UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH
Computational Biology and Complex Systems Research Group
CMCiB Comparative Medicine & Bioimage Centre of Catalonia
IGTP Germans Trias i Pujol Research Institute

Daily report 04-12-2020

Analysis and prediction of COVID-19 for EU-EFTA-UK and other countries

Situation report 163

Contact: clara.prats@upc.edu

With the financial support of European Commission *and* Ministerio de Ciencia e Innovación, "la Caixa" Foundation

Foreword

The present report aims to provide a comprehensive picture of the **pandemic situation of COVID-19** in the EU countries, and to be able to foresee the situation in the next coming days. We provide some figures and tables with several **indexes and indicators** as well as an **Analysis** section that discusses a specific topic related with the pandemic.

As for the predictions, we employ an **empirical model**, verified with the evolution of the number of confirmed cases in previous countries where the epidemic is close to conclude, including all provinces of China. The model does not pretend to interpret the causes of the evolution of the cases but to permit the **evaluation of the quality of control measures made in each state** and a **short-term prediction of trends**. Note, however, that the effects of the measures' control that start on a given day are not observed until approximately 7-14 days later.

We show an individual report with 8 graphs and a summary table with the main indicators for different countries and regions. We are adjusting the model to **countries and regions** with at least 4 days with more than 100 confirmed cases and a current load over 200 cases.

Martí Català
Pere-Joan Cardona, PhD
Comparative Medicine and Bioimage Centre of Catalonia; Institute for Health Science Research Germans Trias i Pujol

Clara Prats, PhD
Sergio Alonso, PhD
Enric Álvarez, PhD
Miquel Marchena, PhD
David Conesa
Daniel López, PhD
Computational Biology and Complex Systems; Universitat Politècnica de Catalunya – BarcelonaTech

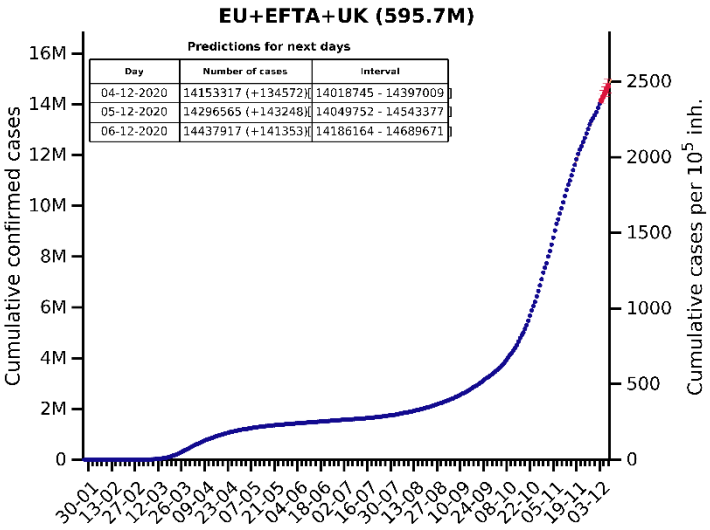
With the collaboration of: Daniel Molinuevo, Pablo Palacios, Tomás Urdiales, Aida Perramon, Inmaculada Villanueva

These reports are funded by the European Commission (DG CONNECT, LC-01485746)

PJC and MC received funding from "la Caixa" Foundation (ID 100010434), under agreement LCF/PR/GN17/50300003; CP, DL, SA, MC, received funding from Ministerio de Ciencia, Innovación y Universidades and FEDER, with the project PGC2018-095456-B-I00;

