate log-linear random effects models, growth rates
207 merged, and 95% confidence intervals calculated. For time-series projection, log-cumulative cases
208 were lagged by 7- and 14-days prior to fitting a log-linear mixed-effects model with random
209 intercept and slopes. Log-cumulative deaths was fitted to predicted number of cases, also with
210 random intercept and slope. The epidemio-statistical Gompertz model was fitted to the same
211 analysis dataset. Predictions (mean and 95% CI) were back-transformed, differenced (for incidence
212 calculation) and overlaid with observation. Model goodness of fit was assessed visually and by Bayes
213 Information Criteria. All analyses were conducted in SAS 9.4 proc mixed and proc nlmixed,
214 automated by scripting and available from the author. At the time of publication, daily updated
215 global projections are ongoing.

216 **References**

- 217 1. World Health Organisation. (2019). Coronavirus disease (COVID-2019) situation reports.
218 Retrieved from [https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports)
219 [reports](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports)
- 220 2. Li, Q., & Wu, P. (2020). Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-
221 Infected Pneumonia. *N Engl J Med.*, 382(13), 382(13). doi:doi:10.1056/NEJMoa2001316

- 222 3. Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., & Shaman, J. (2020). Substantial
223 undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2).
224 **Science**, eabb3221. doi:DOI: 10.1126/science.abb3221
- 225 4. Wang, D., Hu, B., & Hu, C. (2020). Clinical Characteristics of 138 Hospitalized Patients With 2019
226 Novel Coronavirus-Infected Pneumonia in Wuhan, China. **JAMA**, e201585.
227 doi:doi:10.1001/jama.2020.1585
- 228 5. Zhou, F., Yu, T., & Du, R. (2020). Clinical course and risk factors for mortality of adult inpatients
229 with COVID-19 in Wuhan, China: a retrospective cohort study. **Lancet**, 395(10229), 1054–1062.
230 doi:doi:10.1016/S0140-6736(20)30566-3
- 231 6. Anderson, R. M., & May, R. M. (1992). *Infectious Diseases of Humans: Dynamics and Control*.
232 Oxford: Oxford Science Publications.
- 233 7. Ferguson, N. M., Laydon, D., Nedjati-Gilani, G., & Imai, N. (2020). Impact of non-pharmaceutical
234 interventions (NPIs) to reduce COVID19 mortality and healthcare demand. doi:DOI:
235 <https://doi.org/10.25561/77482>
- 236 8. European Centre for Disease Prevention and Control. (2020). Publications and Data. Retrieved
237 from [https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-](https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide)
238 [distribution-covid-19-cases-worldwide](https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide)
- 239 9. Centers for Disease Prevention and Control. (2017). Clinical Growth Charts. Retrieved from
240 https://www.cdc.gov/growthcharts/clinical_charts.htm
- 241 10. Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to Time Series Analysis*
242 *and Forecasting*, 2nd Edition. Hoboken: Wiley-Interscience.
- 243 11. Demidenko, E. (1987). *Mixed Models: Theory and Applications with R*, Second Edition. New
244 York: Wiley Interscience. doi:10.1002/9781118651537

245 **Acknowledgements**

246 The author would like to acknowledge the European Centre for Disease Prevention and Control,
247 without who's daily updated data, this work would not have been possible. This research is
248 dedicated to all who have lost their lives to this dreadful disease.

249 **Competing Interests**

250 The author is an employee and stockholder of GlaxoSmithKline. This work was completed as part of
251 his ongoing employment duties.

252 **Materials and Correspondence**

253 Correspondence and requests for materials should be addressed to daren.j.austin@gsk.com.

254 Reprints and permissions information is available at www.nature.com/reprints.

255 **Extended data**

256 **Time-Series model**

257 Final parameters for linear mixed-effects Time-Series model of log(Cases); var1 = 7-day lagged

258 log(Cases), var2 = 14-day lagged log(Cases). Geold = Country ID.

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > t
Intercept	3.0981	0.4699	27	6.59	<.0001
var1	0.7920	0.08559	26	9.25	<.0001
var2	-0.07047	0.04846	21	-1.45	0.1607

259

Covariance Parameter Estimates		
Cov Parm	Subject	Estimate
UN(1,1)	geold	4.1598
UN(2,1)	geold	-0.5955
UN(2,2)	geold	0.1002
UN(3,1)	geold	0.1834
UN(3,2)	geold	-0.03779
UN(3,3)	geold	0.01943
Residual		0.002510

260

261 Final parameters for linear mixed-effects Time-Series model of log(Deaths); var1 = 7-day lagged

262 log(Cases), var2 = 14-day lagged log(Cases).

