



HAL
open science

Estimation of covid-19 cases in France and in different countries: Homogeneisation based on mortality

Marc Dhenain

► **To cite this version:**

Marc Dhenain. Estimation of covid-19 cases in France and in different countries: Homogeneisation based on mortality. 2020. hal-02570346

HAL Id: hal-02570346

<https://hal.science/hal-02570346>

Preprint submitted on 12 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **ESTIMATION OF COVID-19 CASES IN FRANCE AND IN DIFFERENT COUNTRIES:**
2 **HOMOGENEISATION BASED ON MORTALITY**

3 *Marc DHENAIN*

4
5 (1) Académie Vétérinaire de France, 34, rue Bréguet, 75011 Paris, France

6 (2) Académie Nationale de Médecine, 16 rue Bonaparte, 75006 Paris, France

7 (3) Centre National de la Recherche Scientifique (CNRS), Université Paris-Sud, Université Paris-Saclay UMR
8 9199, Laboratoire des Maladies Neurodégénératives, 18 Route du Panorama, F-92265 Fontenay-aux-Roses,
9 France.

10 (4) Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), Institut François Jacob, Molecular
11 Imaging Research Center (MIRcen), 18 Route du Panorama, F-92265 Fontenay-aux-Roses, France.

12
13 **Correspondance**

14 Marc Dhenain

15 MIRcen, UMR CEA-CNRS 9199, 18 Route du Panorama, 92 265 Fontenay-aux-Roses CEDEX, France

16 Tel: +33 1 46 54 81 92; Fax: +33 1 46 54 84 51; email: Marc.Dhenain@cea.fr

17
18 **Abstract:**

19 *Every day the authorities of different countries provide an estimate of the number of persons*
20 *affected by Covid-19 and a count of fatality. We propose to use the fatality reported in each*
21 *country to provide a better estimate ($C_{t_0\text{-estimated}}$) of the number of cases at a given time t_0 .*

22
$$C_{t_0\text{-estimated}} = (F_{t_0} / F_{r\text{-est}}) * (F_{t_0} / F_{t_0-3d})^6$$

23 *With F_{t_0} : number of fatalities reported in a country at time t_0 ; F_{t_0} : number of fatalities reported*
24 *in a country at time t_0 minus 3 days; $F_{r\text{-est}}$: estimated fatality rate. Based on $C_{t_0\text{-estimated}}$ calculated*
25 *using a fatality rate of 2%, we assessed the number of cases April 10th, 2020 in Belgium, China,*
26 *France, Germany, Iran, Italy, South Korea, Netherlands, Spain, United Kingdom and USA.*
27 *This number reached 2,872,097 in France and 924,892 persons in Germany. This work*
28 *suggests a very strong underestimation of the number of cases of people affected, with a*
29 *notification index often lower than 5%. The proposed formula also makes it possible to evaluate*
30 *the impact of policies to prevent the spread of epidemic.*

31
32 **Key-Words: Covid-19, Estimated number of cases, Fatality, Prevalence**

33
34 *Version submitted on May 10th 2020*

35 A French first version of this article is "in press" as

36 **Dhenain Marc, Estimation du nombre de cas de Covid-19 en France et dans différents pays :**
37 **homogénéisation basée sur la mortalité, Bulletin de l'Académie Vétérinaire de France, 2020 (provisionally**
38 **accepted), <https://academie-veterinaire-defrance.org/bavf-coronavirus/>**

1 INTRODUCTION

2 The Sars-CoV-2 coronavirus infection that causes Covid-19 has spread worldwide leading to
3 significant deaths (European Center for Disease Prevention and Control 2020). Every day
4 authorities of different countries provide an estimate of the number of affected persons and a
5 count of fatalities (Dong et al. 2020, <https://ourworldindata.org/covid-testing>,
6 https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series).

7 Knowing the number of affected subjects is critical for implementing strategies to protect
8 populations and for ending the crisis. Figures reported by different countries reveal strong
9 differences and only partly reflects the reality (Table I). For example, the day their death toll
10 approached 3,000 people, France had 44,550 people affected versus 80,537 for China and
11 122,171 for Germany. Calculating the case fatality rate (F_r) on a given day (t_0) is another way
12 to objectify differences between countries. At first sight,

$$13 \quad F_r = F_{t_0} / C_{t_0} \quad (\text{Eq. 1})$$

14 With F_{t_0} = number of fatalities reported on day t_0 ; C_{t_0} = number of cases reported on day t_0 .
15 The day (the closest to April 10th, date of redaction of this article) when the death toll of
16 different countries was the closest to 3,000 people, three countries (Germany, South Korea,
17 and the United States) had fatality rates close to 2%; seven countries (Belgium, France, Iran,
18 Italy, the Netherlands, Spain, and the United Kingdom) had rates between 6% and 12%, and
19 China had an intermediate value of 3.7% (Table I).

20 Patients who die on any given day were infected much earlier, and thus the denominator of
21 the fatality rate should be the total number of patients infected at the same time as those who
22 died (Baud et al. 2020). This is particularly true as the rates of evolution of the pandemic
23 evolve differently in various countries: in March 2020, the number of people affected
24 increased sharply from day to day in France, while it was stable in China.

25 A better estimate of fatality rate is thus:

$$26 \quad F_{r-x\text{day}} = F_{t_0} / C_{t_0-x\text{days}} \quad (\text{Eq. 2})$$

27 With $C_{t_0-x\text{days}}$ = number of cases reported on day t_0 minus x days, with x = average time-
28 period from onset of symptoms to death.