Situation and highlights

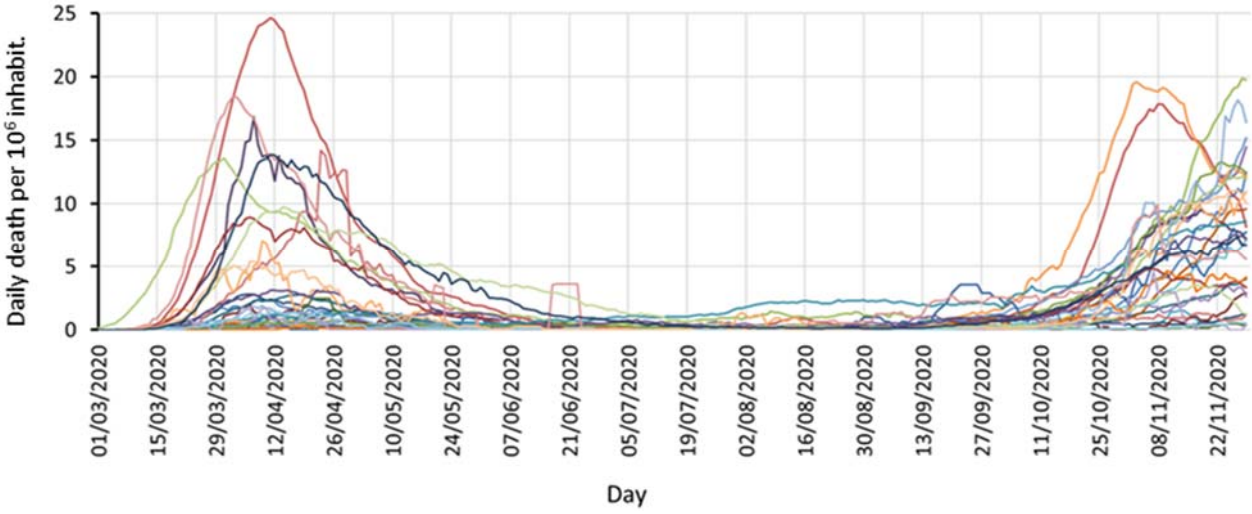
Global situation



Data on deaths caused by Covid-19 can be very different from one country to another, as different criteria can be followed to account for deaths. This makes it very difficult to make comparisons between countries. However, it is interesting information that gives us a picture of the epidemiological situation. To assess the different epidemiological situations, we will look at the number **of daily deaths per 10^6 inh**. This is a striking parameter, but surely a good picture of reality.

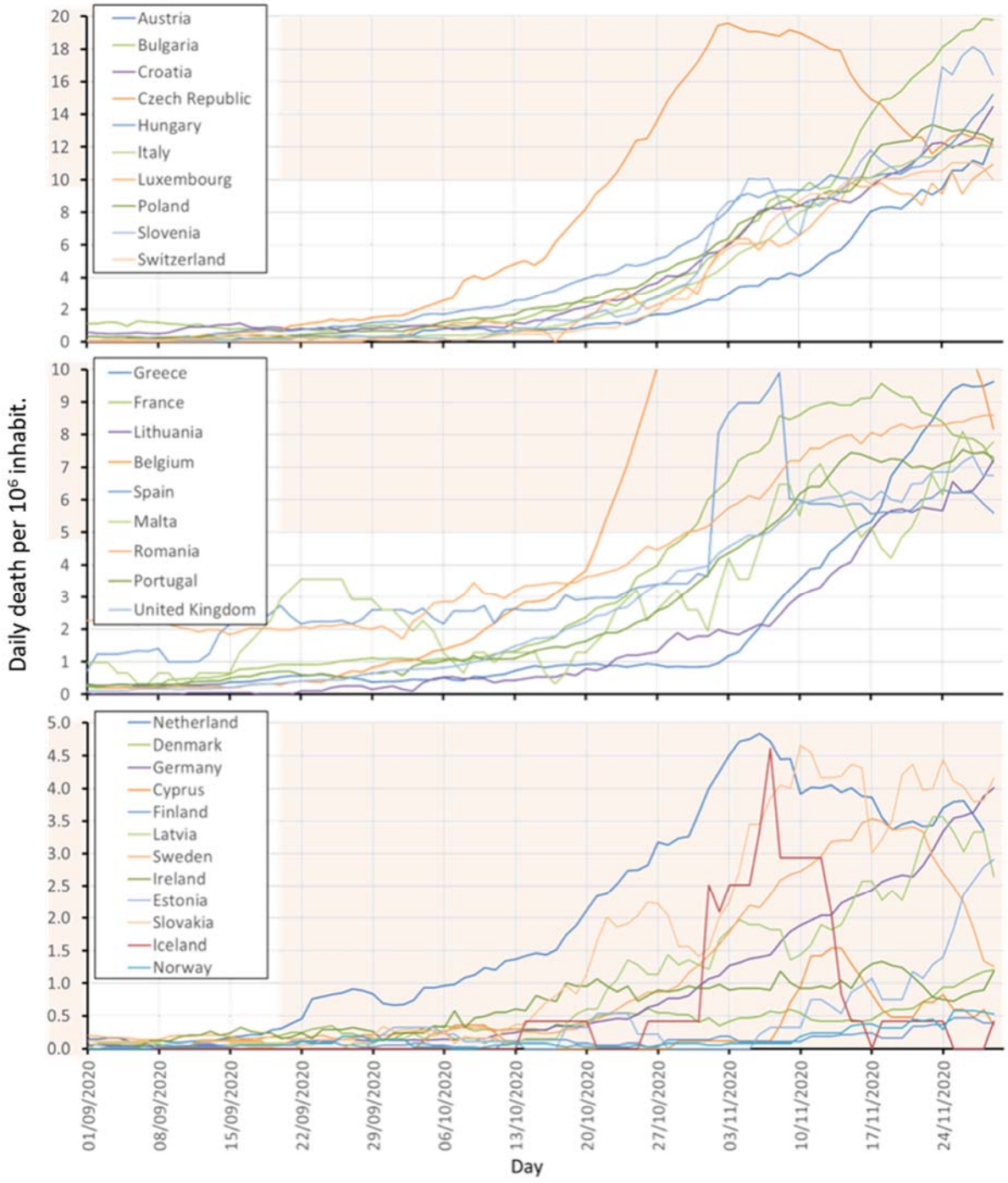
In a first figure we can make a general comparison between the situation of the first months of the epidemic and the current situation, from September 1st. Between the end of March and the middle of April, some countries reached really high values. Belgium, at the peak came to report almost 25 daily deaths per 10^6 inh., Spain 18, France 17 and UK 14. During the autumn the maximum values reached have been also very high Czech Republic and Bulgaria, with almost 20 daily deaths per 10^6 inh., Slovenia 18, and Belgium 17. Looking at these data we see that the intensity of the two waves has been different in each country, but overall the **second wave is being as devastating as the first**.