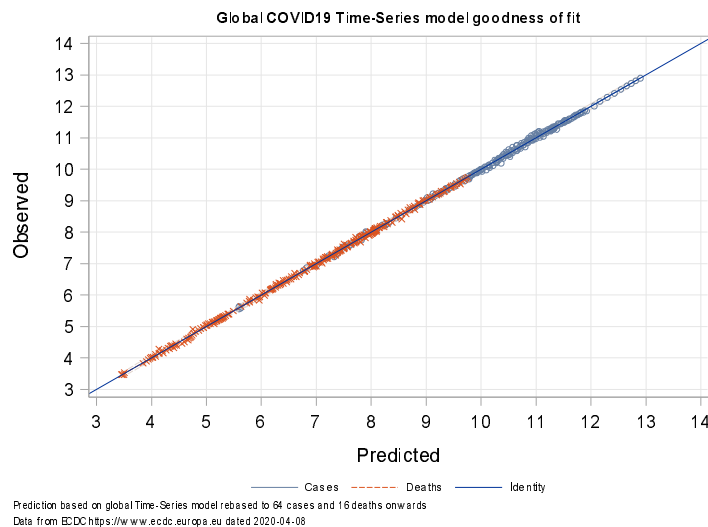
263

Covariance Parameter Estimates		
Cov Parm	Subject	Estimate
UN(1,1)	geold	27.4397
UN(2,1)	geold	-3.2088
UN(2,2)	geold	0.4123
UN(3,1)	geold	0.4031
UN(3,2)	geold	-0.07932
UN(3,3)	geold	0.04143
Residual		0.002019

264

Solution for Fixed Effects								
Effect	Estimate	Standard Error	DF	t Value	Pr > t	Alpha	Lower	Upper
Intercept	-1.5271	1.0530	27	-1.45	0.1585	0.05	-3.6876	0.6334
var1	0.7129	0.1448	26	4.92	<.0001	0.05	0.4152	1.0107
var2	0.1953	0.06690	21	2.92	0.0082	0.05	0.05621	0.3345

265



266

267 **Gompertzian epidemio-statistical model**

268 Estimated model parameters and correlation matrix (population was excluded by BIC criteria)

269

Parameter Estimates								
Parameter	Estimate	Standard Error	DF	t Value	Pr > t	95% Confidence Limits		Gradient
r	0.6919	0.02010	34	34.43	<.0001	0.6510	0.7327	-0.00005
s	0.06620	0.001062	34	62.34	<.0001	0.06404	0.06836	-0.00290
a0	7.9252	0.1091	34	72.67	<.0001	7.7036	8.1469	0.000011
eps	0.01455	0.000672	34	21.65	<.0001	0.01318	0.01591	0.002434
s2b1	0.4351	0.1023	34	4.25	0.0002	0.2273	0.6430	0.000321
s2b2	0.01102	0.002862	34	3.85	0.0005	0.005200	0.01683	0.003451
s2b1b2	0.04661	0.01427	34	3.27	0.0025	0.01761	0.07561	-0.00046
s2b3	0.3981	0.09627	34	4.14	0.0002	0.2024	0.5938	0.000270
ALAG	4.7195	0.1542	34	30.60	<.0001	4.4061	5.0329	-0.00006
CFR	-2.7634	0.1083	34	-25.52	<.0001	-2.9835	-2.5434	-0.00019

