

The Impact of Social Distancing on The Course of The Covid-19 Pandemic in Four European Countries

Dr Andrew Shardlow MB BS, BSc(Hons), MRCP

May 3, 2020

1 Introduction

In this article the mortality data from four European countries arising from the Covid-19 pandemic is modelled using logistic functions. The countries chosen for examination are Spain, Italy, France and the UK. They have been selected because in each the pandemic is advanced, mortality high and any prospect of containment has passed. They have also been selected because in each social distancing has been used in an attempt to reduce peak daily mortality with relatively strict enforcement following a defined date. The choices of data set and model type is justified. The impact, if any, of social distancing is examined.

2 The Selection of a Covid-19 Pandemic Model

It is probable Covid-19 infection has a high proportion of asymptomatic or minimally symptomatic[1][2] persons and is infectious before symptom development.[3] This has allowed the disease to spread widely amongst the population in countries where early containment was not successful or not attempted. In these circumstances it is reasonable to use a simple pandemic model to provide an overview of pandemic progress. This is in contradistinction to the complex numerical modelling with geographic inputs and outputs that provide detailed local morbidity data to inform logistics and planning. Mechanistic epidemiological models are often founded on unknown or imprecisely known parameters and in this sense are not precisely dynamic. They do however inform the mathematical form of kinetic models.

2.1 The Kermack and McKendrick Model

As this disease is of short duration in relation to population life expectancy we might use the *SIR* compartmental model of Kermack and McKendrick.[4] This is a three compartment model in which N represents the size of the population of interest, S represents the number of susceptible people, I those on their infective journey to R , the recovered or dead. In the following equations β represents the product of the average contact rate c and p_i the probability of transmitting infection in a single contact between an infected and a susceptible person. γ is the rate of transition from I to R . It is defined by three ordinary differential equations:

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

$$\frac{dS}{dt} = -\beta \cdot I \cdot S \quad (1)$$

$$\frac{dI}{dt} = \beta \cdot I \cdot S - \gamma \cdot I \quad (2)$$

$$\frac{dR}{dt} = \gamma \cdot I \quad (3)$$

On reflection however, the assumption present in equation 3 is only valid if illness duration is uniformly distributed within I . This is only true if I is constant over time, which in general is not the case.

2.2 A Modified Kermack and McKendrick Model

An alternative SIR model can be developed in which R is a time delayed convolution of $N - S$ with a normalized mask M with its peak shifted forward in time by λ , the latency in I , and representing the probability over time of leaving I after entry into I .

$$R(t) = \int_{-\infty}^{\infty} (N - S(\tau)) M(t - \tau - \lambda) d\tau \quad (4)$$

$$I = N - S - R \quad (5)$$

$$\frac{dS}{dt} = -\beta \cdot I \cdot S \quad (6)$$

This has been solved numerically using a computerised algebra system (Maxima). If this is done with a broad skewed mask, the kinetic pattern described by the original SIR model is recreated with a skewed I . If convolved with a Kronecker delta, $N - S$ and R can be shown to be an identical time shifted logistic functions such that $R(t) = N - S(t - \lambda)$.

