Supplements to *COVID-19 Confirmed Cases Prediction*
This version: April 8, 2020

Jointly published by

[1] Chair of Entrepreneurial Risks, D-MTEC, ETH Zurich
[2] Institute of Risk Analysis, Prediction and Management (Risks-X), Academy of Interdisciplinary and Advanced Studies, Southern University of Science and Technology (SUSTech)
[3] Gavekal Intelligence Software

Contacts:
Dr. Peter CAUWELS (pcauwels@ethz.ch) Dr. Euan MEARNS (emearns@ethz.ch)
Dr. Michael SCHATZ (mschatz@ethz.ch) Prof. Dr. Didier SORNETTE (dsornette@ethz.ch) Dr. Ke WU (kwu@ethz.ch)

This document contains material supplementing the prediction analysis in a daily report published on https://er.ethz.ch/Covid-19. This version replaces an earlier version from April 07 with an update of Figure 14 and a more detailed description of its methodology.

Content:

Daily growth rate of hospitalization, ICU admittance and mortality in Belgium

In each daily report, we focus on the number of confirmed cases, which is the single metric that is **available** on a daily basis for a wide range of countries and regions. We have mentioned, however, that for a thorough understanding of the Covid-19 epidemic, it is essential to

1. keep track of other key variables such as **mortality, hospitalizations** and **ICU admittance**;
2. and make a serious effort to relate **reported** to **actual** numbers (which, for the confirmed number of cases indicates a careful consideration of testing rates).

Figure 1 reports the daily growth rate of hospitalization, ICU admittance and mortality in Belgium. Looking at the data on confirmed cases only, Belgium followed exponential growth with stable daily growth rate until recently. This, however, is heavily influenced by the varying number of tests conducted. On the other hand, the daily growth rates of hospitalization, ICU admittance and mortality have been decreasing throughout, and provide a much clearer picture of Belgium approaching the peak of the epidemic.

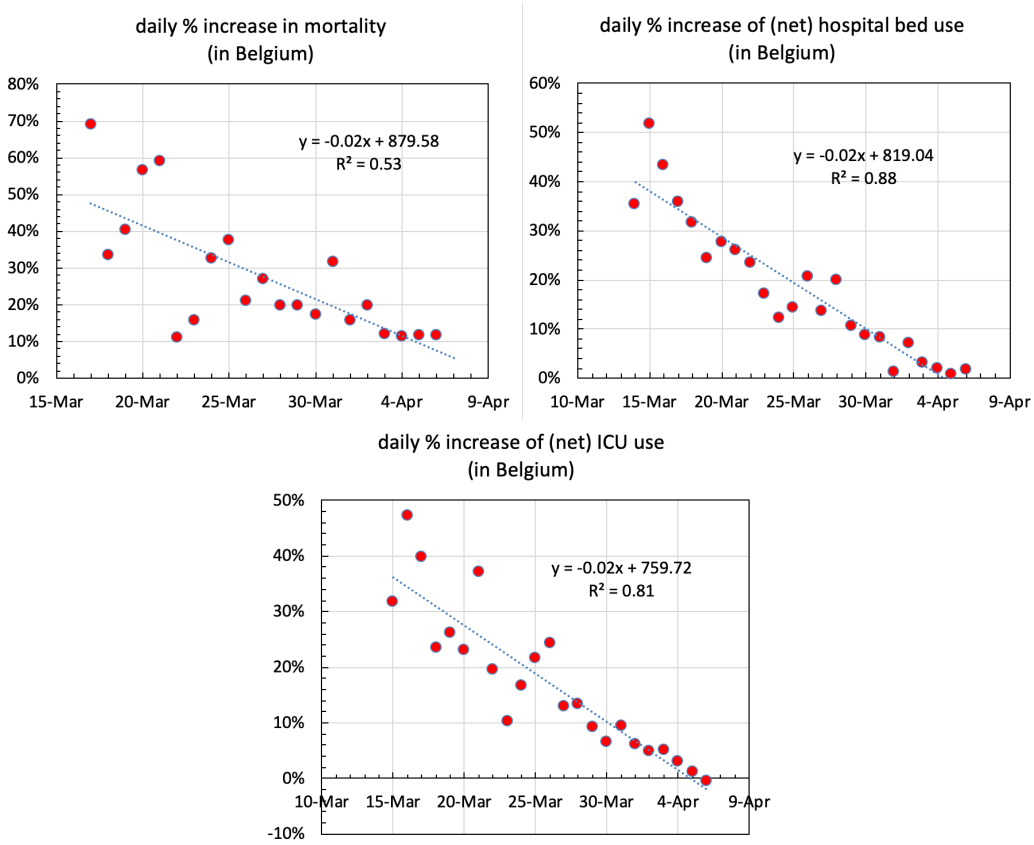

Figure 1. Daily growth rate of mortality, hospitalization and ICU admittance in Belgium.

Relation between the number of confirmed positive cases in the US to the number of tests conducted

Based on the findings for Belgium, in Figure 2, we try to relate the number of confirmed positive cases in the US to the number of tests conducted.[1] In particular, we normalize the number of reported cases with the number of tests to get a better understanding of the growth rate of actual infections in the population. We find that a large part of the exponential growth of reported cases may be due to an exponential increase in testing. To make such an analysis more reliable, random testing (polling) in the general population would be necessary (and is highly recommended).