If we consider together all the countries of the EU+EFTA+UK, around 7 April the value of 8 daily deaths per 10^6 inhabitants was reached, almost the same value that was reached around 27 November (7.8). These values reiterate that the second wave is clearly of the same magnitude as the first.



If we focus on the current situation, as we can see in the following figures (next page), we can divide the countries into three groups. In the first group, countries with an incidence of less than 5 deaths per day per 10^6 inhab. especially noteworthy are Finland and Norway, which since September have never exceeded the value of 0.6 deaths per 10^6 inh. In the second group, with values between 5 and 10, what seems worrying is the fact that mostly countries show a fairly stable behaviour. If the epidemiological situation is prolonged over time, the effects, such as deaths, will accumulate. Finally, the group of countries with the highest mortality, between 10 and 20. The situation of these 10 countries is worrying. Bulgaria stands out with a

value of almost 20, and Slovenia with 18, although Slovenia seems to have started a process of improvement. Of concern is the growth shown by Hungary and Croatia with values above 14 and Austria with 12.



Highlights

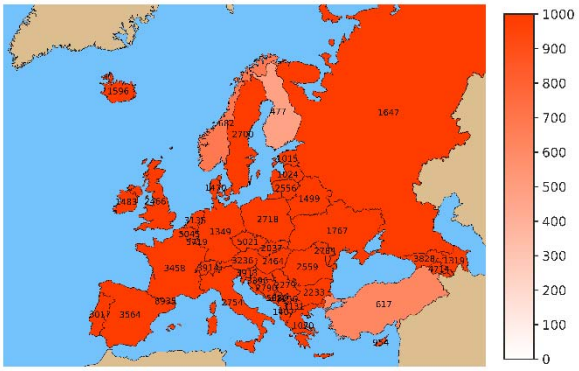
- We are **progressively changing the methodology of assessing the reproduction number and the predictions** so that **anomalies in countries reporting and weekend effect are minimized**. This may affect some parts of the reports these days that will be resolved soon. The methodology will be updated once it has been fully implemented.

Situation and trends per country

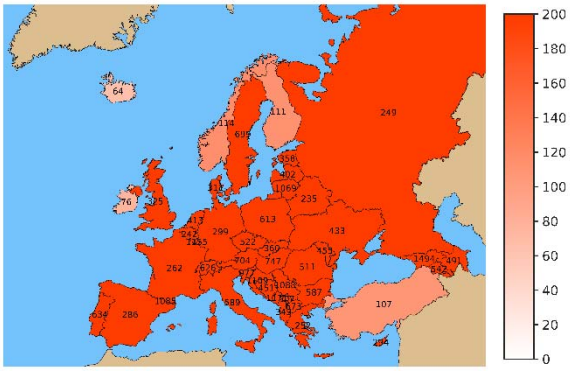
Maps of current situation in EU countries. Colour scale is indicated in each legend.

- Cumulative incidence: total number of reported cases per 100,000 inhabitants
- A_{14}: Cumulative incidence last 14 days per 100,000 inhabitants (active cases)
- ρ_7: Empiric reproduction number
- EPG: Effective Potential Growth ($EPG = A_{14} \cdot \rho_7$)