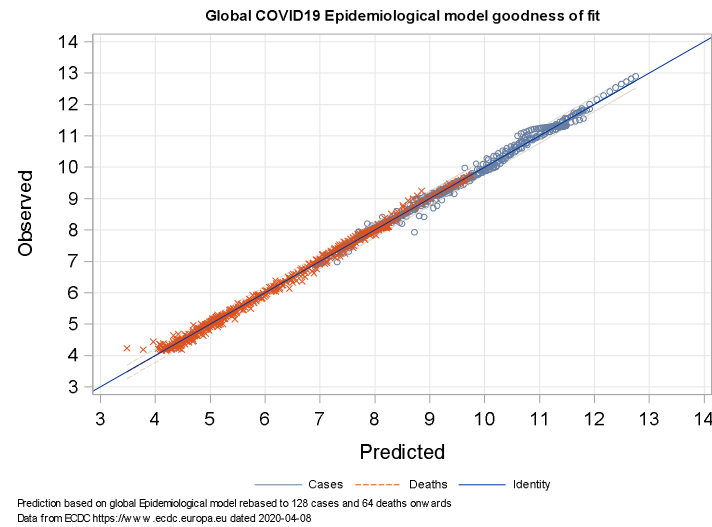
270

Correlation Matrix of Parameter Estimates										
	r	s	a0	eps	s2b1	s2b2	s2b1b2	s2b3	ALAG	LNCFR
r	1.0000	0.4697	0.5587	-0.0006	0.0019	0.0473	0.0234	-0.0148	-0.2684	-0.0092
s	0.4697	1.0000	-0.0235	-0.0070	-0.0092	0.0970	0.0320	-0.0118	-0.5106	-0.0890
a0	0.5587	-0.0235	1.0000	-0.0012	0.0005	0.0061	0.0045	-0.0009	0.0488	-0.0027
eps	-0.0006	-0.0070	-0.0012	1.0000	-0.0045	-0.0047	-0.0043	-0.0048	0.0021	0.0037

Correlation Matrix of Parameter Estimates										
	r	s	a0	eps	s2b1	s2b2	s2b1b2	s2b3	ALAG	LNCFR
s2b1	0.0019	-0.0092	0.0005	-0.0045	1.0000	0.3962	0.7580	0.0167	0.0134	0.0081
s2b2	0.0473	0.0970	0.0061	-0.0047	0.3962	1.0000	0.7896	-0.0685	-0.0635	-0.0122
s2b1b2	0.0234	0.0320	0.0045	-0.0043	0.7580	0.7896	1.0000	-0.0467	0.0027	0.0083
s2b3	-0.0148	-0.0118	-0.0009	-0.0048	0.0167	-0.0685	-0.0467	1.0000	0.0157	-0.0075
ALAG	-0.2684	-0.5106	0.0488	0.0021	0.0134	-0.0635	0.0027	0.0157	1.0000	0.1780
LNCFR	-0.0092	-0.0890	-0.0027	0.0037	0.0081	-0.0122	0.0083	-0.0075	0.1780	1.0000

271

272 Parameters are: r = growth rate, r/s = log-carrying capacity, a_0 = initial condition, eps = residual, s_{2b1} =
 273 variance(r), s_{2b2} = variance(a_0), s_{2b3} = variance(logCFR), ALAG = lag time between cumulative cases
 274 and deaths, LNCFR = log(case fatality rate).