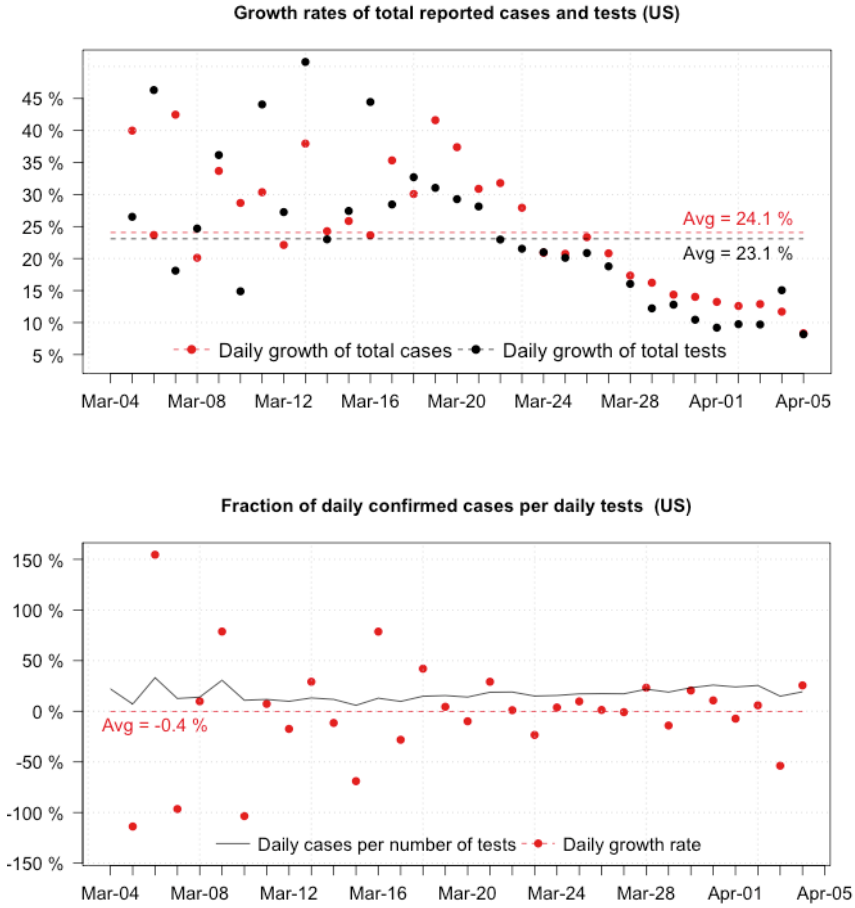

Figure 2. Upper plot: Growth rates of confirmed cases and of total tests in the US; lower plot: daily fraction of confirmed cases per number of tests and daily growth rate of confirmed cases.

[1] Using data from https://covidtracking.com.

A quick look at the death per capita in countries from different regions of the world.

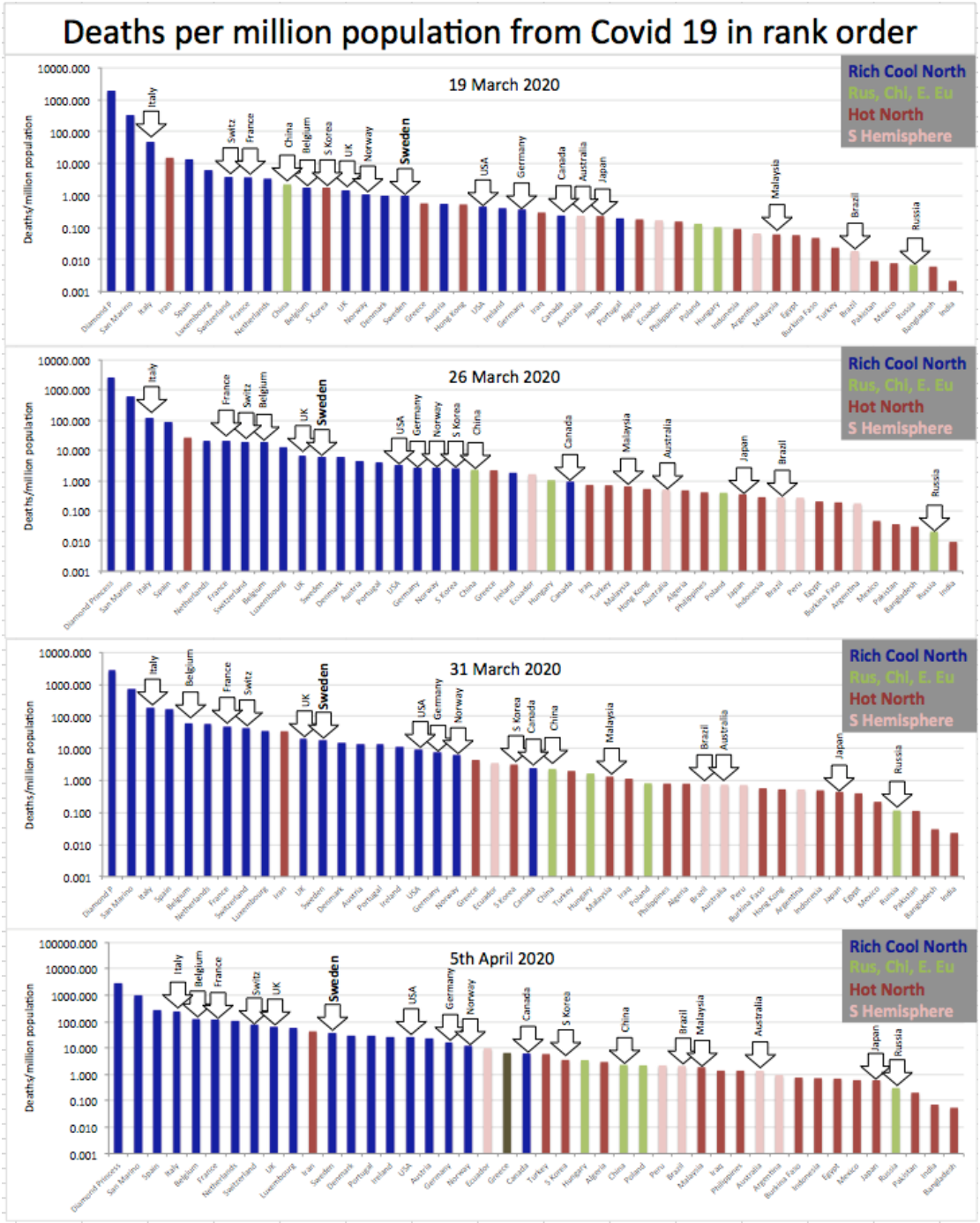

Figure 3. Total deaths per million population for selected countries ranked by death as of 19, 26, 31 March and 5 April 2020, which are subjectively divided into 4 groups: Rich North countries (dark blue), Hot North countries (dark red), Southern Hemisphere countries (pink), and the East Block countries (Russia, China and East Europe) in pink. This view provides an account of the severity of the epidemic in different regions. It seems that the epidemic mainly lies in Rich North countries, while Hot North, East Block countries and southern hemisphere countries do not have large scale epidemics yet. There are several possible (and highly speculative) explanations for this, see the detailed break down below. An exception is Iran, whose classification in Hot North countries is likely incorrect as a large part of the population lives in high plateaus and mountains where the weather is cold, confirming the likely role of temperature and humidity in the severity of the epidemics.