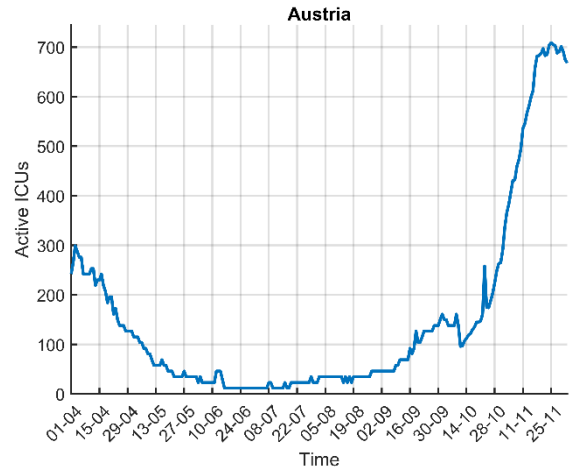
Cumulative incidence

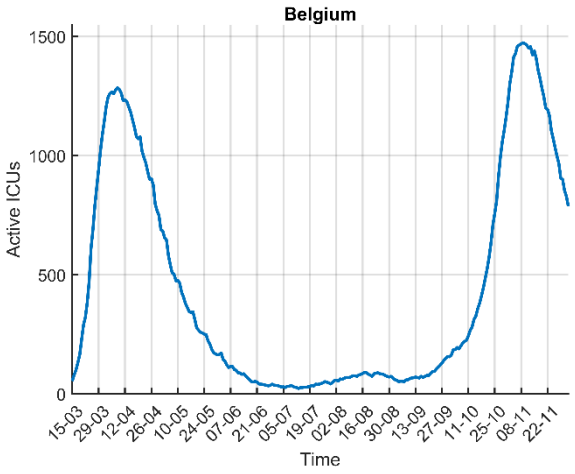


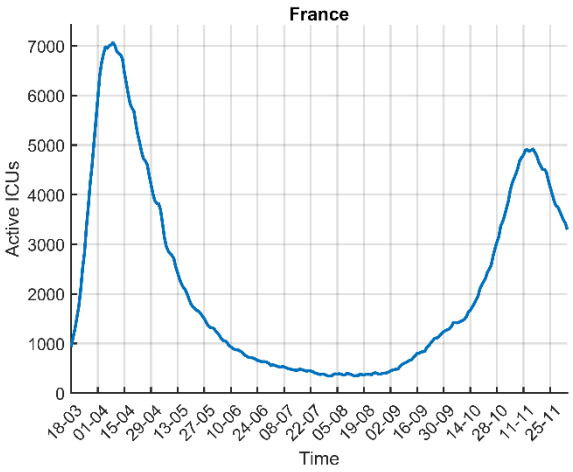
A_{14}

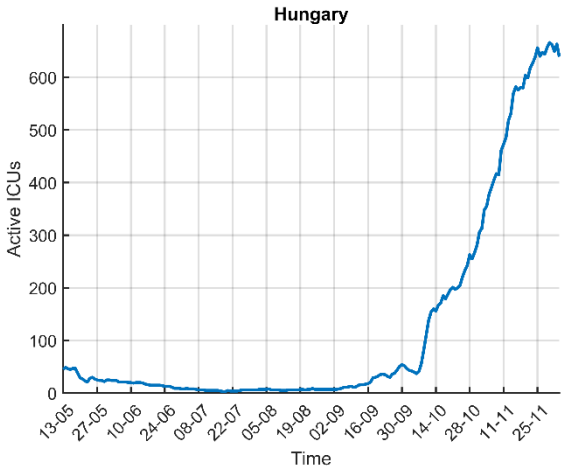


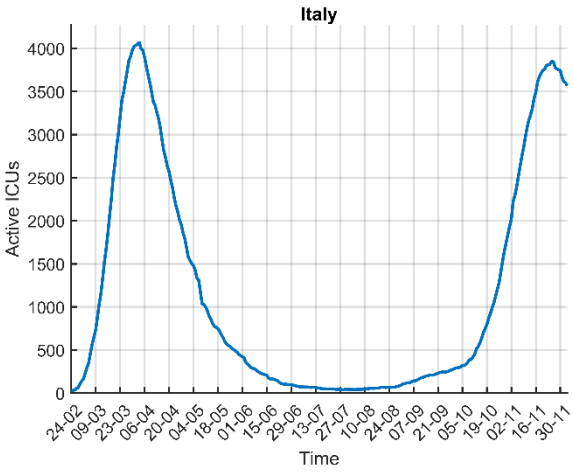
Evolution of of active ICUs in some EU countries.

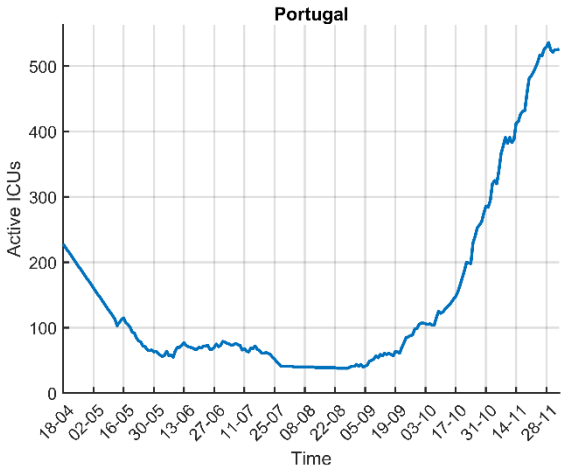




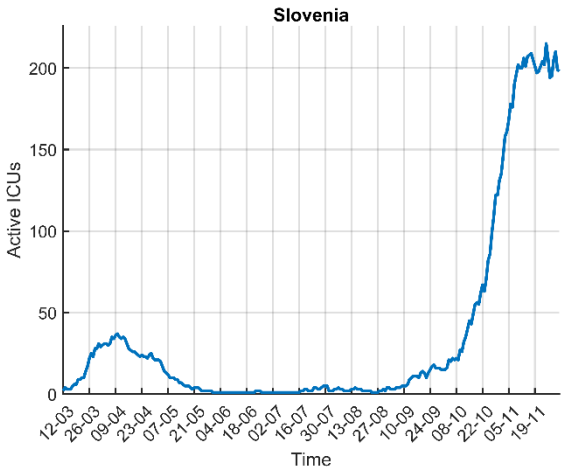


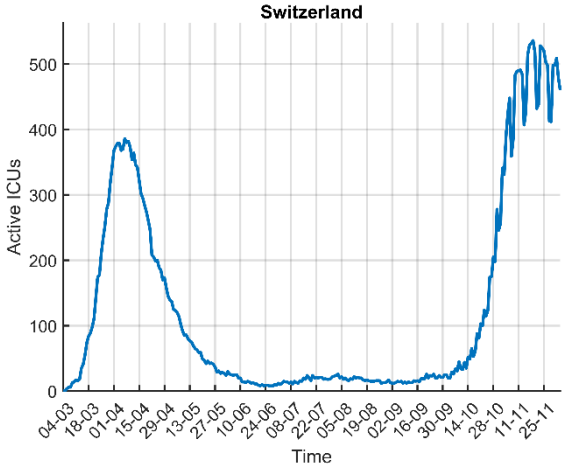







Slovakia
Active ICUs
0
20
40
60
80
100
120
24-04
08-05
22-05
05-06
19-06
03-07
17-07
31-07
14-08
28-08
11-09
25-09
09-10
23-10
06-11
20-11
04-12
Time