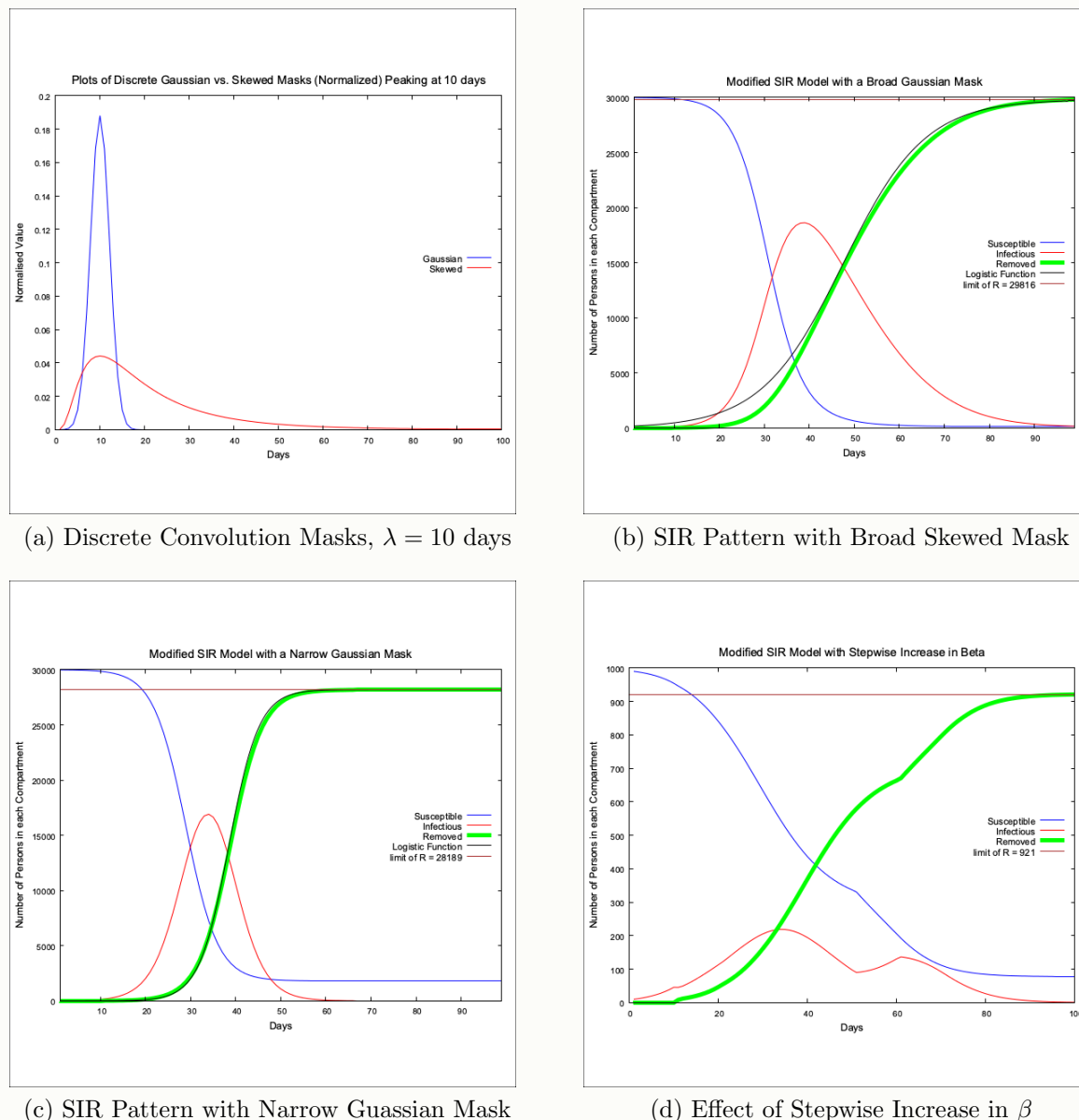


Figure 1: Plots of S,I and R from a numerical solution to a modified version of the tri-compartmental SIR model of Kermack and McKendrick

The original SIR model was used to model diseases that might well have a broad skewed probability distribution of latency in travelling between S and R . Many viral illnesses will have a more predictable λ . We should in the first instance attempt to model R with a simple logistic function.

It should be noted that, although logistic in shape, the model does not exhaust all of the susceptible population over time. In this way it differs dynamically from Verhulst's original concept.[5] The limiting value of cumulative mortality is therefore not fixed, increasing with increasing β and λ though must be less than the initial susceptible population. This will allow for further smaller peaks of infection should β subsequently increase while I is not zero.

3 Modelling The Covid-19 Pandemic using a Logistic Curve

The development of suitable serological and PCR tests and the rolling out of testing has understandably trailed the rapid evolution of the Covid-19 pandemic.[6][7] Morbidity data is bound to be influenced by testing bias. Mortality data is likely to be more reliable because the probability of testing will be highest in those recently deceased following the exhibition of symptoms of the illness. Mortality $D(t)$ to be proportional to $R(t)$, related by an as yet unknown case fatality probability. We should, therefore, study mortality data to inform our pandemic modelling. Of course we need to be aware that mortality data will lag behind morbidity data.

Cumulative mortality is an increasing monotone. The daily geometric increase in mortality D is the growth factor G or alternatively $(1 + g)$ and so $D_{n+1} = GD_n = (1 + g)D_n = D_n + gD_n$. We therefore have a window on the day to day value of g using the equation

$$\frac{\Delta D}{D} = g$$

The logistic function is the solution to an ordinary differential equation $\frac{dD}{dt} = g_0 \cdot (1 - \frac{D}{L}) \cdot D$ where g_0 is the growth factor when $D = 0$, L is the limiting value of D . We can see that $\frac{\Delta D}{D} = g = g_0 (1 - \frac{D}{L})$ and so

$$\frac{\Delta D}{D} = g_0 - \frac{K}{L}D$$

If the mortality data is modelled by a logistic we should expect a plot of $\frac{\Delta D}{D}$ vs. D to be a straight line. The D intercept is the predicted total mortality of the pandemic. The $\frac{\Delta D}{D}$ intercept is the value of g_0 .

We would expect the effect of social distancing to be a stepwise function modulating g so:

$$g = s \cdot g_0 \left(1 - \frac{D}{L}\right)$$

where $s = 1$ before the impact of social distancing and $s < 1$ thereafter. s is a stepwise function of time but as cumulative mortality vs. time is an increasing monotone the step also occurs at a particular cumulative mortality.

For any line segment in the space of $\frac{\Delta D}{D}$ vs. D there is an equivalent logistic function segment in the space of D vs. $Date$.¹

¹There are actually an infinite number of identically shaped time shifted logistic segments in D vs. $Date$ which would transform under this operator to a single line segment, the kernel, in $\frac{\Delta D}{D}$ vs. D . The choice of time shift has been selected algorithmically to minimise RMSE

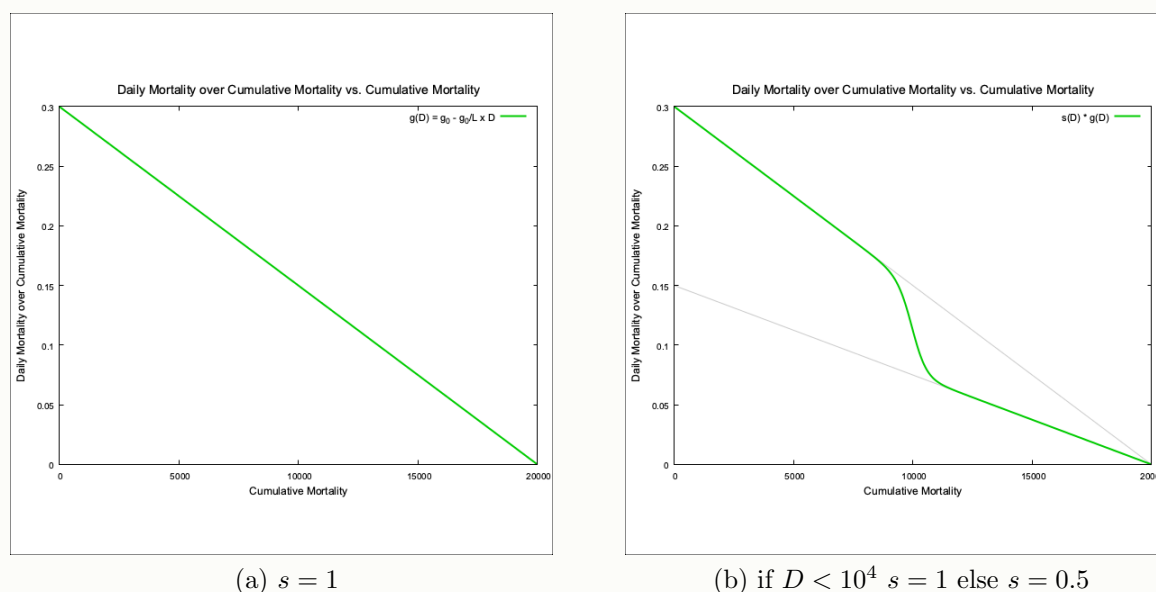


Figure 2: Plots of Predicted $\frac{\Delta D}{D}$ vs. D using a Logistic Function