29 An average duration of 18 days is reported between the onset of symptoms and the death of
30 Covid-19 patients (Ruan et al. 2020; Verity et al. 2020; Zhou et al. 2020). Thus, the adjusted
31 Fatality rate (F_{r-18d}) that takes into account this average delay is (Flaxman et al. 2020).

$$32 \quad F_{r-18d} = F_{t_0} / C_{t_0-18d} \quad (\text{Eq. 3, Table I})$$

1 With C_{t_0-18d} = number of cases reported on day t_0 minus 18 days. The calculation of F_{r-18d}
2 reveals widening gaps between countries compared to F_r with variations ranging from 2.3%
3 (South Korea) to more than 700% for Spain.
4 When comparing F_{t_0} and C_{t_0-18d} in different countries (Fig. 1), we see a linear relationship
5 between mortality at t_0 and the number of cases at $t_0-18days$ for all countries (Pearson linear
6 correlation test, $p < 0.05$ except for Belgium ($p = 0.07$) due to the small number of points ($n = 3$)).
7 The slopes of the regression lines fitting the data vary widely between countries, which is
8 consistent with variable F_{r-18} . The values of fatality rate F_{r-18} based on the cases reported by
9 the different countries are therefore unreliable, in part because the number of cases reported in
10 different countries is not reliable (different testing strategies in different countries).
11 How to assess more precisely the number of people affected using a similar method for
12 different countries? We offer a simple method using the number of deaths reported by each
13 country to estimate and compare the rate of people affected by Covid-19. This method relies
14 on three first assumptions: 1. The number of deaths reported by each country is reliable; 2.
15 The fatality rate (F_r) is known and similar in different countries; 3. The average time between
16 the onset of symptoms and death is known (here considered 18 days). Based on these
17 assumptions, one can calculate the number of cases presented eighteen days before a given
18 day (t_0). Two methods are then proposed to infer the number of cases, eighteen days later, at
19 time t_0 . The first one relies on the time-dependent increase in the number of cases reported in
20 databases during these 18 days. The second one models the evolution of the number of cases
21 based on daily rate of changes of the number of estimated cases 18 days before t_0 .

22

23 **METHODS**

24 **Estimation of the number of cases at time t_0 minus 18 days based on fatality rate**

25 One way to estimate the number of Covid-19 cases is to infer the number of cases based on
26 the number of death and the fatality rate calculated from well-controlled studies using the
27 following formula

$$28 \quad C_{(est-18d)} = F_{t_0} / F_{r-est} \quad (\text{Eq. 4, Table II})$$

29 With $C_{(est-18d)}$: number of cases estimated 18 days before t_0 ; F_{t_0} = number of deaths reported
30 on day t_0 ; F_{r-est} = estimated fatality rate from well-controlled studies. F_{r-est} can be assessed
31 from well-controlled studies based on residents of mainland China, travelers returning from
32 mainland China, repatriated from China, passengers on the *Diamond-Princess* cruise ship
33 (values of 0.7 to 3.6% (Verity et al. 2020)). Here, based on this last study, we proposed to use
34 $F_{r-est} = 2\%$.

1 Knowing $C_{(est-18d)}$ eighteen days before t_0 , one needs to assess the progression rate of the
 2 cases during the 18 last days to estimate the number of cases at day t_0 ($C_{t_0-estimated}$). We tested
 3 two methods to assess this progression.

4 **Estimation of progression of cases for 18 days based on reported number of cases**

5 One can assume that progression of estimated cases (P_{18d}) reflects the time-dependent
 6 increase in the reported number of cases during the same time-period. In that case,

7
$$P_{18d} = C_{t_0} / C_{t-18d} \quad (\text{Eq. 5})$$

8 With C_{t_0} : number of cases reported in a country at time t_0 ; C_{t-18d} : number of cases reported in
 9 the country at time t_0 minus 18 days. Thus,

10
$$C_{t_0-estimated} = C_{(est-18d)} * P_{18d} \quad (\text{Eq. 6, Table II})$$

11 **Estimation of progression of cases for 18 days based on 3-day rate of change of the**
 12 **estimated $C_{(est-18d)}$**

13 Another option to assess the progression of cases for 18 days is to calculate the daily rate of
 14 change of the number of estimated cases (R_d) or alternatively 3-day rate of change (R_{3d}) of the
 15 number of estimated cases

16
$$R_{3d} = (C_t / C_{t-3d}) - 1 \quad (\text{i.e. } (C_t - C_{t-3}) / C_{t-3d}) \quad (\text{Eq. 7})$$

17 With C_{t_0} : number of cases reported at time t_0 ; C_{t_0-3d} : number of cases reported at time t_0 minus
 18 3 days.

19 The last day when this calculation is possible is 18 days before t_0 .

20
$$R_{3d-18d} = (C_{(est-18d)} / C_{(est-21d)}) - 1 \quad (\text{i.e. } (C_{(est-18d)} - C_{(est-21d)}) / C_{(est-21d)}) \quad (\text{Eq. 8})$$

21 With $C_{(est-18d)}$: estimated cases 18 days before t_0 (see Eq. 4); $C_{(est-21d)}$: estimated cases three
 22 days before. Assuming that the progression of the estimated cases follows an exponential
 23 model then

24
$$C_{t_0-estimated} = C_{(est-18d)} * (1 + R_{3d-18d})^6 = C_{(est-18d)} * (F_{t_0} / F_{t_0-3d})^6 \quad (\text{Eq. 9})$$

25 With F_{t_0-3d} : number of fatalities reported in a country at time t_0 minus 3 days. The exponent 6
 26 represents the period of the model as $6 * 3 \text{ days} = 18 \text{ days}$. This estimation supposes that R_{3d}
 27 does not evolve with time, during the last 18 days.