1. First we need to declare an error on Sweden (report of April 1): unfortunately, we used a wrong data point to normalize for population. Sweden has puzzled us with the lightest lockdown in Europe and in earlier reports the lowest deaths per capita that made no sense.
2. Now that we have fixed that error, we provide summaries of the rank order at 4 time horizons – 19 March, 26 March, 31 March and 5 April (Figure 3). For each time horizon Sweden is located adjacent to Scandinavian neighbor, Denmark. Since 26th March, Sweden's position in the rank order has barely changed (position 11, 11, 12) and it sits in the middle of its peer group of Rich Cool North countries. There is no evidence to suggest that Sweden's light lockdown policy is having detrimental outcomes on mortality. Admittedly, Norway is doing a little better, but this could be linked to geography and other factors such as a smaller urban population. The UK, with draconian lock down measures, is working up through the rank order, with worsening performance relative to Sweden.
3. All countries to the right of Ecuador have <10 deaths per / million population (0.001%).
4. The Hot North countries do not have large epidemics yet. Greece remains the worst case in the hot North with 6.52 deaths / million compared with 4.41 a week ago. Notably Greece lies at the boundary of hot north and rich north in the rank order reflecting also its geographic location.
5. The East Block countries do not have large epidemics yet either. The worst country is Hungary with 3.52 deaths per million compared with 1.66 deaths / million a week ago. Many of these countries have climates similar to West Europe. There are three options to explain this.
   - The early lock down prevented the spread of the disease;
   - They are simply at an earlier stage and a larger epidemic is coming (not applicable to China);
   - They are using a different reporting standard, for example recording cause of death as pneumonia, or different treatment.
6. China continues to march to the right through the ranks. The horror story of January and February that everyone in the west watched is going to turn out to be a fairy tale for China. While the rich north countries wither on the vine. How all of China outside of Hubei Province allegedly escaped this pandemic is a burning question.
7. The S Hemisphere doesn't have a large epidemic yet either. The worst hit country is Ecuador with 9.75 deaths / million compared with 3.51 deaths / million a week ago – near trebling is a sign of acceleration. Ecuador is now adjacent to Norway in the ranking and the cold Southern Hemisphere countries should pay close attention to what happened in the cold rich North. But Ecuador lies on the Equator, but is also a high altitude Andean nation. Looking at the charts we can see the pink colours of the S Hemisphere countries moving to the left through the ranks of Hot North countries. This could reflect the changing seasons as the Northern hemisphere moves into Spring and the southern hemisphere moves into Autumn. If the epidemic comes in the Southern hemisphere winter, they have time to prepare. Locations like S Chile, Patagonia and S Island New Zealand could be most at risk.
8. The large epidemic lies in Rich Cool North countries where land locked San Marino is by far the hardest hit with 941 deaths / million. With 32 deaths, the population is reduced from 34,000 to 33,968 (0.09%).
9. Of the large countries Italy (254) and Spain (266) swap places, with Spain now the worst hit large country. Italy's epidemic is now waning with the growth rate in deaths declining. The rank order of the Rich North countries in part will be affected by the stage of progression of the outbreaks. Being a large country with low population density and isolated communities might help in the cases of Norway and Canada. The reason that Rich North countries have been hit so hard could be due to the cold climate and aging populations, as well as a remarkably mild winter with a much smaller death rate due to the "normal" influenza in the vulnerable elderly cohort.
10. Iran is the only hot north country to have a large epidemic. It was cold over Teheran during the early stage and a large part of the population lives in high plateaus and mountains where the weather is cold. And we speculate that old men sharing bubble pipes may have spread the virus among the vulnerable who then imported it to the family home.

11. India, Pakistan and Bangladesh are all big countries that should have been in the thick of it early on given their location. They might not have enough resources to test and find the cause of death. But they appear to be seeing something going on now since India and Pakistan have gone into lockdown. We speculate that the very low death rate in India, Pakistan and Bangladesh is in part because there populations lack large cohorts in the over 65 groups where most deaths in the Rich North occur (Figure 4).

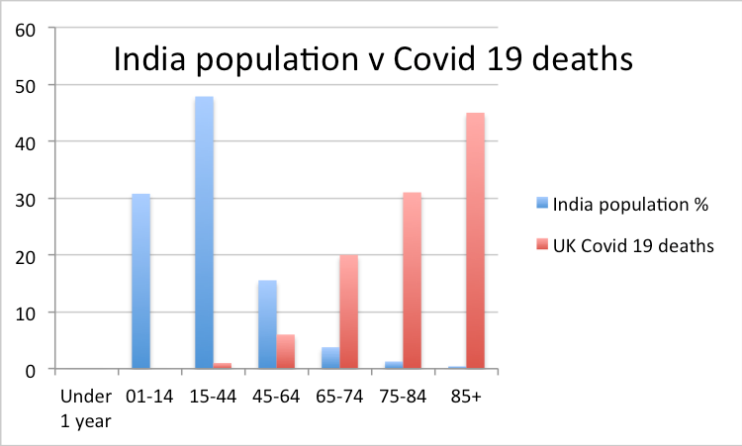

Figure 4. Demographics of India and distribution of Covid-19 deaths in the UK.

Per Capita Mortality Trends

Summary: We present an analysis of cumulative mortality trends for 10 countries – Italy, Spain, France, Switzerland, Belgium, UK, Portugal, USA and Germany – based on mortality statistics normalised to population size, reported as deaths per million people. The curves are displaced in time because the epidemics broke out at different times in these countries with a spread of ~ ±10 days. We describe a method where we attempt to remove these time lags. France, Switzerland, UK, Portugal, USA and Germany all seem to be following similar trends suggesting that variations in lockdown policies are having minimal impact on outcomes. Italy is doing a little worse, and Spain and Belgium currently appear to have the worst mortality rates. This does not necessarily mean that they will have the worst final outcomes.