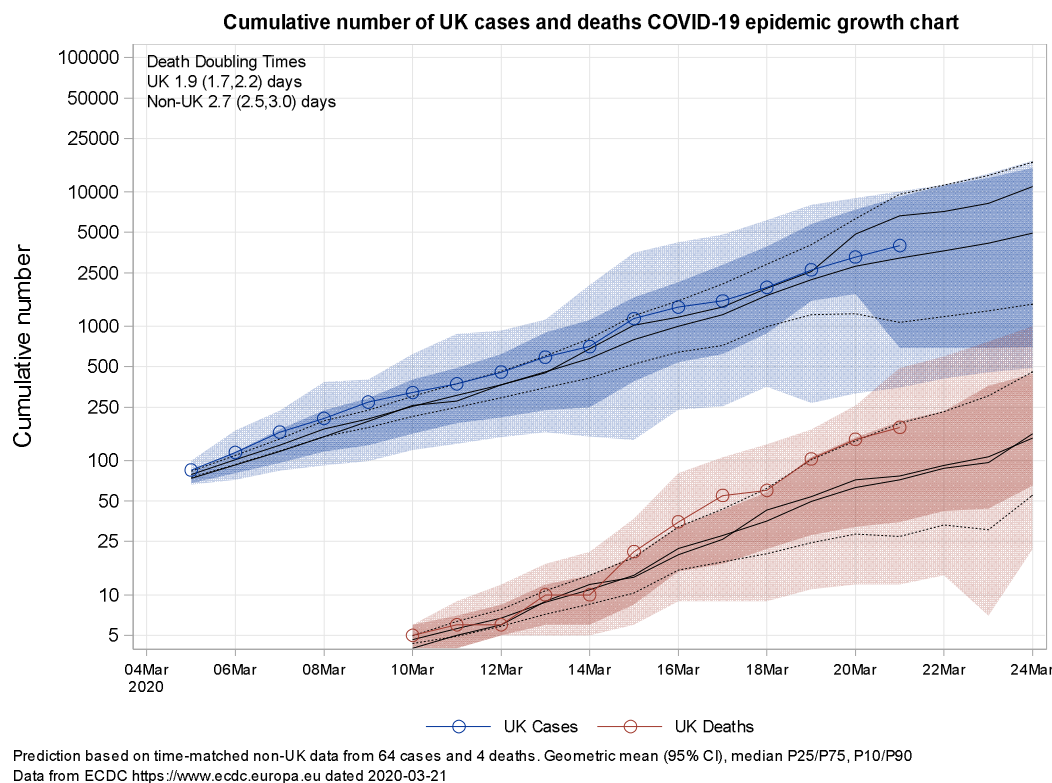


275

276 **Figure 1**

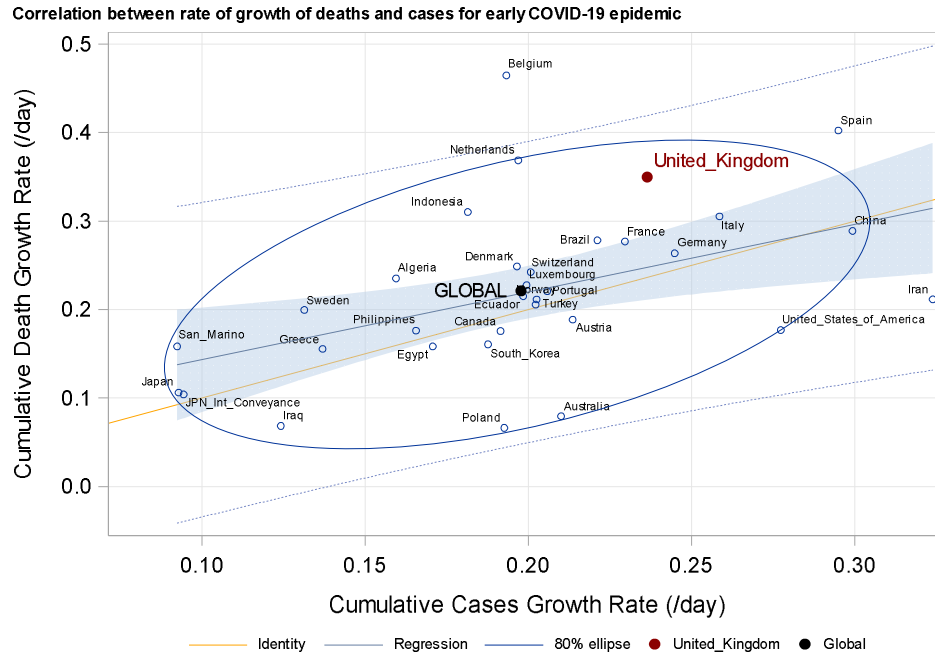
277 a) Growth curve for 2020 SARS-COV-2 global epidemic with UK data (not included in curve
278 estimation) overlaid rebased to 64 cases and 4 deaths. Geometric mean and 95% CI, median,
279 interquartile and interdecile ranges indicated. Data to 21/MAR/2020. b) Global exponential growth
280 rates for cumulative cases and deaths with UK indicated (21/MAR/2020, data rebased to 64 cases
281 and 4 deaths and fitted with a log-linear mixed effects model). c) Growth Rate Ratios for all
282 countries, values greater than unity indicate that the rate of growth of deaths is greater than cases
283 and may signify strains on healthcare (Data to 21/MAR/2020).