4 Testing the Model Fit

If we look at the plots of mortality data[8] we are initially confronted by the considerable variation in the day to day reporting from each country. This is presumably related to national testing and recording procedures. For this reason the author has chosen to focus initially on the mortality data from Italy as it appears the least erratic.

4.1 Italy

On looking at the $\frac{\Delta D}{D}$ vs. D plot for Italy as in figure 3 we see that the data points appear to fall into two separate groups which appear highly correlated to their distinct regression lines. The segmentation into these two “phases” of the pandemic was performed algorithmically so as to avoid subjective bias, splitting the data into two at each possible location and finding a break, if present, where the 95% confidence intervals for gradient did not overlap.

As previously stated, for any line segment in the space of $\frac{\Delta D}{D}$ vs. D there is an associated logistic function segment in the space of D vs. $Date$ (which minimises the sum of residual errors).

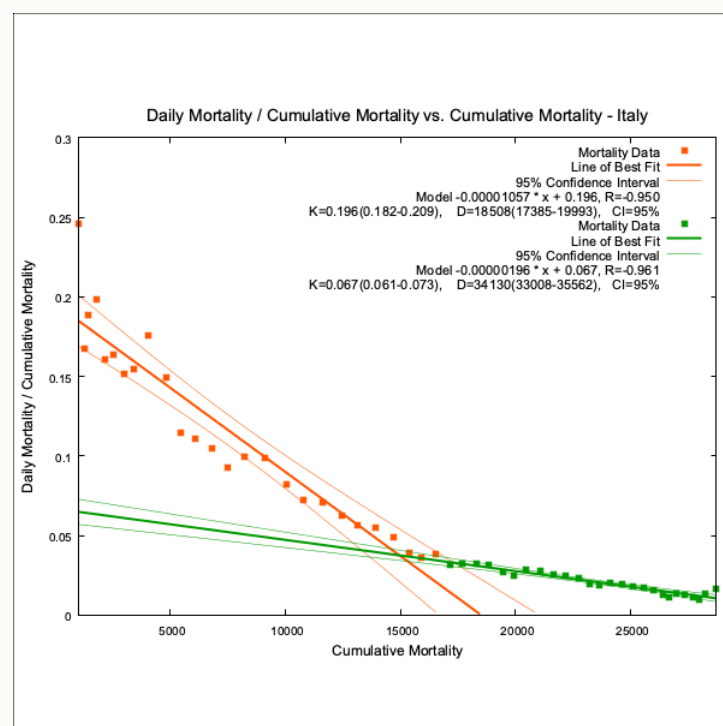


Figure 3: Italy, Linear Regression, Daily Mortality / Cumulative Mortality vs. Cumulative Mortality.

From figures 3, 4 and 5 we can see that the cumulative mortality data for Italy is modelled by two logistic functions spliced together in a piecewise fashion:

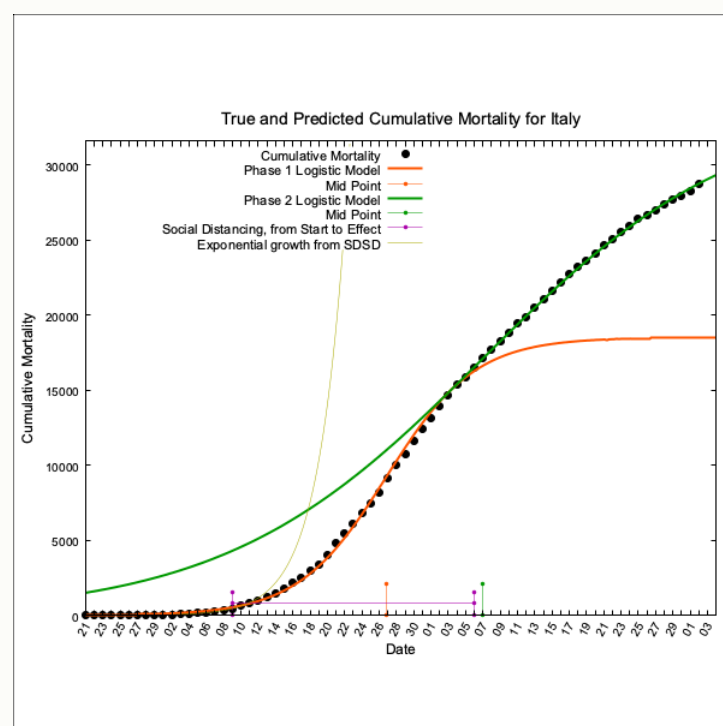


Figure 4: Italy, True and Modelled Cumulative Mortality vs. Date.