Methods: We present cumulative mortality curves for 10 countries. Mortality data have been recorded from the John's Hopkins dashboard [1] on a daily basis since 20 February 2020. For the USA, we use the aggregate from nCoV2019 Live [2]. Data are normalised to deaths/million population (deaths / mil) using the UN population statistics [3].

The outbreaks across the world began at different times. We therefore try to adjust for date differences by time-shifting the curves to align them on a particular value of deaths/mil. We performed this exercise on 0.1, 1, 4 and 10 deaths per mil and found that 1 and 4 gave the best curve matches, 4 being marginally better than 1. The time offsets are consistent with ~±10 days variance in the stage of outbreak in the selected countries (Figure 9). As one final iteration, we took the fit produced using the 4 deaths / mil datum (Figure 5) and adjusted the time displacements by eye to produce the best overall fit of the curves (Figure 7).

We selected 10 countries for analysis. This exercise began from an interest to compare Spain with Portugal. We then added our home countries – Switzerland, France, UK and Belgium. We then supplemented with other important countries – Italy, Germany, USA and S Korea. In future we may add more countries but stress that many countries (E Europe and S hemisphere) are still at a very early stage making trend comparisons impossible at this time.

Results: Figures 5-6 present the evolution of cumulative number of deaths per million population, time-shifted to align at 4 deaths/million. This improves comparability of the effectiveness of lockdown strategies in countries that are in different stages of the epidemic. Italy, where the epidemic started early, provides a reference for the evolution of deaths in different European countries and the US (group 1 - Rich North countries). For South Korea, as an example of a country in group 2, we see a substantially different trajectory. For more details, key observations and conclusions; see the text below.

It has been clear from the onset of the epidemic that Covid-19 disproportionately affects the elderly population in terms of mortality. Representative for the situation in other countries, in Figures 7 and 8 we show the distribution of deaths for the UK and Italy and find a mean age of death equal to 80.2 and 78, respectively. Note that these countries are in different stages of the epidemic. To get a better picture of the relative likelihood to die from Covid-19 for different age groups, in Figure 9, we show an estimate of the likelihood of death given infection. To account for the high number of unreported cases in most countries, we use data from Iceland, where a broad testing of the general population was conducted. This allows for a different perspective on reported case fatality rates of Covid-19.

[1]https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6

[2] https://ncov2019.live/data

[3] https://population.un.org/wpp/Download/Standard/Population/

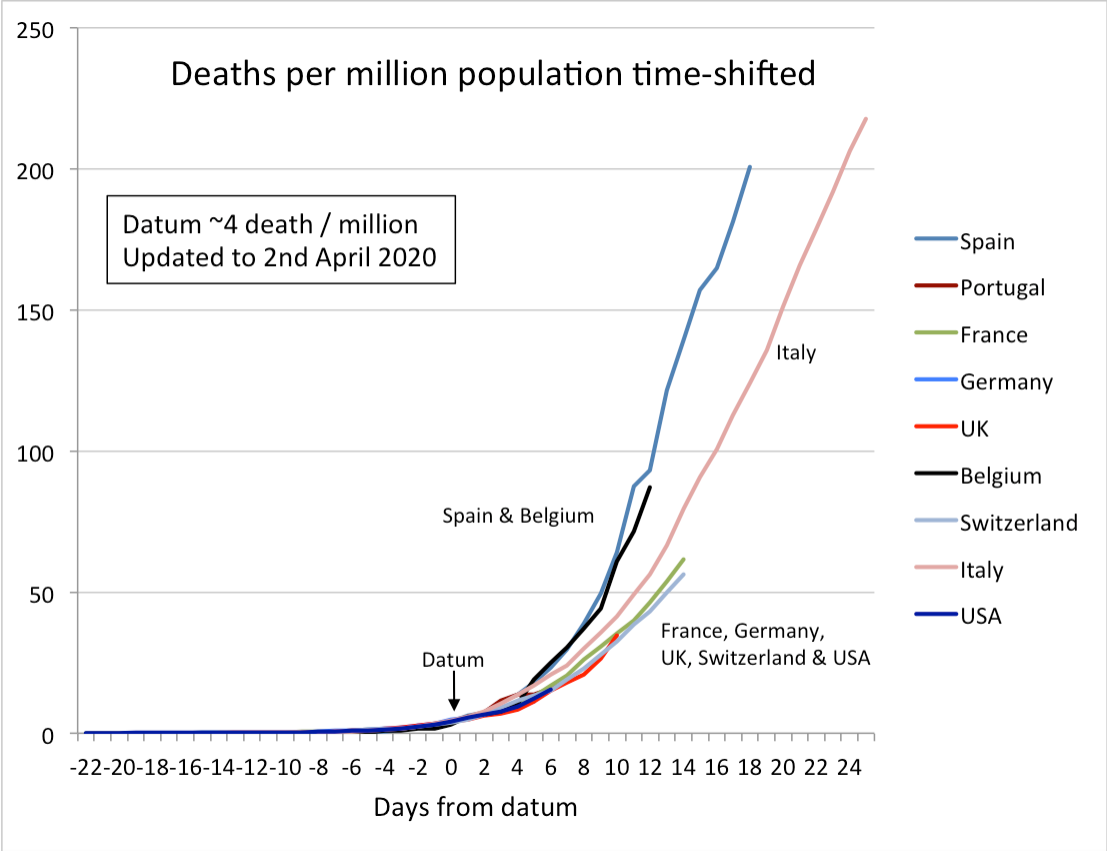

Figure 5. Cumulative mortality curves for 10 countries aligned on the day each country reached 4 deaths / mil (the datum). For comparing curves we have a strong preference to use this linear plot. A log scale plot is shown in Figure 6. S Korea has not yet reached 4 deaths / mil and is not plotted here but it appears on Figure 7.