28
$$\text{Thus } C_{t_0-estimated} = (F_{t_0} / F_{t-est}) * (F_{t_0} / F_{t_0-3d})^6 \quad (\text{Eq. 10})$$

29

30 **Comparison with basic reproductive potential of the pathogen (R_0)**

31 The basic reproductive potential of the pathogen (R_0) is an important index in epidemiology.
 32 It is defined as the average number of secondary cases arising from a primary case in an
 33 entirely susceptible population. Another critical parameter is the mean generation time (T_g)

1 *i.e.*, the time between the infection of a primary case and the infection of a secondary case
2 (Flaxman et al. 2020; Keeling & Rohani, 2008).

3 Assuming that the progression of Covid-19 follows a SEIR model based on four
4 compartments (Susceptible, Exposed, Infectious, and Recovered subjects (Keeling & Rohani,
5 2008)), the increase in prevalence during the invasion phase of the disease is estimated as

$$6 \quad C_{SEIR(t)} \approx C_{(t_0)} \exp \{[(\sqrt{R_0} - 1) * \gamma] * t\} \quad (\text{Eq. 11, Keeling \& Rohani, 2008})$$

7 With C: number of infected cases at a time t or t₀; R₀: the basic reproductive potential of the
8 pathogen; γ : the recovery rate γ with $1/\gamma$: the infectious period (Di).

9 This equation allows to estimate the doubling time for cases (T₂) if $C_{SEIR(t)}/C_{(t_0)}=2$.

$$10 \quad T_2 = \ln(2) / [(\sqrt{R_0} - 1) * \gamma] \quad (\text{Eq. 12})$$

$$11 \quad \text{thus } R_0 = [[\ln(2) / (T_2 * \gamma)] + 1]^2 \quad (\text{Eq. 13})$$

$$12 \quad T_g \approx D_e + D_i/2 \quad (\text{Eq. 14})$$

13 if Covid-19 is modeled with a SEIR model (with D_e = exposition time during which a subject
14 is exposed but not infectious), and $D_e \approx D_i$ (Li et al., 2020).

15 One can thus estimate that

$$16 \quad R_0 = [[2 * T_g * \ln(2) / (3 * T_2)] + 1]^2 \quad (\text{Eq. 15})$$

17 Using the daily rate of change (R_d) (or R_{3d} / 3), and the "rule of 70", one can estimate

$$18 \quad T_2 = 0.7 / R_d \quad (\text{Eq. 16})$$

$$19 \quad \text{Thus } R_0 = [[2 * T_g * \ln(2) / (3 * 3 * 0.7 / R_{3d})] + 1]^2 \quad (\text{Eq. 17})$$

20 T_g can be estimated to be 6.5 according to (Flaxman et al. 2020).

21

22 **RESULTS**

23 **Estimations relying on number of cases at time t₀ minus 18 days**

24 Using the (Eq. 4) and international databases (Dong et al. 2020,

25 https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series),

26 we estimated the number of people that have been affected by Covid-19 eighteen days before
27 April 10st 2020 in different countries (Table II). This estimation was 659,850 persons in

28 France and 138,350 in Germany. We then proposed two different methods to infer the number
29 of cases 18 days latter (Fig. 2). Method based on P_{18d} evaluation using the reported number of
30 cases during the same time-period suggested a time-dependent decrease of number of cases in
31 some countries (*e.g.* Germany (Fig. 2C) or USA (Fig. 2D)), which is not consistent. Method

32 based on the evaluation of R_{3d} provided a better correspondence with the estimated cases in

33 all tested countries (Fig. 2). We thus retained results from the R_{3d} method for further analyses.

1 Estimation of the number of cases based, April 10th was 2,872,097 in France, 924,892 in
2 Germany, 1,811,469 in Spain, 4,240,198 in the United Kingdom and 9,035,229 in the United
3 States (Table II, Fig. 2, Fig. 3A). R_{3d} evaluation can further be used to assess the basic
4 reproductive potential of the pathogen (R_0) based on (Eq. 17). We reported R_0 values from 1.0
5 (China) to 2.86 (Belgium) (Table II).

6 This analysis is based on an estimated fatality rate from well-controlled studies (M_{r-est}) of 2%.
7 The estimated number of cases must be halved if the mortality rate used jumps from 2 to 4%
8 (Table II). It must be doubled if the fatality rate used drops from 2 to 1%. Some authors
9 suggest that the real fatality rate for Covid-19 could be 5.6 to 15.6% (Baud et al. 2020). If the
10 calculation uses a fatality rate of 15%, then the estimated number of cases drops to 382,946
11 for France, but it becomes lower than the number of cases actually reported for some
12 countries (e.g. 2,242 *versus* 10,450 for South Korea), which is not consistent (Table II).

13 In our study, we set the delay between the onset of symptoms and death at 18 days based on
14 robust data from the literature (Ruan et al. 2020; Verity et al. 2020; Zhou et al. 2020) and
15 delays used in other models (Flaxman et al. 2020). Lowering this delay, for example to 12
16 days, sharply decreases the number of estimated cases (e.g. 1,645,302 for France (Table II))
17 although it remains high compared to figures reported by most countries.

18 Using estimations based on R_{3d} model, with delay of 18 days between the onset of symptoms
19 and death and fatality rate of 2%, we could thus compare estimated cases of Covid-19 in
20 different countries (Fig. 3A), proportion of cases in different countries (Fig. 3B), as well as
21 notification indexes which is the ability to report cases (Fig. 3C-D). These data highlight
22 strong discrepancies between countries. It suggests a high proportion of affected persons in
23 Belgium. It also shows notification indexes that varies from 60 to 80% in Korea while it is
24 below 5% in most countries.