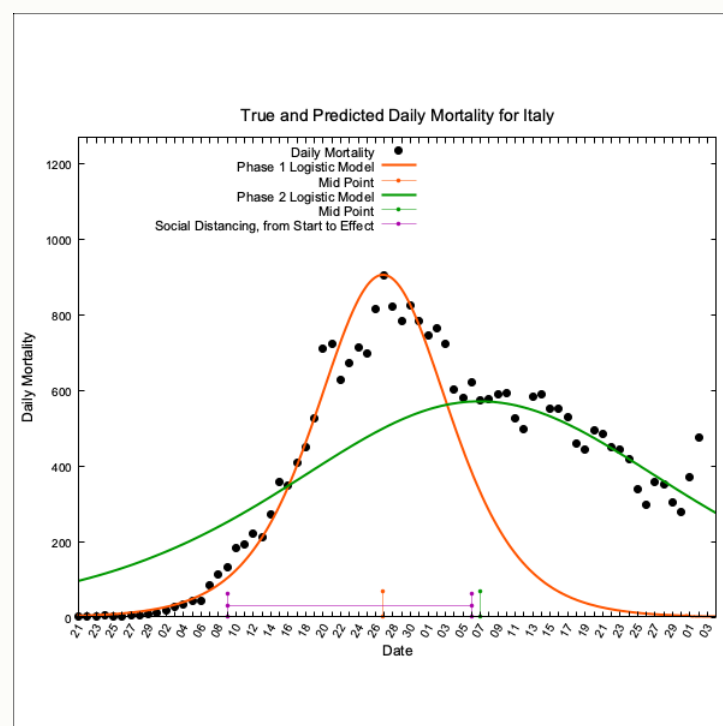


Figure 5: Italy, True and Modelled Daily Mortality vs. Date.

The date of peak mortality for each logistic is marked on the date axis in orange (the “natural peak”) and green (the modified peak).

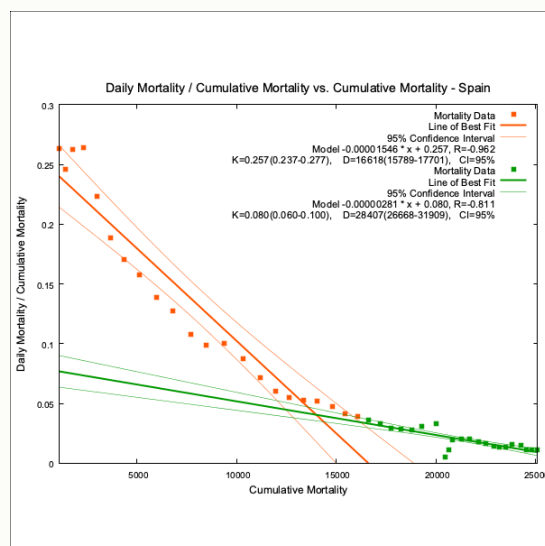
An exponential growth curve is attached to the data point on the day of implementation of “lockdown” with a growth factor based on the gradient of the regression line of the preceding week’s log mortality data. It is labelled Exponential Growth from Social Distancing Start Date (SDSD). It is evident that in the period encompassing social distancing implementation, mortality growth was not exponential but rather a logistic function, which unperturbed by any other factor, would naturally flatten (and daily mortality reach its peak).

The date of the start of social distancing through “lockdown” has been marked on the plot with a purple T bar between it and the date of the natural juncture of the two logistic curves. This marks the point at which social distancing appears to have produced an abrupt stepwise reduction in g . For Italy this date occurs 25 days after the start of social distancing. We must assume that in Italy the average duration of time from contracting Covid-19 to death for those unfortunate enough to succumb is around 3.5 weeks.

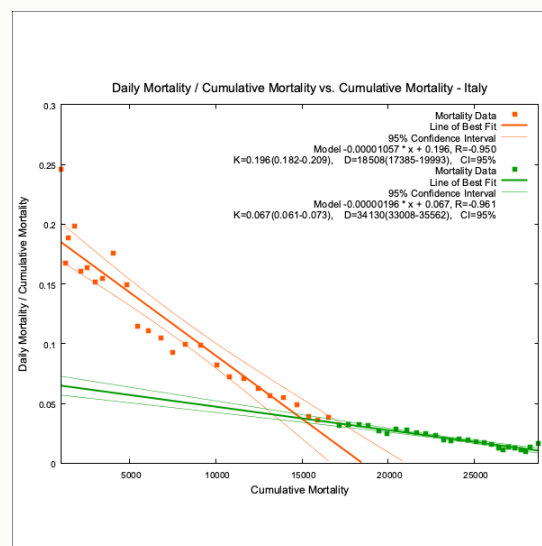
4.2 Plots for Italy, Spain, France and The UK

The plots for Spain, Italy, France and the UK follow for side by side comparison:

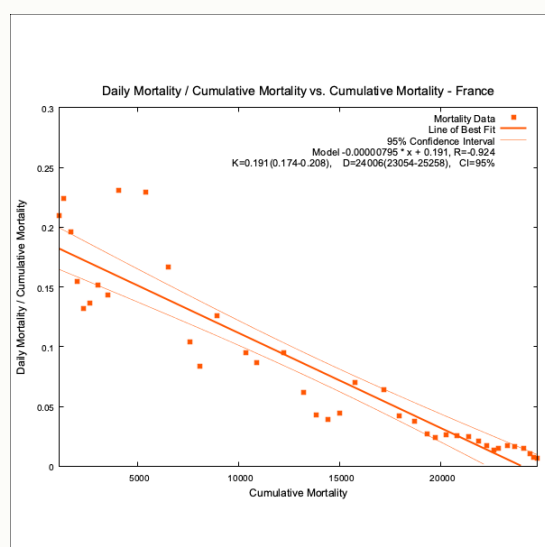
The overall pattern of a “natural” peak followed by the remains of a modified peak is present in the data from both Italy and Spain. In the UK this pattern is just about discernible though the data has much more day to day variability. France appears to have a single phase to its pandemic with daily mortality climbing and descending a sharper “natural” peak.