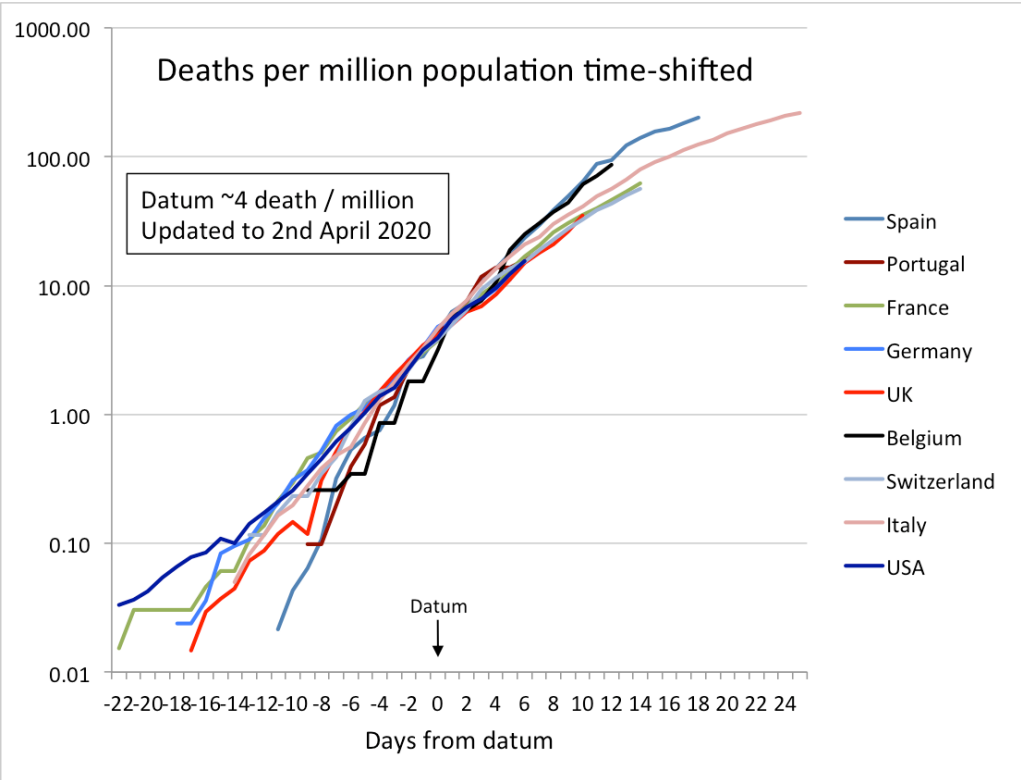

Figure 6. Log plot of the data shown in Figure 5.

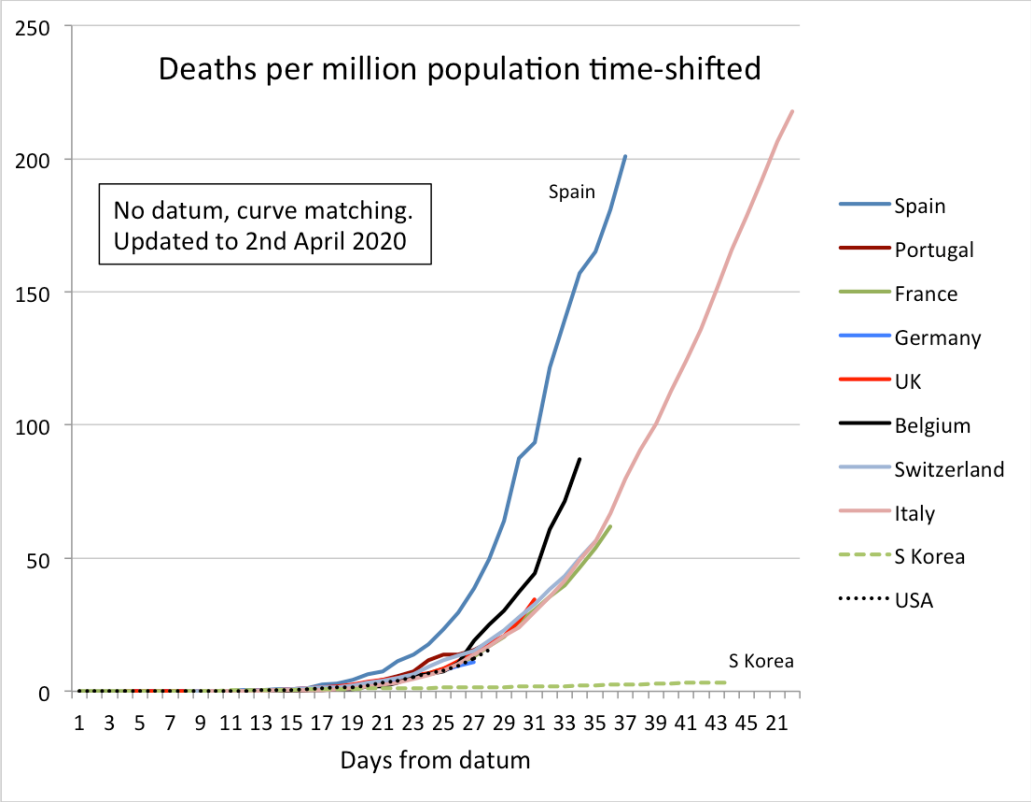

Figure 7. This plot is an adjusted version of Figure 5 where curves were time shifted to produce the best visible match. The time displacements relative to Figure 5 are small, usually just 1 or 2 days (Figure 9). This plot illustrates how different the outbreak has been in S Korea. We have reasons to believe that China, Japan, and Singapore will follow similar trajectories to S Korea.