284 a)



285

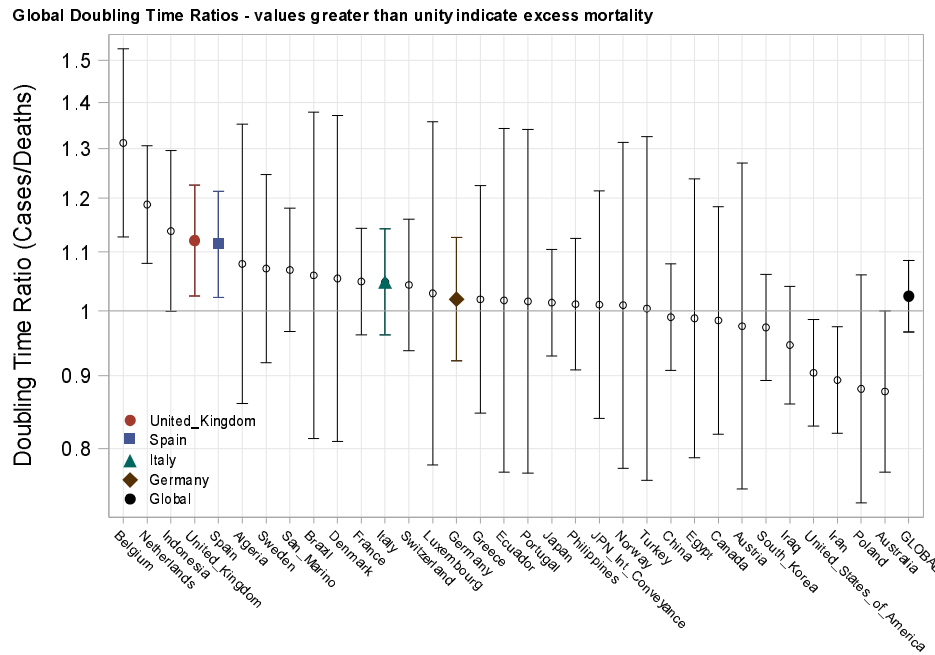
286 b)



Log-linear mixed-effects analyses of data rebased to 64 cases and 4 deaths. Analysis is for 14 days from rebase
Data from ECDC <https://www.ecdc.europa.eu> dated 2020-03-21

287

288 c)



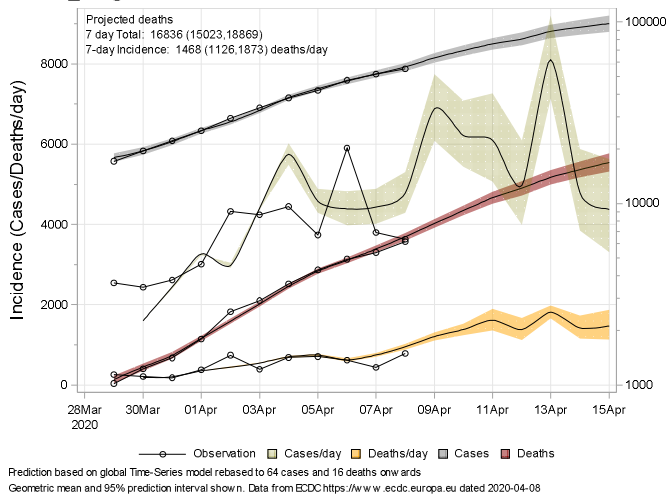
Log-linear mixed-effects analyses of data rebased to 64 cases and 4 deaths. Analysis is for 14 days from rebase
Data from ECDC <https://www.ecdc.europa.eu> dated 2020-03-21

289

290 **Figure 2**

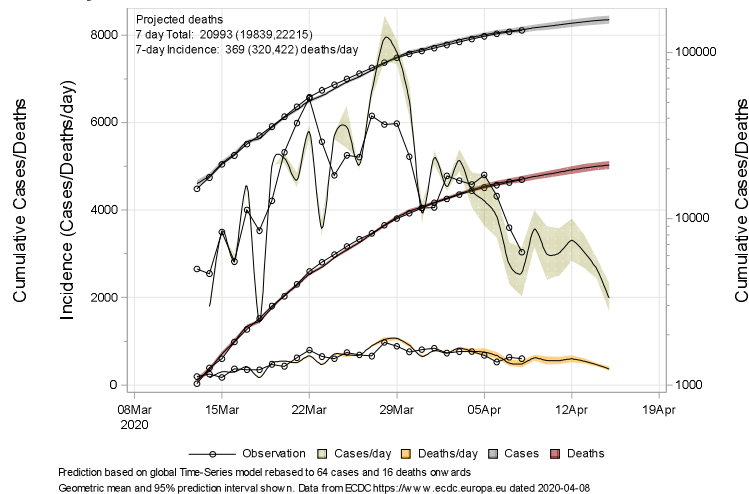
291 Near-term prediction of the ongoing UK SARS-COV-2 epidemic (mean and 95% prediction interval)
292 for the dual-lagged linear mixed effects model. Model shows cumulative cases and deaths and daily
293 case and mortality incidence. Near-casting (7-day) projection indicated. Data to 08/APR/2020.