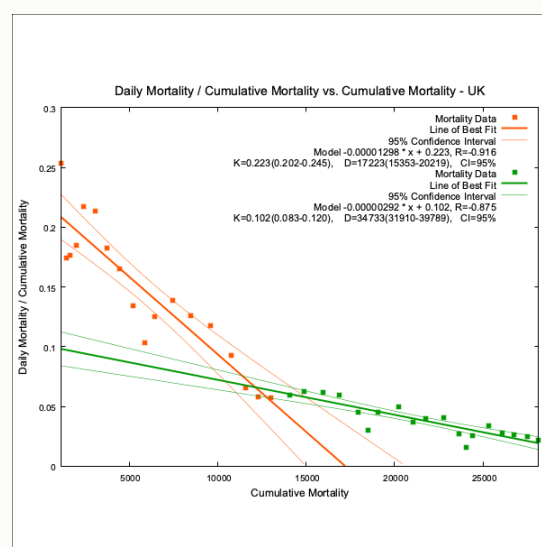
(a) Spain



(b) Italy

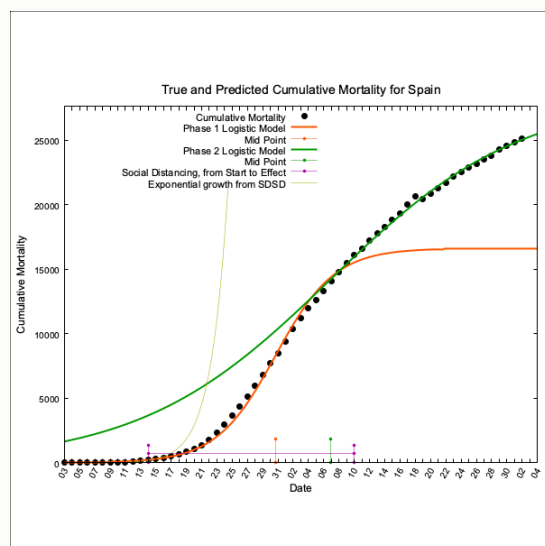


(c) France

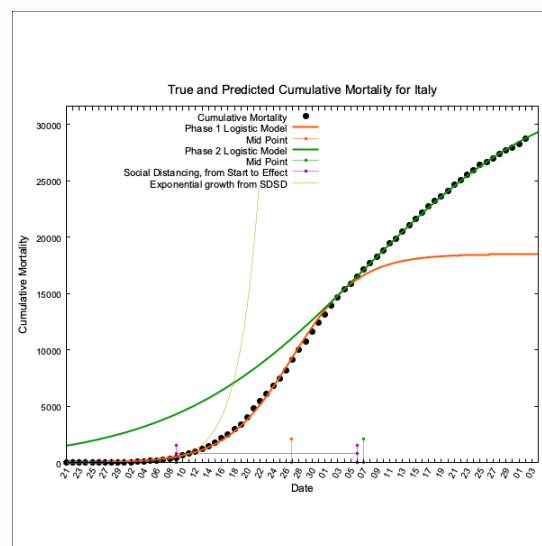


(d) UK

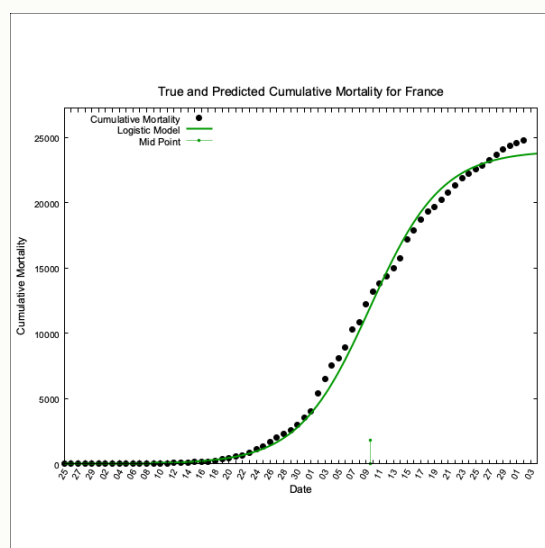
Figure 6: Plots of $\frac{\Delta D}{D}$ vs. D and their Regression Lines