25 **Comparison of estimated cases with the number of cases reported afterwards**

26 The number of cases evaluated between March 16 and April 10, 2020 from the number of
27 cases at time t_0 minus 18 days and the R_{3d} model was compared with cases estimated from the
28 number of cases at time t_0 minus 18 days (without the R_{3d} -based model) calculated from
29 mortality data collected between April 11 and 28, 2020 (Fig. 4). The cases estimated with the
30 R_{3d} model were higher than those estimated a posteriori. This can easily be explained by the
31 reduction in the spread of the disease (and thus of R_{3d}) in the past 18 days following the
32 containment measures in many countries. The three-day rate of change in the number of
33 estimated cases (R_{3d}) we used is overstated.

1 The measurements that we have made allow to assess the impact of containment policies. For
 2 example, for France, the cases estimated on April 10 were based on an R_{3d} of 0.28 (measured
 3 on March 23). Five days before this date (March 19), R_{3d} was 0.50. Using an R_{3d} of 0.50
 4 (instead of 0.28) leads to an estimate of 7,516,104 cases on April 10. The reduction of R_{3d}
 5 from 0.50 to 0.28 thanks to containment therefore prevented the appearance of 4,644,007 new
 6 cases in France (7,514,104 - 2,872,097). The number of cases estimated from the deaths
 7 which occurred on April 28, 2020 was in fact 1,141,450. Thus, the number of cases actually
 8 avoided is $7,514,104 - 1,181,450 = 6,332,654$. This represents 126 653 avoided deaths .

9

10 **DISCUSSION**

11 It is essential to assess the number of persons affected by Covid-19 in all countries of the world
 12 to stem this crisis. We propose to use the mortality reported by each country at a time t_0 to create
 13 an index of the number of real cases at this same time t_0 . Mortality at t_0 makes it possible to
 14 estimate the number of cases 18 days earlier ($C_{(est-18d)}$). Then, the rate of change over time of
 15 the estimated cases ($C_{(est)}$) is evaluated over 3 days (R_{3d}). This rate is used to modulate $C_{(est-18d)}$,
 16 and calculate the number of cases at time t_0 . This calculation leads to the following equation

$$17 \quad C_{t_0-estimated} = (F_{t_0} / F_{r-est}) * (F_{t_0} / F_{t_0-3d})^6 \quad (Eq. 10)$$

18 With F_{t_0} : number of fatalities reported in a country at time t_0 ; F_{t_0} : number of fatalities reported
 19 in a country at time t_0 minus 3 days; F_{r-est} : estimated fatality rate. This analysis is based on
 20 four assumptions: 1. The number of deaths reported by each country is reliable, 2. The
 21 estimated fatality rate among people affected is known (F_{r-est} , here considered as 2%), 3. The
 22 average time between the onset of symptoms and death is known (here considered 18 days). 4.
 23 The rate of variation over three days of the estimated cases (R_{3d}) does not change during the
 24 last 18 days. This last condition is not entirely exact because, thanks to containment policies,
 25 the rate of change over three days decreased continuously until the beginning of May 2020.

26 $C_{t_0-estimated}$ is thus overestimated as shown by the comparison of the values obtained with
 27 measurements a posteriori, that is to say without taking into account the modeling of the last
 28 18 days. Estimating the 3-day rate of change in the number of estimated cases ($C_{(est-18d)}$ -
 29 Eq.4) can be used to assess R_0 . The values of R_0 that we have reported (from 1.0 (China) to
 30 2.86 (Belgium)) are consistent with data from the literature (for example $R_0 = 4.0$ in Flaxman
 31 et al. 2020, at the beginning of pandemic). Conversely, the estimation of R_0 on the basis of an
 32 epidemiological model (cf. for example (Flaxman et al. 2020)) could be used to calculate R_{3d}
 33 (cf. Eq. 17) and refine the estimation of the number of cases. It would also be possible to

1 smooth out the risks of daily variations in the calculation of R_{3d} by using average measures of
2 R_{3d} over a longer period of time.

3 Our analyses showed that the number of Covid-19 cases in several country greatly exceeds
4 the number of cases presented in international databases (2,872,097 (or 1,181,450 cases with a
5 posteriori measures) versus 124,869 for France on April 10th, 2020). The very high values of
6 estimated cases that we report are consistent with those evaluated with another method by
7 (Flaxman et al. 2020). For example, we report 1.8 million cases in Spain while Flaxman
8 reports 7.0 million on March 28th. Our calculation relies on a relatively simple method while
9 that of Flaxman uses more complex analyzes (hierarchical semi-mechanistic Bayesian model).

10 Our model used a fatality rate of 2% while several strongly controlled international studies
11 reported rates of 0.7 to 3.6% (Verity et al. 2020). Values from 0.5 to 4% could thus be other
12 reasonable options to estimate fatality rate. One of the limitations of our model is that fatality
13 rates can change from one country to another, for example depending on the distribution of
14 the population of different age groups that have different susceptibility to Covid-19. Also, it is
15 possible that death rate changes over time in a given country, for example because of the
16 saturation of hospitals or the correction of mortality figures to include non-counted cases (as
17 done in France between April 1 and 4, 2020 to include mortality in nursing homes). We fixed
18 a single value for the time between symptom occurrence and death (18 days). In reality, this
19 time is variable with a 95% credible interval of 16.9 to 19.2 or more (Verity et al. 2020). We
20 however considered that using such interval would make the model more complicated without
21 strongly adding reliability compared to other potential sources of errors. Our analysis is based
22 solely on the number of people who have died with confirmed Covid-19 cases. It is therefore
23 essential that all countries are able to provide very reliable Covid-19 death values. Finally,
24 note that to know the number of actual cases in a country at a given time, we must subtract
25 from the estimates presented here the number of people healed, including those whose disease
26 has not been identified.