United_Kingdom

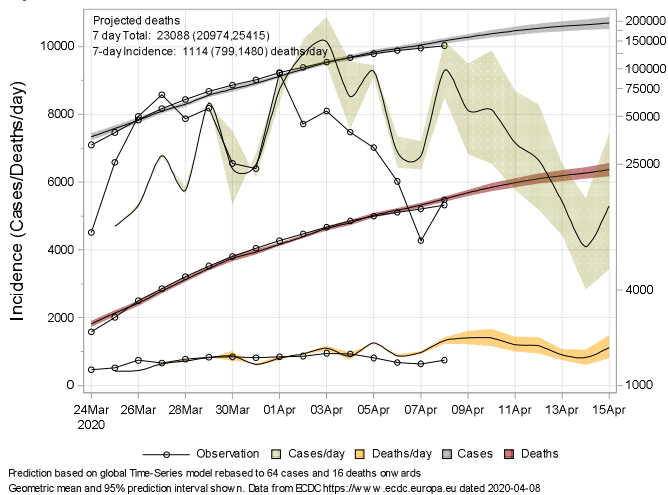


294

Italy

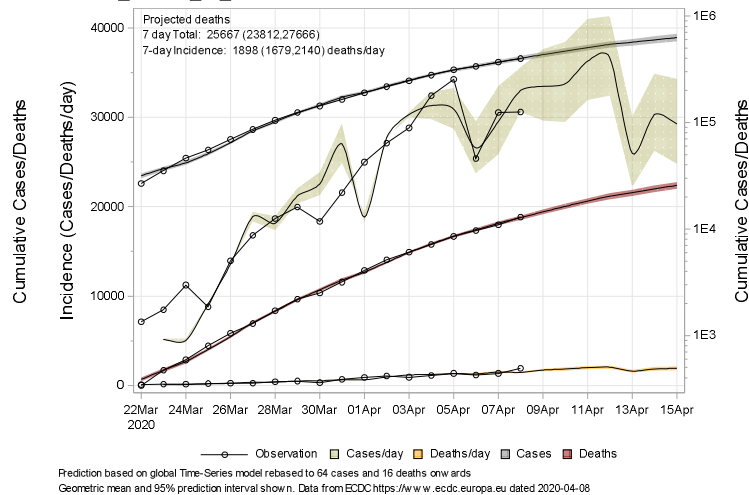


Spain



295

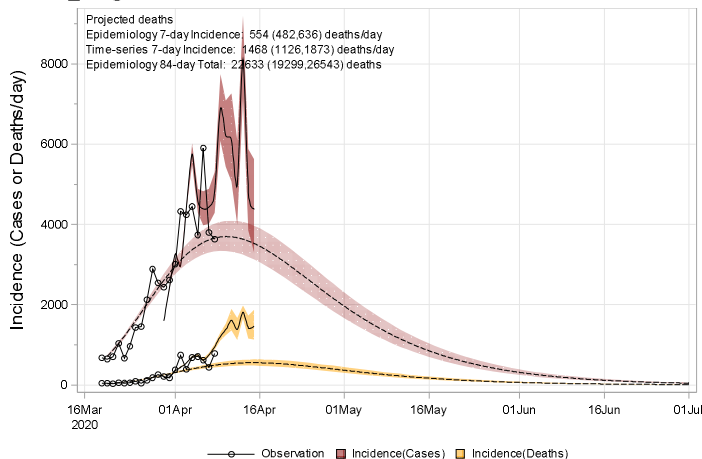
United_States_of_America



296 **Figure 3**

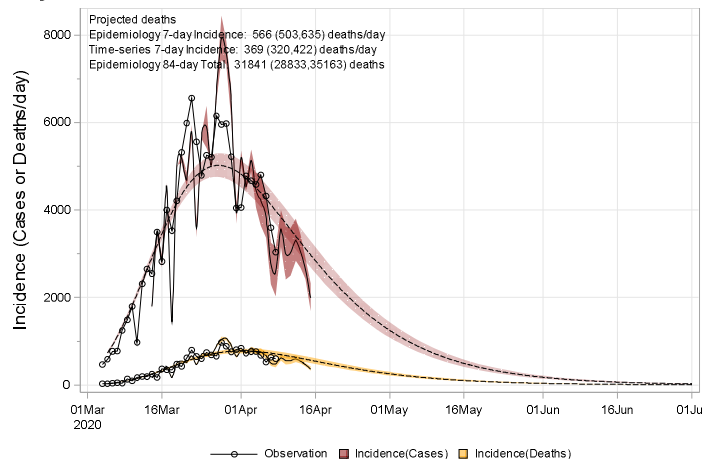
297 Posterior predicted course of the SARS-COV-2 epidemic for selected countries with near-term dual
298 lagged model predictions overlaid. Concordance in short-term predictions improves confidence in
299 the long-term projection Lack of concordance between prediction methods means that it is too early
300 to call the final size and shape of the UK epidemic.

United_Kingdom



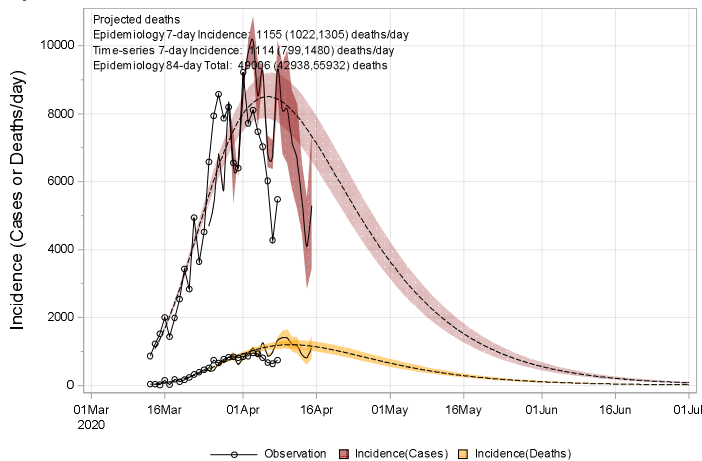
Prediction based on global Epidemiological model rebased to 128 cases and 64 deaths onw ards