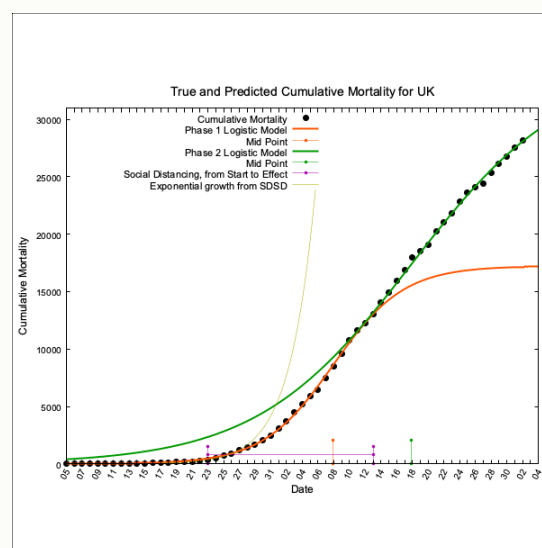
(a) Spain



(b) Italy

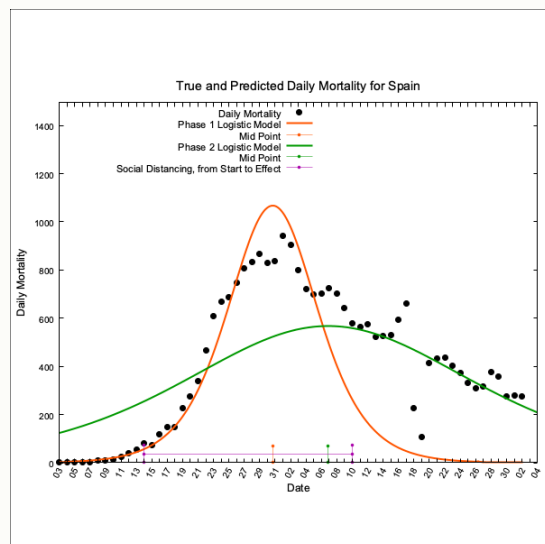


(c) France

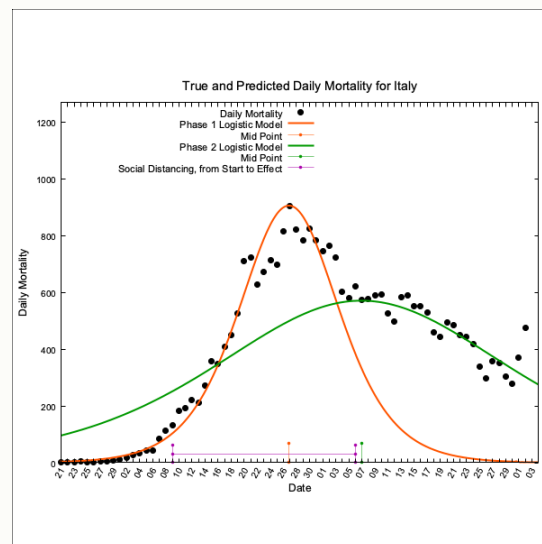


(d) UK

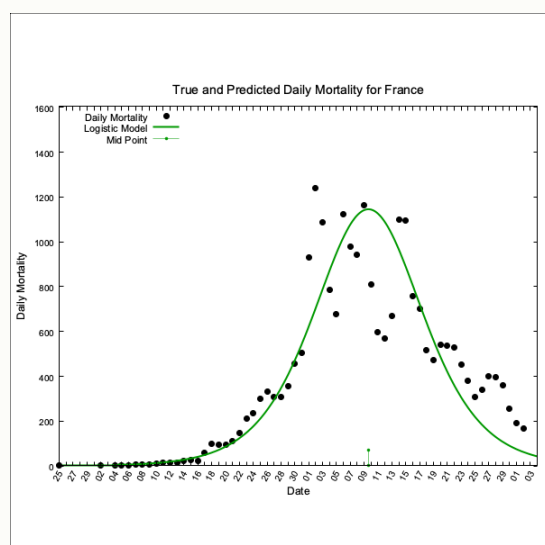
Figure 7: Cumulative Mortality, Actual and Predicted



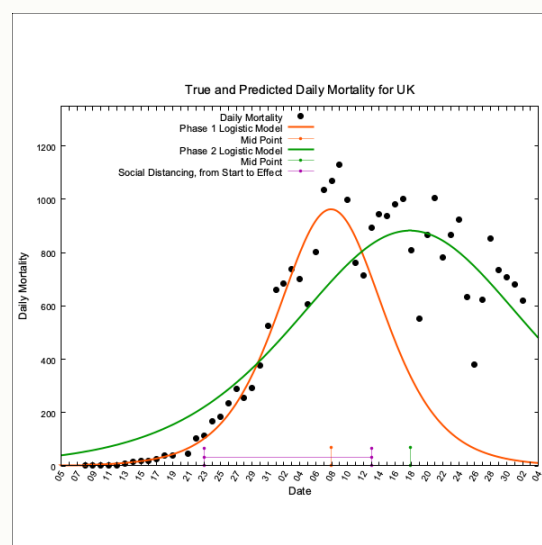
(a) Spain



(b) Italy



(c) France



(d) UK

Figure 8: Daily Mortality, Actual and Predicted

5 Model Predictions

The data for Italy, Spain and the UK appear biphasic. It is not possible to discern a clearly biphasic pattern in the mortality data from France. The following tables contain the model predictions compared with true values.

Spain	Phase 1	Phase 2
Epoch	Before 10 Apr 2020	After 10 Apr 2020
R	-0.962	-0.811
g_0	0.257 (0.237 to 0.277)	0.080 (0.060 to 0.100)
Predicted Total Mortality and (95% CI)	16618 (15789 to 17701)	28407 (26668 to 31909)
Predicted Peak Daily Mortality and (95% CI)	1067 (936 to 1223)	567 (397 to 800)
Peak Daily Mortality Recorded	961	961
Predicted Date of Peak Daily Mortality	31 Mar 2020	07 Apr 2020
Date of Highest Recorded Daily Mortality	02 Apr 2020	
Social Distancing Latency	27 days	

Table 1: Details for Spain

Italy	Phase 1	Phase 2
Epoch	Before 06 Apr 2020	After 06 Apr 2020
R	-0.950	-0.961
g_0	0.196 (0.182 to 0.209)	0.067 (0.061 to 0.073)
Predicted Total Mortality and (95% CI)	18508 (17385 to 19993)	34130 (33008 to 35562)
Predicted Peak Daily Mortality and (95% CI)	905 (791 to 1046)	570 (505 to 645)
Peak Daily Mortality Recorded	919	919
Predicted Date of Peak Daily Mortality	27 Mar 2020	07 Apr 2020
Date of Highest Recorded Daily Mortality	27 Mar 2020	
Social Distancing Latency	28 days	

Table 2: Details for Italy

France	Single Phase
Epoch	N/A
R	-0.924
g_0	0.191 (0.174 to 0.208)
Predicted Total Mortality and (95% CI)	24006 (23054 to 25258)
Predicted Peak Daily Mortality and (95% CI)	1145 (1000 to 1313)
Peak Daily Mortality Recorded	1438
Predicted Date of Peak Daily Mortality	10 Apr 2020
Date of Highest Recorded Daily Mortality	15 Apr 2020
Social Distancing Latency	N/A

Table 3: Details for France

UK	Phase 1	Phase 2
Epoch	Before 13 Apr 2020	After 13 Apr 2020
R	-0.916	-0.875
g_0	0.223 (0.202 to 0.245)	0.102 (0.083 to 0.120)
Predicted Total Mortality and (95% CI)	17223 (15353 to 20219)	34733 (31910 to 39789)
Predicted Peak Daily Mortality and (95% CI)	962 (773 to 1240)	882 (665 to 1191)
Peak Daily Mortality Recorded	1172	1172
Predicted Date of Peak Daily Mortality	08 Apr 2020	18 Apr 2020
Date of Highest Recorded Daily Mortality	21 Apr 2020	
Social Distancing Latency	21 days	

Table 4: Details for UK

6 Discussion

Although we might be pleased with the fit of the model, the conclusions that inevitably follow if it correctly describes the pandemic in these countries, are both surprising and disquieting.