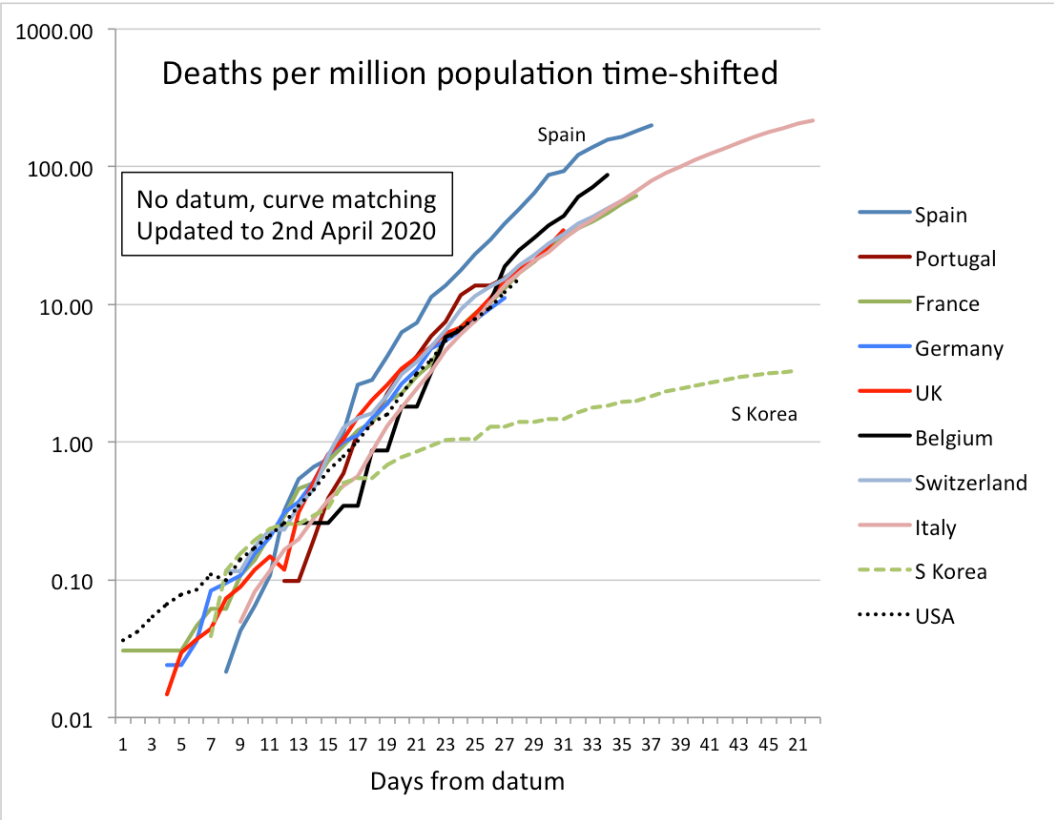

Figure 8. Log plot of the data shown in Figure 7.

	Days Lagging Italy		
	4 / million datum	Curve matching	Difference
Italy	0	0	0
Spain	7	11	4
France	11	12	1
Switzerland	11	13	2
Belgium	13	14	1
UK	15	17	2
Portugal	18	20	2
USA	19	20	1
Germany	20	21	1

Figure 9. This figure effectively shows the order that countries became seriously infected relative to Italy that was first. The order is the same for both the 4 / mil datum and the curve matching methods. The time offsets are small, normally 1 to 2 days.

Discussion

Looking at the 4 deaths / mil normalised data (Figure 5) we see 3 groups. France, Germany, UK, USA, Switzerland and the USA are all following similar trajectories suggesting that variances in the lockdown strategies in theses countries is making little difference to the mortality outcomes so far. Italy is on a slightly steeper trend and Spain+Belgium on a steeper trend still.

Looking at the time adjusted "curve matching" chart, it is possible to converge Italy with the France, Germany, UK, USA, Switzerland and USA group. It is possible to pull Belgium towards this group but impossible to achieve full convergence as is the case for Spain.

It would be a mistake to assume that Spain and Belgium are doing less well than other countries at present since the trajectory could come down to those countries following steeper but narrower curves.

S Korea appears on Figure 7 and stands out as being totally different to the European and N American countries. There has been no large-scale epidemic recorded in S Korea. The reasons for this could include:

1. Warmer climate suppresses the disease
2. Genetic / racial differences are reflected in higher immunity
3. Different social standards, for example the Japanese eat a very healthy diet and features like obesity are near absent in their population
4. Strong adherence to social distancing, masks and other protective measures (without quarantine or lockdown)
5. A different standard is applied to how deaths are recorded. These countries are perhaps following what we understand to be the correct medical protocol and record the cause of death as pneumonia (or a variant thereof)

More detailed understanding about why these East Asian countries have such low deaths is part of our on-going research.

Our intention is to use this analysis to try and estimate the value of measures taken to abate the crisis compared with the costs of abatement. Over time, we look forward to adding Sweden to this analysis since Sweden has had a much lighter, and therefore less costly lock down.

Age distribution of deaths

The UK has begun to report the demographic profiles of those dying from Covid 19 (https://www.ons.gov.uk). The first report was for week 12 (ending 20 March), see Figure 10 (and figure 11 for Italy). There is promise of more detailed reports to come. It has to be noted that week 12 is a very early stage of the epidemic in the UK. Similar to other countries, Covid 19 is killing the elderly to very elderly in the UK. The mean age of death is computed to be 80.2 years compared with life expectancy in the UK of 81 years. In this calculation an age of 90 years is assumed for the 85+ cohort

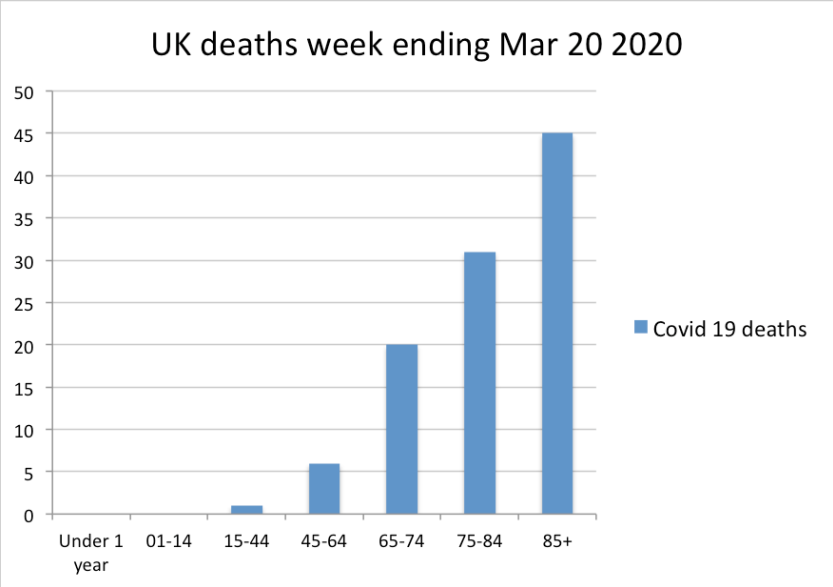

Figure 10. Age distribution of confirmed deaths in the UK (Mean: 80.2), Source: https://www.ons.gov.uk and https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/weeklyprovisionalfiguresondeathsregisteredinenglandandwales