Slovenia
Active ICUs
0
50
100
150
200
12-03
26-03
09-04
23-04
07-05
21-05
04-06
18-06
02-07
16-07
30-07
13-08
27-08
10-09
24-09
08-10
22-10
05-11
19-11
Time


Switzerland
Active ICUs
0
100
200
300
400
500
04-03
18-03
01-04
15-04
29-04
13-05
27-05
10-06
24-06
08-07
22-07
05-08
19-08
02-09
16-09
30-09
14-10
28-10
11-11
25-11
Time

Analysis: On a theoretical analysis of massive testing with antigen tests (V). Estimation of the lives saved and final tables.

In the last four reports, we have described the proper framework to classify the different types of massive testing proposals that have been put forward by different governments. We explain that, besides the sensitivity and specificity of the rapid tests, the classification of the different proposals should take into account whether the analysis of its success would need the development of an important infrastructure to do the tests or the sociological responses of the people in front of requests/advice/possibility of taking those tests.

We explained that compulsory tests of a continuous population in a country, city, or neighborhood need the proper infrastructure in place so that it can be carried out. In this type of massive test, personal inclinations are not relevant since all the people take the test and they can move or must remain at home according to the result. We explained that in compulsory continuous frameworks everything is a matter of costs versus benefits of the proposal.

We identified the main costs associated with the massive testing arise from the infrastructure to obtain and deploy the massive testing scheme, plus the vigilance structure needed to make it compulsory, plus the economic loss associated with putting in quarantine a group of people who are false positive in the screening. We also described how the benefit analysis is done using specificity and sensibility values of the test and how these are the main sources of uncertainty regarding the cost-benefit analysis of the measure.

We showed that the total cost of doing test is roughly 10 euros per antigen test, once all the cost associated with performing the tests are included. The estimated error of this cost is roughly 30%. On the other hand, for the benefit side, we found a large uncertainty in its benefits. For a 1% weekly incidence (2% A_{14}) the confidence interval is so large that the cost per truly positive case detected (not including false positive) could go anywhere from 1,000 to 6,000 euros. If we included the error in the cost, it is easy that the cost can go anywhere from 800 euros to 8,000 euros per properly detected person. The former case is associated with specificities around 99.8-99.9% and lower end of cost and higher end of sensibility and the later with specificities around 99.1% and higher end of cost and lower end of sensibility.

Finally, we showed that a simple way to do **cost/benefit estimation analysis** was focusing of the **evaluation of how many people you saved thanks to reduction in risk**. The reason is that the literature on price of risk reduction is quite clear about threshold. If saving a life costs more than one million dollars it is almost sure that you can find other ways to save more lives with the same money. Arguably, in the epidemics we face now, which is new, where this limit is probably lower since it has not been enough time to find all the easy low-cost risk reduction measures. So, anything that costs more than half a million euros per person saved thanks to risk reduction is a clear sign that there are other ways to save much more people with less money. On the other hand, if this price tag is 5,000€, then we find very difficult to entertain the idea that spending this money is not worthy.

Given these two clear extremes where massive testing would clearly work and not, the intermediate price tag needs further clarification. It seems reasonable to expect that anything that it slightly higher but close to 5,000 € would probably be worthy, while something slightly lower than half a million will probably not. The great question is, of course, what to do in intermediate cases.

The complexity of intermediate risk reduction

When the result of our analysis points to price tags between 10,000 and 100,000 euros per person saved thanks to risk reduction, our estimation cannot render useful policy-making advise. At this level of cost, the opportunity cost of the measures becomes key. Are there better ways to save lives? How many resources does a country have? Is this kind of expenditure a priority way to deploy a complex infrastructure plan? Are there other options?

More importantly, **our analysis has only focused on the reduction of deaths thanks to risk reduction, we do not take into account the benefits in reduced hospitals costs or ICU occupied beds.** We also do not take into account the benefits of a possible delay of lockdown measure that could tilt the argument in favor of doing massive testing. Similarly, **we do not take into account the negative economic impact of a hard lockdown needed for the preparation of the massive screening, nor the complexity of the diverting human capital to deploy this plan**. Human capital can be the resource with higher scarcity and, therefore, larger opportunity cost. Employing hours in the design of the process can prevent other key projects to be deployed. We would not be surprise that this could tilt heavily against massive testing.

In any case, it is clear that the scope of this analysis cannot be carried out by our group. It requires a full staff of economic planification to assess whether it is worthy to carry out and feasible. Government agencies should use their knowledge to assess if in intermediate risk reduction, a massive testing process can be carried out

Key tables estimating saved lives thanks to risk reduction

In the following section we describe how a reasonable estimation of the number of people saved can be carried out under certain important assumptions. The goal is to obtain an estimation of how many people are saved per 100,000 people thanks to massive testing intervention as a function of the incidence of the country/city (7-day incidence since it is the that most closely resemble the days the antigens test can detect an infected person) and the fraction of the people that the test cannot detect. This later number is directly related with the sensibility of the test. If the sensibility of the test is 80%, the fraction not detected is 20% (0.2 in our head of the table).