The principle stated aim of social distancing is to reduce peak daily mortality, allow health services to cope with those seriously affected by the disease and so save lives.[9] It

is clear, however, that social distancing has not modified the attainment of the “natural” peak in daily mortality predicted by the phase 1 logistic function (the only logistic function for France) in these 4 countries. It has neither reduced the predicted peak in daily mortality nor has it altered its timing. With an average latency from the implementation of strict social distancing to an effect on mortality of 3.5 weeks, we can see that for it to have had an effect strict social distancing would have had to have been introduced at around the time of the first three deaths. This was nearly achieved in Germany with a closure of schools and colleges at 8 deaths and a blanket “lockdown” at 94 deaths. Their mortality has been much lower. We might introduce the term “Delayed Social Distancing” to distinguish it from “Timely Social Distancing”, used where social distancing occurs much later than the date of the third death from a pandemic condition.[10]

In contrast, the natural resolution of the pandemic has been delayed in Italy, Spain and the UK which now proceed along a second broader logistic. This contrasts with France’s “natural” fast up and down daily mortality peak. The predicted total mortality is higher if based on the second phase logistic (of those countries with a biphasic pattern) than that based on the first. To explore this further the ratio of the growth factors g_0 for phases 1 and 2 was plotted vs. normalised predicted total mortality. $\frac{g_{0,1}}{g_{0,2}}$ is found from the ratio of the gradient of the regression lines for pandemic phases 1 and 2 in the plots of $\frac{1}{D} \frac{dD}{dD}$ and provides a dimensionless measure of the effectiveness of strict social distancing.

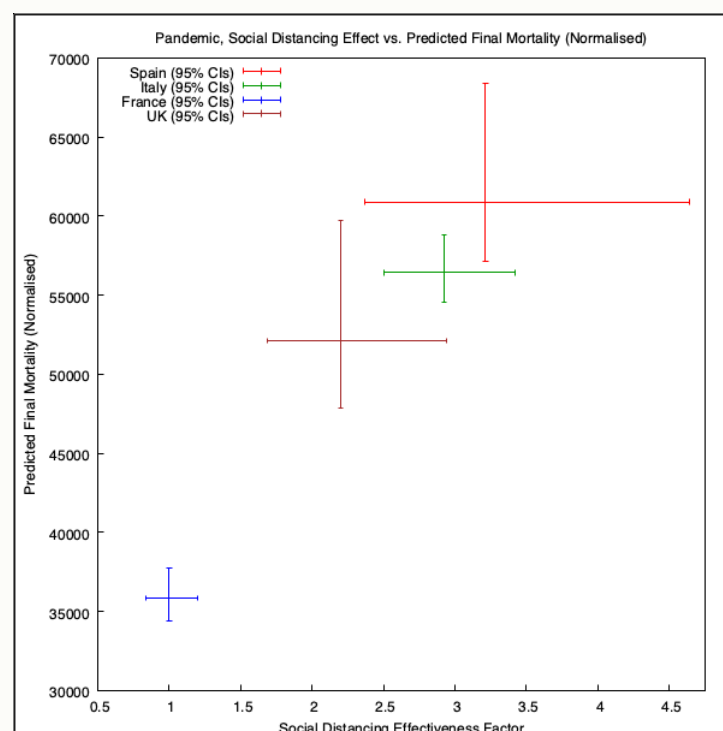


Figure 9: Italy, True and Modelled Cumulative Mortality vs. Date.

The old “...stable door...” adage seems surprisingly apposite. The increase in predicted total mortality with increasing social distancing effect (delayed) is quite unexpected. The pattern of mortality data might provide insight into possible underlying mechanisms. We know that certain groups are at higher risk of dying from Covid-19. Age, obesity, diabetes, gender and socio-economic status are all known factors influencing mortality in addition

to the strong effect from co-morbidities. A low aerobic fitness might be a common thread amongst all these factors and is likely to be adversely affected by a “lock down”. This will be particularly true where the bulk of a person’s exercise arises from the performance of their routine activities rather than any specific training, and where there are limited opportunities at home to maintain fitness because of a lack of space or overcrowding. We already know that a modest increase in 6 minute walking distance (6MWD) has a significant impact on survival in COPD.[11]

7 Conclusions

If the proposed model correctly describes and predicts the evolution of the pandemic in the four countries studied then we may draw the following conclusions:

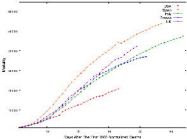
- The mortality from Covid-19 was not growing exponentially prior to the implementation of social distancing measures in any of the countries studied. The mortality data closely followed a logistic curve.
- The gradient of the logistic curve modelling the mortality data culminates in a “natural” peak. Delayed social distancing has had no impact on this “natural” peak of daily mortality which closely fits the phase 1 model peak in all the countries studied.
- To have had an effect on peak daily mortality, one of the stated principle aims of social distancing, it would have had to have been implemented at around the time of the third death from Covid-19. In countries where this was achieved, mortality has been much lower. We might use the terms timely and delayed social distancing to describe this.
- Social distancing, when implemented too late to influence the “natural” peak (as is the case in the four countries studied), appears to delay the subsequent resolution of the pandemic. The area under the curve of daily mortality is increased and so therefore is the final predicted total mortality.
- Delayed social distancing, where effective, appears to increase overall mortality. The increase in overall mortality seems directly related to the effectiveness of social distancing.
- The modified SIR model predicts that not all susceptible persons will be infected. When social distancing is relaxed we should expect a further peak in daily mortality from Covid-19. Mini-peaks of infection have been noted after the relaxation of regional social distancing used to control influenza outbreaks.[12] This is an inevitable consequence of increasing β in our model. It certainly does not validate the use of delayed social distancing, nor should it influence its relaxation.