Case fatality rates in perspective

To reach a conclusion regarding the mortality of Covid-19 for different age groups, it is essential to normalize death numbers with the number of infections to get the case fatality rate. Iceland has been the only country so far to take a broad sample of tests in the general population to get an unbiased view of infection numbers. In Figure 12, we combine this (see below for method) with the age distribution of deaths in Italy (Figure 11) to derive an estimate of the likelihood to die from Covid-19. More precisely, the likelihood of dying once infected at a given age is given by the ratio of the distribution of ages for those that died divided by the distribution of ages over the infected population.

Note that these numbers do not reflect absolute probabilities and only allow interpretation of relative occurrence (for example, someone in the age group 80+ is roughly 30 times as likely to die upon infection than someone in the age group 60-69). If we compare this with the reported case fatality rate from Italy, we get an idea of the bias introduced by unreported cases of infection in Italy (see Figure 13).

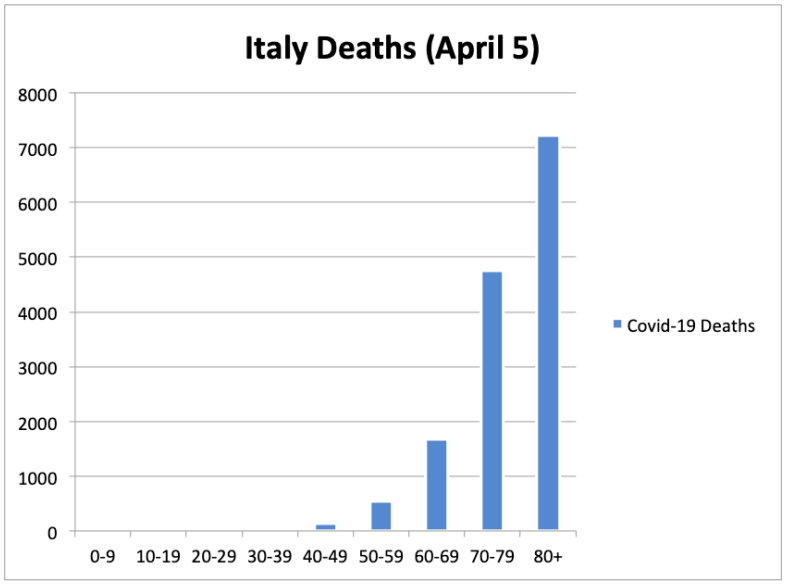

Figure 11. Age distribution of confirmed deaths in the Italy (Mean: 78), Source: https://www.epicentro.iss.it.

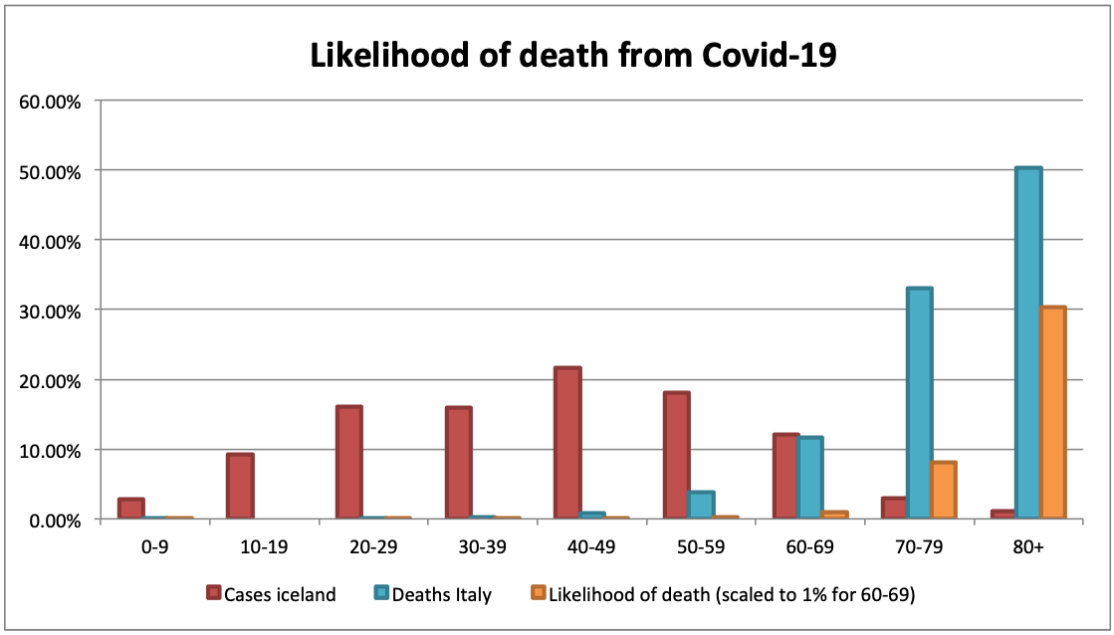

Figure 12. Likelihood of death from Covid-19, Source Island: https://www.covid.is/data. The likelihood of dying once infected at a given age is given by the ratio of the distribution of ages for those that died divided by the distribution of ages over the infected population. For example, someone in the age group 80+ is roughly 30 times as likely to die upon infection than someone in the age group 60-69.

.