We show below that it is necessary to make assumptions about the R. Different R leads to different number of people saved. We consider simple cases where R is constant (although more complex evolutions of R can indeed be computed). We provide three examples R=0.8, R=1.4 and R=2.0.

In the estimation we do not consider complex interactions between measures, so no complex interaction between lockdown decision or implementations are considered. In this sense, our estimation is more like an upper limit of how many people you save following this procedure since do not account for any interference with any other method. Given that massive testing has been a national event, the interaction with the previous short-lockdown to massive testing is the strongest assumption. Having said that, since we are only interested in estimations, the basic hypothesis is that these interactions would not affect importantly the order of magnitude of people saved.

The tables below show how critical knowing the sensitivity of antigen test. When only 10% of the infected are detected the cost-benefit analysis improve strongly. On the other hand, for 50%, the number of people saved is highly reduced. **Similarly, incidence is key. Low incidence makes massive testing not worthy, and high incidence worthy. Notice also that the close one is to R=1 the more effective risk reduction exists.** Closer to R=1 it is more difficult for the population not detected to grow again to the same level the society had before massive testing. The reduction in risk lasts longer and, therefore, the number of people saved.

Tables: *Estimation of the deaths per 10^5 inhabitants that could be saved because of a mass screening campaign, as a function of the reproduction number (0.8, 1.4 and 2.0), the weekly incidence (20 to 20,000 cases per 100,000 inh.) and the fraction of cases that are non-detected by the tests (10 % to 50 %). Color scale indicate the feasibility according to the cost/benefit analysis. In red, those situations where the cost of saving a life is above half a million (≤2). In green, those situations where the cost of saving a life is lower than 5,000€ (≥200). In red-grey scale, the 500,000-100,000 € range. In green-grey scale, the 5,000-10,000€ range. In grey, the intermediate situations where there is no a clear cost/benefit positioning (10,000-100,000€ range).*

R=0.8		Non detected fraction				
		0.1	0.2	0.3	0.4	0.5
7-day incidence	20	2	1	1	1	1
	50	5	4	3	2	1
	100	9	7	5	4	3
	200	18	14	11	8	6
	500	45	36	27	20	14
	1000	91	72	55	40	28
	2000	181	143	110	81	56

R=1.4		Non detected fraction				
		0.1	0.2	0.3	0.4	0.5
7-day incidence	20	12	5	2	1	1
	50	30	12	6	3	2
	100	60	24	12	7	4
	200	120	48	24	13	7
	500	301	119	61	33	19
	1000	602	238	121	67	37
	2000	1204	476	243	134	74

R=2.0		Non detected fraction				
		0.1	0.2	0.3	0.4	0.5
7-day incidence	20	6	2	1	1	0
	50	15	6	3	2	1
	100	29	12	6	3	2
	200	58	23	12	6	4
	500	146	58	29	16	9
	1000	292	115	59	32	18
	2000	584	231	118	65	36

Methodology to obtain risk reduction

Here we explain the methodology to obtain the previous tables. We have an exponential growth which defines the parameter R after the typical reproduction time of the Covid-19 (T=5days), for an initial value of N_0 cases we obtain the next evolution:

$$N = N_0 e^{\beta t}$$

Therefore, after this reproduction time T we obtain $N(t+T) = R \cdot N(t)$, which can be used to obtain the relation between the exponent β and the reproductive number:

$$R = \frac{N_0 e^{\beta(t+T)}}{N_0 e^{\beta t}} = e^{\beta T}$$

and therefore

$$\beta = \frac{ln(R)}{T}$$

Under uncontrolled conditions we have an exponential growth in the number of new cases where the exponent is related with the reproductive number. When a massive testing is performed, we can assume that a certain percentage of undetected remains. We set the time after the massive test as t=0 and the fraction of undetected as Φ. Therefore, from the time of the massive test we have the new growth given by:

$$N = \phi N_0 e^{\beta t}$$

where ϕN_0 is the actual undetected individuals. We arrive to the same risk situation as t=0 when the evolution of the previous equation arrives to N_0 again:

$$\phi N_0 e^{\beta T_F} = N_0$$

and the time necessary to arrive to such point is:

$$T_F = \frac{-ln\phi}{\beta}$$

People saved with the massive test

We are going to estimate the total of contagious eliminated with the massive test by the comparison of the total infected people during the exponential growth after the massive test from the undetected infected following the exponential growth:

$$N_2 = \phi N_0 e^{\beta t}$$

with respected the estimated infected people expected if the test where not performed

$$N_1 = N_0 e^{\beta t}$$

till the time where the same risk is obtained, see Figure 1.