27 To conclude, our model questions the small number of people reported to be affected by
28 Covid-19 in most countries compared to the large numbers we estimate. This difference could
29 be explained by a large underestimation of the "mortality rate". For example, in France, the
30 "estimated mortality rate" should be changed from 2 to ~ 46% to decrease the estimate of the
31 number of cases from 2,872,097 to 124,869. Obviously, a mortality rate of 46% is not
32 observed. Thus, the only explanation that remains is that most countries strongly
33 underestimate the number of affected people. The secondary interest of our model is that it

1 takes into account the rate of change over 3 days, over the past 18 days. Thus it can be used to
2 model the effectiveness of policies to prevent the spread of Covid- 19.

3

4 **ACKNOWLEDGMENTS**

5 We thank Matthieu Domenech de Cellès for constructive and supportive advices during the
6 redaction of this manuscript.

7

8 **REFERENCES**

9 Baud D, Qi X, Nielsen-Saines K, Musso D, Pomar L, Favre G. Real estimates of mortality
10 following COVID-19 infection [published online ahead of print, 2020 Feb 19]. Lancet Infect
11 Dis. 2020. doi: 10.1016/S1473-3099(20)30195-X.

12

13 Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real
14 time [published online ahead of print, 2020 Feb 19]. Lancet Infect Dis. 2020;S1473-
15 3099(20)30120-1. doi:10.1016/S1473-3099(20)30120-1. Database available on:

16 https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series

17 Consulted 2020/04/20

18

19 Flaxman S, Mishra S, Gandy A, Unwin JT, Coupland H, Mellan TA et al. 2020. Estimating
20 the number of infections and the impact of nonpharmaceutical interventions on COVID-19 in
21 11 European countries. Imperial College London (30-03-2020). doi: 10.25561/77731

22

23 European Centre for Disease Prevention and Control. Situation Update-Worldwide.

24 Disponible à: <https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases>.

25 Consulted 2020/04/20

26

27 Keeling, M.J. & Rohani, P. 2008. Modeling infectious diseases in humans and animals.

28 Chapter 2- Introduction to simple epidemic models. Princeton University Press.

29

30 Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., Shaman, J. 2020. Substantial
31 undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-
32 CoV2). Science. doi: 10.1126/science.abb3221

33

- 1 Ruan Q, Yang K, Wang W, Jiang L, Song J. Clinical predictors of mortality due to COVID-
2 19 based on an analysis of data of 150 patients from Wuhan, China [published online ahead of
3 print, 2020 Mar 3]. *Intensive Care Med.* 2020;1–3. doi:10.1007/s00134-020-05991-x
4
- 5 Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N et al. Estimates of the
6 severity of coronavirus disease 2019: a model-based analysis [published online ahead of print,
7 2020 Mar 30]. *Lancet Infect Dis.* 2020;S1473-3099(20)30243-7. doi:10.1016/S1473-
8 3099(20)30243-7
9
- 10 Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z et al. Clinical course and risk factors for mortality
11 of adult in patients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet.*
12 2020 ; 395: 1054-1062. doi: 10.1016/S0140-6736(20)30566-3

1 **FIGURES AND TABLES**

2

Country	Day (t_0)	Fatalities at t_0 (F_{t_0})	Cases at t_0 (C_{t_0})	Fatality rate F_r	Cases at t_0-18j (C_{t_0-18d})	Fatality rate F_{r-18d}
Belgium	April 10	3 019	26 667	11.3%	3743	80.7%
China	March 5	3 015	80 537	3.7%	70 513	4.3%
France	March 30	3 024	44 550	6.8%	2 281	132.6%
Germany	April 10	2 767	122 171	2.3%	29 056	9.5%
Iran	April 1	3 036	47 593	6.4%	12 729	23.9%
Italy	March 18	2 978	35 713	8.3%	1 128	264.0%
South-Korea	April 10	208	10 450	2.0%	8 961	2.3%
Netherlands	April 10	2 511	23 097	10.9%	4 749	52.9%
Spain	March 24	2 808	39 885	7.0%	400	702.0%
United Kingdom	April 2	2 921	33 718	8.6%	1 140	256.2%
USA	March 30	2 978	161 807	1.8%	1 163	256.1%

3

4 *Table I: Fatality rates in different countries when the number of deaths approached 3,000*

5 *people (or the last figure available when the 3,000 deaths were not reached April 10th 2020).*

6 (https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series)