Geometric mean and 95% prediction interval show n. Data from ECDC <https://www.ecdc.europa.eu> dated 2020-04-08

Italy



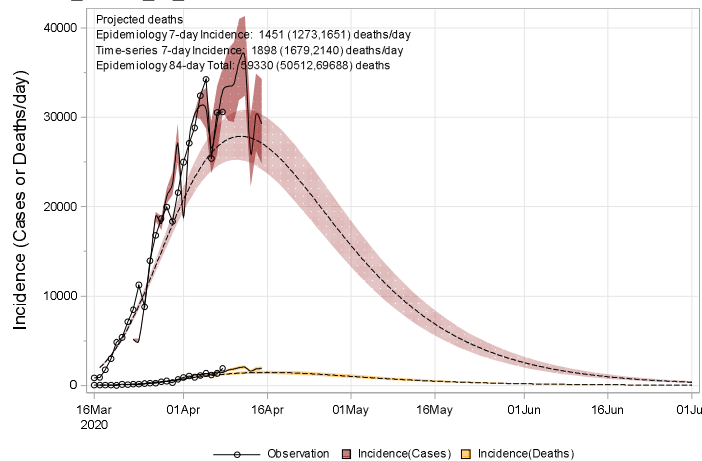
Prediction based on global Epidemiological model rebased to 128 cases and 64 deaths onw ards
Geometric mean and 95% prediction interval show n. Data from ECDC <https://www.ecdc.europa.eu> dated 2020-04-08

Spain



Prediction based on global Epidemiological model rebased to 128 cases and 64 deaths onw ards
Geometric mean and 95% prediction interval show n. Data from ECDC <https://www.ecdc.europa.eu> dated 2020-04-08

United_States_of_America



Prediction based on global Epidemiological model rebased to 128 cases and 64 deaths onw ards
Geometric mean and 95% prediction interval show n. Data from ECDC <https://www.ecdc.europa.eu> dated 2020-04-08

301

302