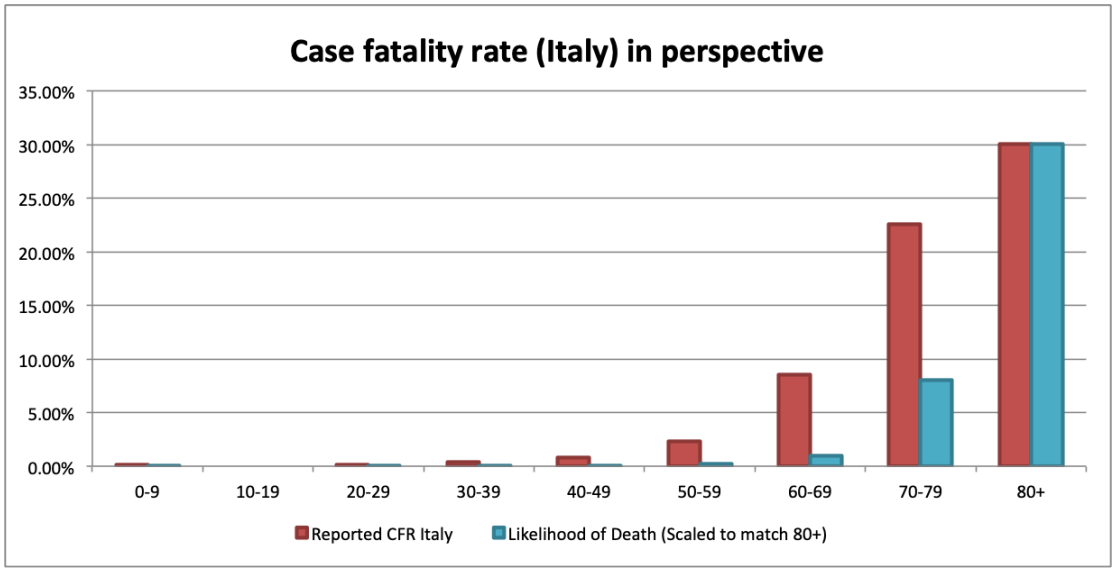

Figure 13. Reported case fatality rate in Italy compared to the likelihood of death as calculated in Figure 12. Source: https://www.epicentro.iss.it. The likelihood of dying once infected at a given age is given by the ratio of the distribution of ages for those that died divided by the distribution of ages over the infected population.

Fatality given severe symptoms

One can use a similar approach to get an idea of the fatality of Covid-19 given that someone shows severe symptoms. This has implications for the focus of medical attention if resources are scarce. We use reported cases in the Netherlands (April 2) as a distribution for showing severe symptoms in a given age group. This can be rationalised by the fact that the Netherlands initially tested very narrowly and only cases that were hospitalised. By combining this (see below for method) with the distribution of deaths from Italy, we see that the likelihood of recovering from severe symptoms is substantially larger for younger ages (Figure 14). Note that, as in Figure 12, the numbers can only be interpreted as relative occurrence (for example, with severe symptoms someone in the age group 70-79 or 80+ is roughly 45 times as likely to die than someone in the age group 30-39).

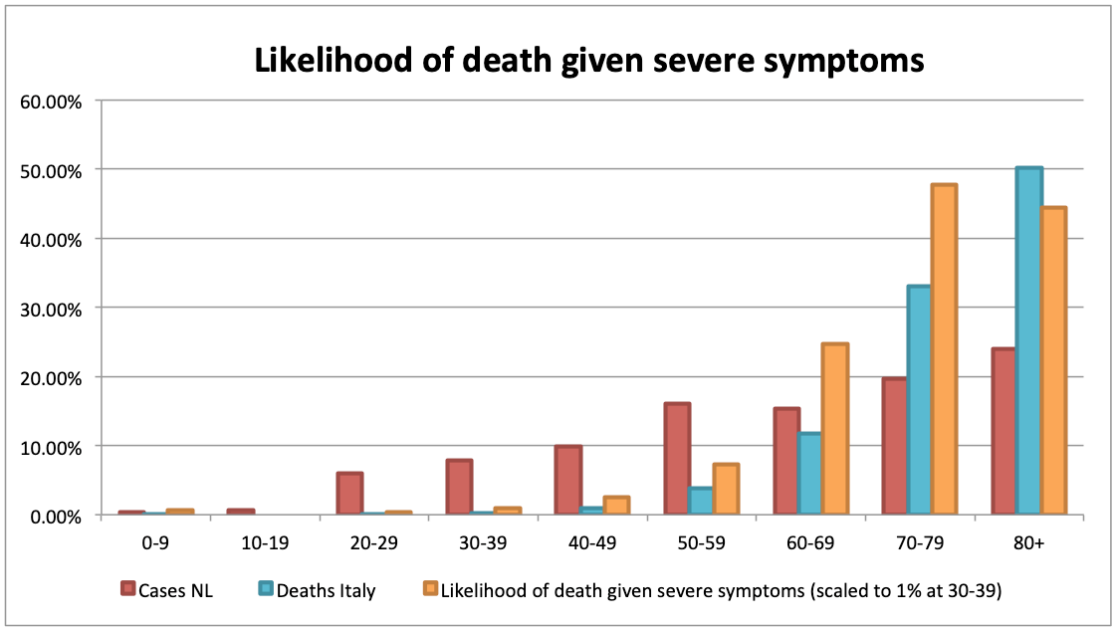

Figure 14. Likelihood of death given severe symptoms, Source Netherlands: https://www.rivm.nl.

Method to construct figures 12-14:

Among a population of recorded Covid-19 cases, we can write the case fatality rate

(1) p(death|age) p(age) = p(age|death) p(death)

where
-p(age|death) is the age distribution of deaths from Covid-19, for which we use the data from Italy, being in a mature state of the epidemic.
-p(age) is the distribution of ages of Covid-19 cases, which we propose to proxy by the Iceland distribution, and account for the demographic difference of Iceland and Italy to get an estimate of the distribution of ages of Covid-19 cases in Italy,
-p(death) is the age-unconditional probability of death given one is tested positive => this is just a constant.
From (1) and our proposition for the proxies, we get

(2) p(death|age) = C distribution_Italy(age|death)/distribution_Iceland(age).

In other words, the ratio of the two distributions gives us, up to a normalisation coefficient, the probability of death as a function of age.

For figure 14, we use
-p(age|severe) as the distribution of ages of Covid-cases with severe symptoms, which we propose to proxy by the Netherland distribution based on narrow testing of hospitalized cases. Combining this with the arguments that led to (2) – and account for the demographic difference of Italy and the Netherlands, we get

(3) p(death|severe,age) = C distribution_Italy(age|death)/distribution_Netherlands (age).