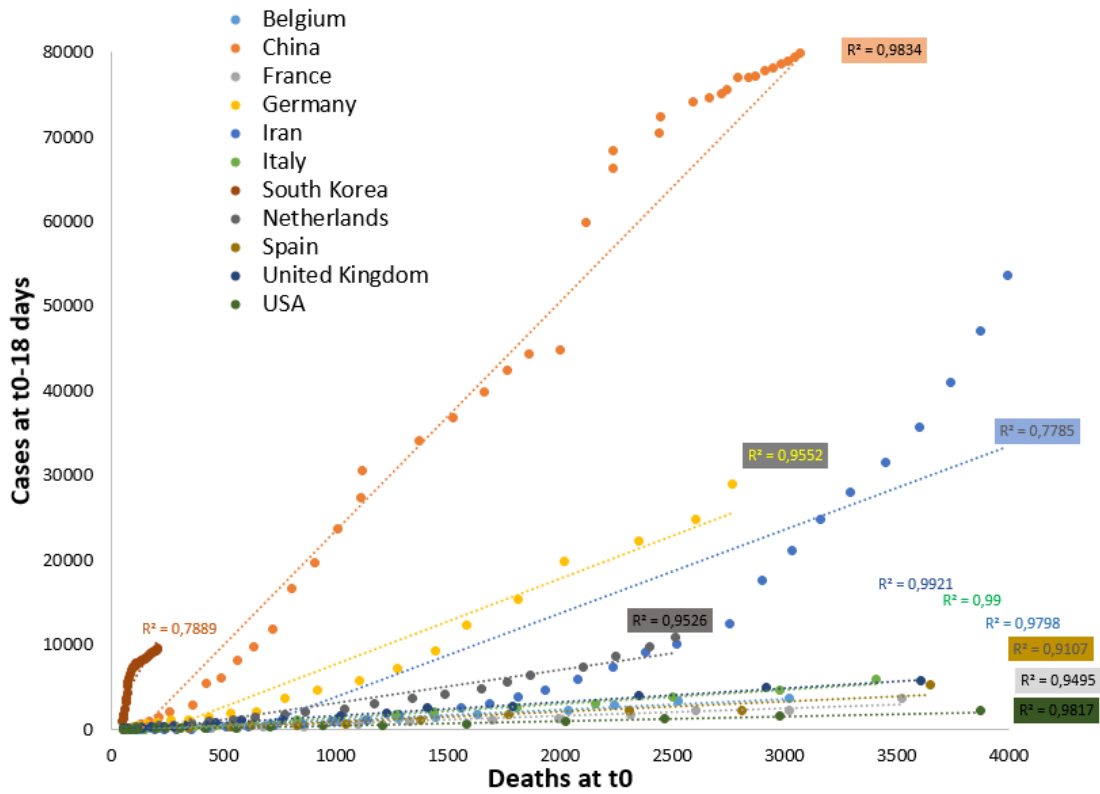
7

1

Country	Population (million)	Reported cases at t_0 (April 10) (C_{t_0})	Reported deaths at t_0 (April 10) (F_{t_0})	Reported cases at t_{0-18d} (March 23) ($C_{t_{0-18d}}$)	Method 1&2 Estimated cases at t_{0-18d} (March 23) $C_{(est-18d)}$	Method 1 Progression rate from actual cases		Method 2 Estimation from 3-days rate of change					
						Progression rate from t_{-18} to t_0 $P_{18d} = C_{t_0} / C_{t_{0-18d}}$ (April 10) P_{18d}	Estimated Cases (April 10) $C_{t_0-estimated}$	R_{3d-18j} (April 10) $C_{t_0-estimated}$	R_0 (April 10) $C_{t_0-estimated}$	Estimated cases (April 10) $C_{t_0-estimated}$	Estimated cases (April 10) $C_{t_0-estimated}$	Estimated cases (April 10) $C_{t_0-estimated}$	Estimated cases (April 10) $C_{t_0-estimated}$
Delay							t_{0-18d}	t_{0-18d}	t_{0-18d}	t_{0-18d}	t_{0-18d}	t_{0-18d}	t_{0-12d}
M_{r-est}							2%	2%	2%	2%	4%	15%	2%
Belgium	11,476	26 667	3019	3 743	150 950	7,12	1 075 443	0,48	2,86	1 609 257	804 628	214 568	683 894
China	1,384,688	82 941	3340	81 498	167 000	1,02	169 957	0,00	1,00	168 508	84 254	22 468	168 003
France	67,795	124 869	13197	19 856	659 850	6,29	4 149 618	0,28	1,95	2 872 097	1 436 049	382 946	1 645 302
Germany	83,073	122 171	2767	29 056	138 350	4,20	581 717	0,37	2,35	924 892	462 446	123 319	482 669
Iran	82,022	68 192	4232	23 049	211 600	2,96	626 033	0,09	1,28	360 726	180 363	48 097	294 241
Italy	60,360	147 577	18849	63 927	942 450	2,31	2 175 668	0,10	1,31	1 674 559	837 279	223 275	1 364 771
South-Korea	51,709	10 450	208	8 961	10 400	1,17	12 128	0,08	1,25	16 811	8 406	2 242	14 306
Netherlands	17,282	23 097	2511	4 749	125 550	4,86	610 619	0,20	1,64	365 882	182 941	48 784	240 570
Spain	46,935	158 273	16081	35 136	804 050	4,50	3 621 909	0,14	1,46	1 811 469	905 735	241 529	1 345 788
United Kingdom	65,761	73 758	8958	6 650	447 900	11,09	4 967 851	0,45	2,72	4 240 198	2 120 099	565 360	2 041 524
USA	328,240	496 535	18586	43 847	929 300	11,32	10 523 638	0,46	2,75	9 035 229	4 517 615	1 204 697	4 086 903