In order to evaluate the total infected people for both cases from $t = 0$ to $t = T_F$ we basically sum all the cases between these two times, which is equivalent to integrate the exponential curves. The integration of the exponentials provides the next values:

$$A_1 = \int N_0 \, e^{\beta t} = \frac{N_0}{\beta}\left(\frac{1}{\phi} - 1\right)$$

and

$$A_2 = \int \phi \, N_0 e^{\beta t} = \frac{\phi N_0}{\beta}\left(\frac{1}{\phi} - 1\right)$$

The value of A_2 corresponds to the total number of infected people if no measurements were performed, in comparison with the value of A_1 corresponding to the actual infected people form the undetected individuals. The subtraction of both values informs about the **infections avoided due to the massive test**:

$$A = A_2 - A_1 = \frac{N_0}{\phi\beta}(1-\phi)^2$$

which in terms of the reproductive number (R) and the time of reproduction (T) is:

$$A = \frac{N_0 T}{\phi ln(R)}(1-\phi)^2$$

Finally, assuming that the IFR is around 0.5% the **total deaths avoided with the massive test would correspond to**

$$S = 0.005 \cdot \frac{N_0 T}{\phi ln(R)}(1-\phi)^2$$

There are some obvious limits corresponding to

- Φ→ 0 corresponds to the case of completely detected infected. There is potential save of infinite lives due to the theoretical exponential growth
- R → 1 in such cases the reduction in a fraction Φ extremely reduces the number of deaths because the continuous generation of new infected people proportional to N_0.

For R<1 note that β <0 and then the condition to obtain the same risk changes:

$$N_0 e^{\beta t} = \phi N_0$$

and the time necessary to arrive to such point is :

$$T_F = \frac{ln\phi}{\beta}$$

Therefore, the integrals change:

$$A_1 = \int N_0\, e^{\beta t} = \frac{N_0}{\beta}(\phi - 1)$$

and

$$A_2 = \int \phi\, N_0 e^{\beta t} = \frac{\phi N_0}{\beta}(\phi - 1)$$

Now, infections avoided will be:

$$A = A_2 - A_1 = \frac{-N_0}{\beta}(1-\phi)^2$$

Which, in terms of the reproductive number (R) and the time of reproduction (T) is :

$$A = \frac{-N_0 T}{ln(R)}(1-\phi)^2$$

There are some obvious limits corresponding to

- Φ→ 0 corresponds to the case of completely detected infected. However, in this case the exponential decrease of the exponential dynamics implicates a finite number of lives saved.
- R → 1 in such cases the reduction in a fraction phi extremely reduces the number of deaths because the continuous generation of new infected people proportional to N_0.