References

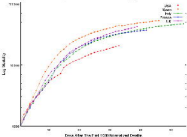
- [1] The Covid Tracking Project. The covid tracking project. <https://covidtracking.com>. Accessed: 2020-05-03.

- [2] Michael Day. Covid-19: identifying and isolating asymptomatic people helped eliminate virus in italian village. *BMJ*, 368:m1165, 2020.
- [3] Yan Bai, Lingsheng Yao, Tao Wei, Fei Tian, Dong-Yan Jin, Lijuan Chen, and Meiyun Wang. Presumed asymptomatic carrier transmission of covid-19. *Jama*, 323(14):1406–1407, 2020.
- [4] Roy M Anderson. Discussion: the kermack-mckendrick epidemic threshold theorem. *Bulletin of mathematical biology*, 53(1-2):1, 1991.
- [5] Dávid Tátrai and Zoltán Várallyay. Covid-19 epidemic outcome predictions based on logistic fitting and estimation of its reliability. *arXiv preprint arXiv:2003.14160*, 2020.
- [6] The UK Survey. Covid-19 serology surveillance strategy. <https://www.reuters.com/article/us-health-coronavirus-britain-tests/uk-to-test-prevalence-of-covid-19-in-general-population-idUSKCN2243IZ>. Accessed: 2020-05-03.
- [7] Atlanta Georgia CDC. Covid-19 serology surveillance strategy. <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/serology-surveillance/index.html>. Accessed: 2020-05-03.
- [8] Worldometer. Worldometer. <https://www.worldometers.info>. Accessed: 2020-05-03.
- [9] GOV UK. Staying at home and away from others (social distancing). <https://www.gov.uk/government/publications/full-guidance-on-staying-at-home-and-away-from-others/full-guidance-on-staying-at-home-and-away-from-others>. Accessed: 2020-05-03.
- [10] Martin CJ Bootsma and Neil M Ferguson. The effect of public health measures on the 1918 influenza pandemic in us cities. *Proceedings of the National Academy of Sciences*, 104(18):7588–7593, 2007.
- [11] Janet B Bowen, John J Votto, Roger S Thrall, Margaret Campbell Haggerty, Rebecca Stockdale-Woolley, Tapas Bandyopadhyay, and Richard L Zu Wallack. Functional status and survival following pulmonary rehabilitation. *Chest*, 118(3):697–703, 2000.
- [12] Richard J Hatchett, Carter E Mecher, and Marc Lipsitch. Public health interventions and epidemic intensity during the 1918 influenza pandemic. *Proceedings of the National Academy of Sciences*, 104(18):7582–7587, 2007.

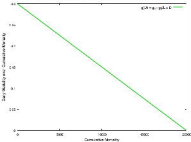
Comparison of Deaths Normalized by Population



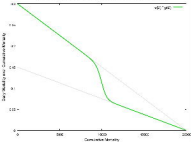
Comparison of Deaths Normalized by Population



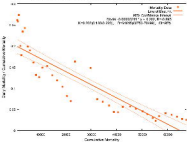
Daily Mortality over Cumulative Mortality vs. Cumulative Mortality



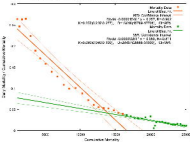
Daily Mortality over Cumulative Mortality vs. Cumulative Mortality



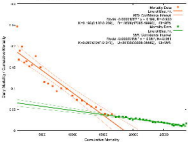
Daily Morality / Comparing Morality vs. Commercial Morality - 10/28



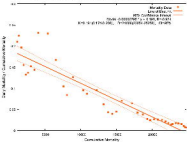
Daily Mortality / Cumulative Mortality vs. Cumulative Mortality - Spain



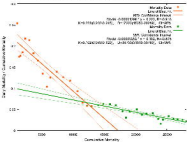
Daily Mortality / Cumulative Mortality vs. Cumulative Mortality - Italy



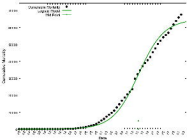
Early Mortality / Cumulative Mortality vs. Cumulative Mortality - France



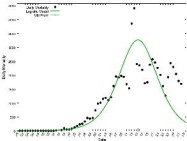
Daily Mortality / Cumulative Mortality vs. Cumulative Mortality - UK



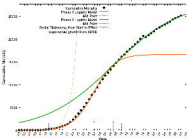
True and Predicted Cumulative Mortality for U.S.A.

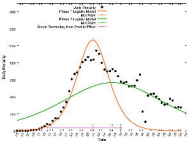


True and Predicted Daily Mortality for USA

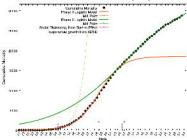


True and Predicted Cumulative Mortality for Spain

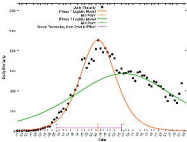


Topic and Prerequisite Data: Locality for Agents

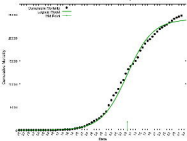
True and Predicted Cumulative Mortality for Italy



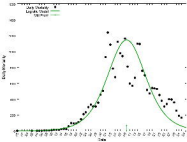
True and Predicted Daily Mortality for Italy



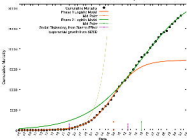
True and Predicted Cumulative Mortality for France



True and Predicted Daily Mortality for France



True and Predicted Cumulative Mortality for UK



True and Predicted Daily Mortality for UK

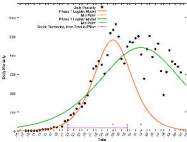


Figure 1: Global Reliability Effects on Predicted Reliability (Normality)