2

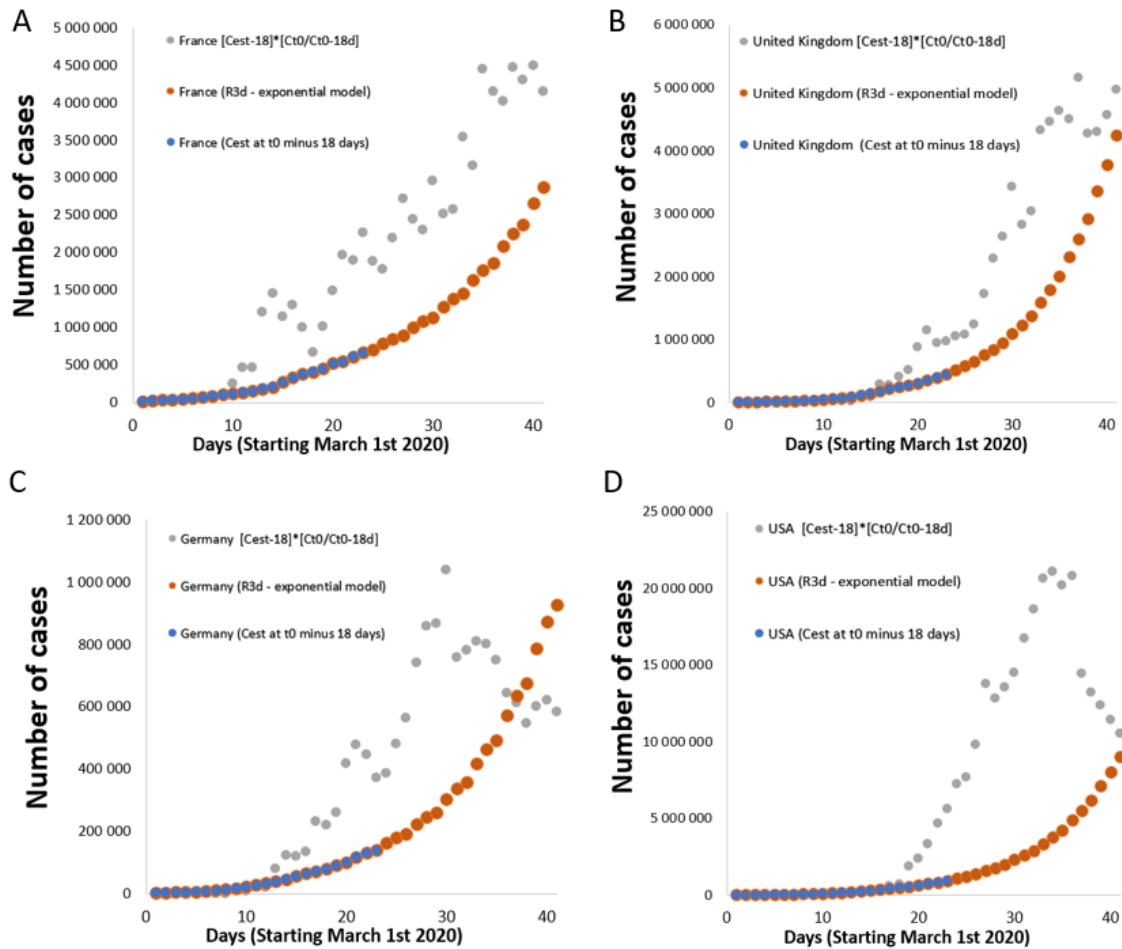
3 Table II: Estimation of the number of cases in different countries April 10st (t_0) using different methods and an estimated Fatality rate (F_{r-est}) of 2%. Numbers
4 of cases estimated with different methods are provided using delays of 18 or 12 days between symptom occurrence and death.



1

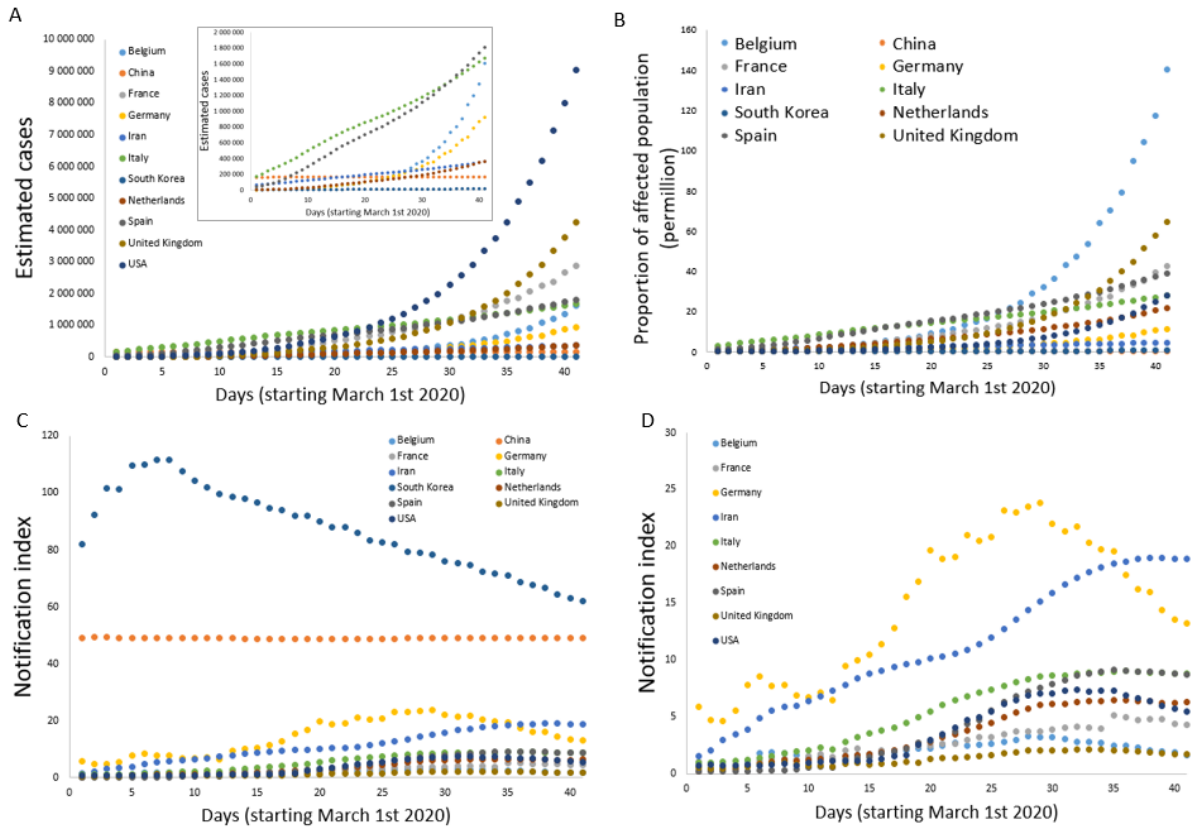
2 *Figure 1: Relationships between deaths a given day (t_0) and the number of cases eighteen*
 3 *days before ($t_0-18days$) in different countries. The figure includes only values between 50 and 4*
 4 *000 deaths (or less if the number of deaths was lower in the country April 10th).*