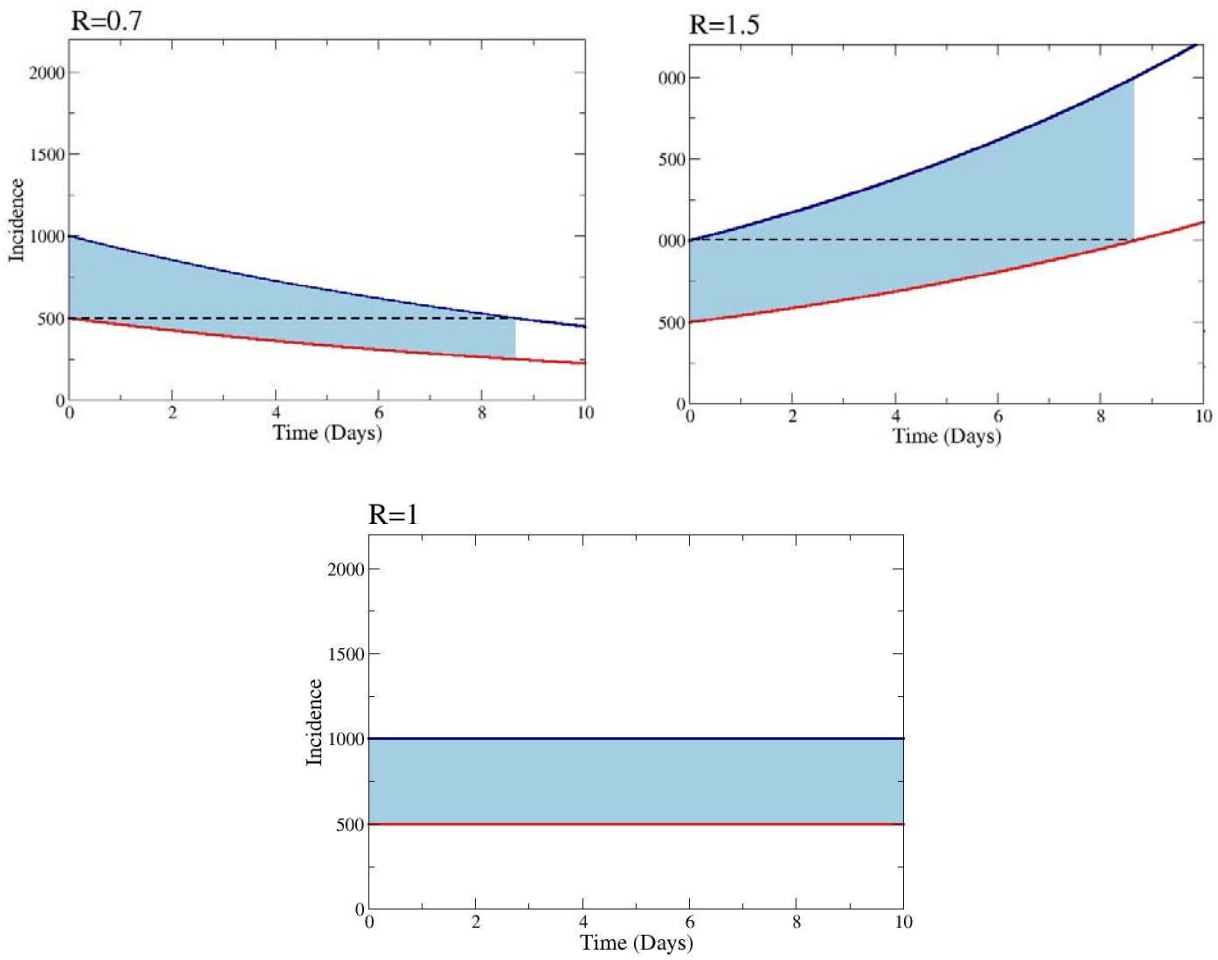


Figure 1. Graphical estimation of the effects of massive testing for three different epidemiological situations. *Blue and red line correspond respectively to the estimated dynamics of the new cases under constant R, without and with massive testing. Light blue area corresponds to the estimated total number of lives saved due to the massive testing. Dashed black line connects values of same epidemiological risk. For the visualization, we assume a 50% of detection (Φ=0.5).*

To summarize, the number if lives saved will be:

$$S = 0.005 \cdot \frac{N_0 T}{\phi ln(R)}(1-\phi)^2 \quad \text{for R>1} \qquad\qquad S = -0.005 \cdot \frac{N_0 T}{ln(R)}(1-\phi)^2 \quad \text{for R<1}$$

Anomalous case where risk analysis cannot be employed since and infinity number of people would be saved. This ideal case never happens in reality for different reasons:

- First, it is never exactly 1, but it can be very close to 1.
- Second, there is a clear cut-off, at some point vaccines or mobility reductions will give a finite bound to this number

In any case, values of R close to 1 are very good in terms of risk reduction since it takes a lot of time to recover the same epidemiological situation you had before. In this case, one should include the cost of not reducing the number of cases and the death count since the epidemics is persistent. Computing the cost of persistent cases in society is outside of our aim in this assessment, so we do not consider cases very close to R=1.

(4) Fitting a mathematical model to data

Previous studies have shown that Gompertz model[2] correctly describes the Covid-19 epidemic in all analysed countries. It is an empirical model that starts with an exponential growth but that gradually decreases its specific growth rate. Therefore, it is adequate for describing an epidemic wave that is characterized by an initial exponential growth but a progressive decrease in spreading velocity provided that appropriate control measures are applied. Once in the tail, predictions work but the meaning of parameters is lost.

Gompertz model is described by the equation:

$$N(t) = K\, e^{-ln\left(\frac{K}{N_0}\right)\cdot e^{-\alpha\cdot(t-t_0)}}$$

where *N(t)* is the cumulated number of confirmed cases at *t* (in days), and N_0 is the number of cumulated cases the day at day t_0. The model has two parameters:

- ✓ *a* is the velocity at which specific spreading rate is slowing down;
- ✓ *K* is the expected final number of cumulated cases at the end of the epidemic.

This model is fitted to reported cumulative cases of the UE and of countries that accomplish two criteria: 4 or more consecutive days with more than 100 cumulated cases, and at least one datapoint over 200 cases. Day t_0 is chosen as that one at which *N(t)* overpasses 100 cases. If more than 15 datapoints that accomplish the stated criteria are available, only the last 15 points are used. The fitting is done using Matlab's Curve Fitting package with Nonlinear Least Squares method, which also provides confidence intervals of fitted parameters (*a* and *K*) and the R^2 of the fitting. At the initial stages the dynamics is exponential and *K* cannot be correctly evaluated. In fact, at this stage the most relevant parameter is *a*.

It is worth to mention that the simplicity of this model and the lack of previous assumptions about the Covid-19 behaviour make it appropriate for universal use, i.e., it can be fitted to any country independently of its socioeconomic context and control strategy. Then, the model is capable of quantifying the observed dynamics in an objective and standard manner and predicting short-term tendencies.

(5) Using the model for predicting short-term tendencies

The model is finally used for a short-term prediction of the evolution of the cumulated number of cases (3-5 days). The confidence interval of predictions is assessed with the Matlab function `predint`, with a 99% confidence level. These predictions are shown in the plots as red dots with corresponding error bar. For series longer than 9 timepoints, last 3 points are weighted in the fitting so that changes in tendencies are well captured by the model.

(6) Estimating non-diagnosed cases

Lethality of Covid-19 has been estimated at around 1 % for Republic of Korea and the Diamond Princess cruise. Besides, median duration of viral shedding after Covid-19 onset has been estimated at 18.5 days for non-survivors[3] in a retrospective study in Wuhan. These data allow for an estimation of total number of cases, considering that the number of deaths at certain moment should be about 1 % of total cases 18.5 days before. This is valid for estimating cases of countries at stage II, since in stage I the deaths would be mostly due to

[2] Madden LV. Quantification of disease progression. Protection Ecology 1980; **2**: 159-176.
[3] Zhou et al., 2020. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet; March 9, doi: 10.1016/S0140-6736(20)30566-3

the incidence at the country from which they were imported. We establish a threshold of 50 reported cases before starting this estimation.

Reported deaths are passed through a moving average filter of 5 points in order to smooth tendencies. Then, the corresponding number of cases is found assuming the 1 % lethality. Finally, these cases are distributed between 18 and 19 days before each one.