5



1
2 *Figure 2: Comparison of evolution of estimated Covid-19 cases in France (A), United Kingdom*
3 *(B), Germany (C), and USA (D) from March 1st 2020 to April 10th 2020. The estimated values*
4 *corresponding to the number of cases estimated 18 days before t_0 ($C_{est-18d}$), based on the number*
5 *of deaths at t_0 are displayed in blue. The progression of these cases, taking into account a*
6 *multiplying factor (P_{18d}) which reflects the increase over time in the number of cases declared*
7 *during the same period is noted in gray. This calculation method leads to a large number of*
8 *cases and to daily variations. The orange marks correspond to a model based on R3d ($(1 + C_{est-}$*
9 *$3d) * R_{3d}$). It provides curves which follow the values corresponding to C_{est} based on fatalities*
10 *estimated 18 days before a given day (blue marks).*

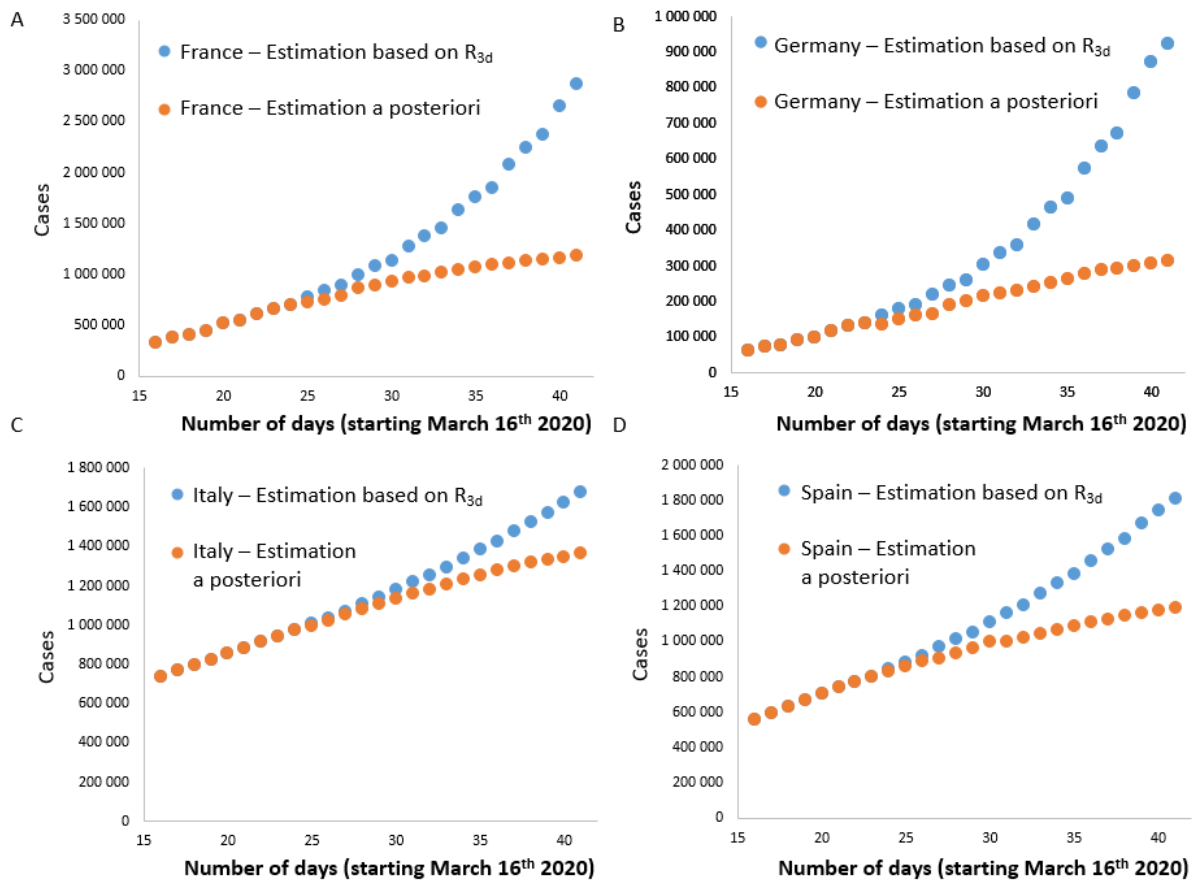
11
12
13



1

2 **Fig. 3.** Comparison of estimated cases and related parameters in different countries. A.
 3 estimated cases in different countries. B. Proportion of affected person compared to the
 4 country population. C-D. Notifications indexes reflecting the number of cases reported by
 5 different countries compared to the estimated number of cases (percentages).

6



1
2 **Fig. 4.** Comparison of cases estimated in different countries from day 16 (March 16th) using a
3 method based on fatality rates reported from March 16 to April 10 with an estimate based on
4 R_{3d} (blue). Results of an a posteriori method following the fatality rates reported from April 3
5 to April 28 are shown in orange.