



**HAL**  
open science

## Simulation of the covid19 epidemic in France

Loïc Pottier

► **To cite this version:**

| Loïc Pottier. Simulation of the covid19 epidemic in France. 2020. hal-02555718v2

**HAL Id: hal-02555718**

**<https://hal.science/hal-02555718v2>**

Preprint submitted on 30 Apr 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.

# Simulation of the covid19 epidemic in France

Loïc Pottier<sup>1</sup>  
(work in progress)

## Abstract

We simulate the evolution of the covid19 epidemic in France on a computer. With approximately 15,000 simulations performed with random epidemic parameters, we obtain a value of the basic reproductive number  $R_0 = 8.8$ <sup>2</sup>. For the mortality rate (infection fatality ratio), we get  $IFR = 2.3\%$ <sup>3</sup>. During the social distancing period (March 9 - March 16), we get  $R_0 = 2.2$ . For the lockdown period, which started on March 17, we obtain  $R_0 = 0.96$ .

We can conclude that lockdown will not stop the epidemic, and that releasing it will cause a new outbreak.

With these values, the simulations carried out give a number of infected persons of approximately 930,000 currently (1.5% of the population) and predict limits of 1.5 million infected persons (2.5% of the population) and 35 000 deaths, limits reached in about 8 months, and 30,000 at the end of July.

Code and results are daily updated on a github depository<sup>4</sup>.

---

<sup>1</sup> HDR, France

<sup>2</sup> all credible intervals and coefficients of determination are given in the following sections

<sup>3</sup> which is compatible with some recent studies ([7] gives 1.29% (CrI 0.89-2.01) in Lombardy)

<sup>4</sup> <https://github.com/loicpottier/covid19>

## Introduction

We simulate the covid19 epidemic in France by computing, for each individual in a population of 60 million people (with a few people infected at the start), their condition over the days (around 100 days), according to various parameters : virus transmission rate ( $R_0$ ), mortality (infection fatality ratio IFR),  $R_0$  during lockdown, etc.

The simulations are performed by a program written in OCaml [6], compiled natively and executed under windows / bash on a PC. It requires 2G of RAM.

Numerous sets of parameter values ( $> 15,000$ ) were taken at random, inside large intervals. Some parameters are fixed and do not vary: the duration of incubation, the duration of the disease, the number of neighbors. For each simulation, we determine the error with reality, from the numbers of deaths each day (deducted from patients in intensive care each day in France). The sets of values of the parameters giving small errors are retained.

It appears that sets of different values can give simulations very close to reality. It is therefore difficult to accurately predict the values of the parameters (such as  $R_0$ , the rate of virus transmission, or IFR, mortality).

We proceed in two steps. First we explore numerous sets of parameter values taken at random inside large intervals. By quadratic regression we get values for the parameters. Second, we explore smaller intervals around these values. Then we obtain better simulations (with smaller errors). Again, by quadratic regression, we get final candidates values for parameters.

We also carried out simulations by taking the values of  $R_0$  and IFR given by [1] (preprint pasteur-02548181, team of Simon Cauchemez). After carrying out simulations with various values for the other parameters, we come to the conclusion that they can be valid until April 14 (date of the end of the data used in this preprint), but not for the period of April 15 to April 22:  $R_0$  between 0.5 and 0.55 during lockdown must be revised upwards: we show that it is around 0.93.

In conclusion, and this is in line with the conclusions of [1]: uncontrolled release of the lockdown would cause a further runaway of the epidemic, and the proportion of the population which has been in contact with the virus is low ( $<4\%$ ).

In view of the simulations carried out, we can also conclude that lockdown is not very effective, and wonder if the rate of virus transmission has been underestimated.

## Method

### Data

The data used are the number of patients in intensive care for covid19, from March 4 to April 25 [2]. It is more regularly updated than the number of hospital deaths. It is closely linked to it: for covid19, the number of patients in intensive care on a given day is equal to 7.5% of the number of deaths on the same day, to within 1%. To study the epidemic, we can therefore reduce ourselves to studying the number of patients in intensive care each day.

The French population is modeled by 60 million individuals, placed in a square of 700 km aside.

Each individual has a state:

- 0 if he has never been in contact with the virus
- 1 or more: this is the number of days since he was infected with the virus
- -1 if he died

## Parameters

They are separated into 2 groups:

Variable parameters:

- **R0**: the number of people contaminated by an infected person during the period in which they are contagious<sup>5</sup>
- **dR0**: the divider of R0 during the social distancing period: from March 9 to March 16, ie R0 becomes  $R0 / dR0$ .
- **R01**: R0 during lockdown, ie from March 17th.
- **pvoy**: the rate of travelers (people who travel far during a day)
- **dpvoy**: the divider of the rate of traveler during lockdown.
- **debi**: the number of days before an infected person becomes contaminant.
- **duri**: the duration during which an infected person is contaminant.
- **IFR**: the mortality rate (Infection Fatality Ratio), ie the probability that an infected person has to die of disease.

---

<sup>5</sup> as recontaminating an already contaminated person has no effect, this number is effective only at the start, when few people are infected: it becomes more and more difficult for the virus to infect new people: the growth in the number of infections is exponential at the start, then becomes constant and finally decreases towards 0, the number of infected persons tends towards a limit.

## Fixed parameters

- **dinc**: the duration of incubation, from 2 to 12 days, on average 5, taken at random for each individual [3]
- **dmal**: the duration of the disease, fixed at 20 days [3] [4]. An infected person dies with the IFR probability in the period between  $dinc + dmal / 2$  and  $dinc + dmal$ .
- **dvoisins**: maximal distance of neighbors, who are the people likely to be met by a person during a day (25 km).

## Simulation

The country is a square matrix of 7746 x 7746 people (it can be seen as a computer screen with resolution 7746 x 7746). Each pixel is 4.5 km wide. The neighbors of a person are those at distance less than 25 km (Manhattan distance). Thus the number of people likely to be contaminated by a non-traveling person will be on average 300,000 (contained in a square of 51 km per side).

If she is a traveler, the whole population is likely to be contaminated by her.

On day 1 we choose 3 people at random who become contaminated (states = 1,2,3), then one calculates day after day the states of the 60 million people

- each day, a person will or will not contaminate people randomly among its geographic neighbors, or throughout the country, for travelers; for this we use a Poisson distribution with parameter  $R_0$ .
- every day, an infected person can die (Bernouilli's distribution of parameter  $IFR / (dmal / 2)$ ), if he is in the second period of the disease (state between  $dinc + dmal / 2$  and  $dinc + dmal$ ).

The day when the number of deaths exceeds that of reality on March 9 is considered to be March 9 for the simulation, which then stops at the end of the actual data.

At the end, the error is calculated<sup>6</sup> with the actual death numbers each day, reduced to a percentage of the actual death number on the last day.

---

<sup>6</sup> average quadratic error for days with at least 100 deaths

## Results

### Random exploration

#### Step 1

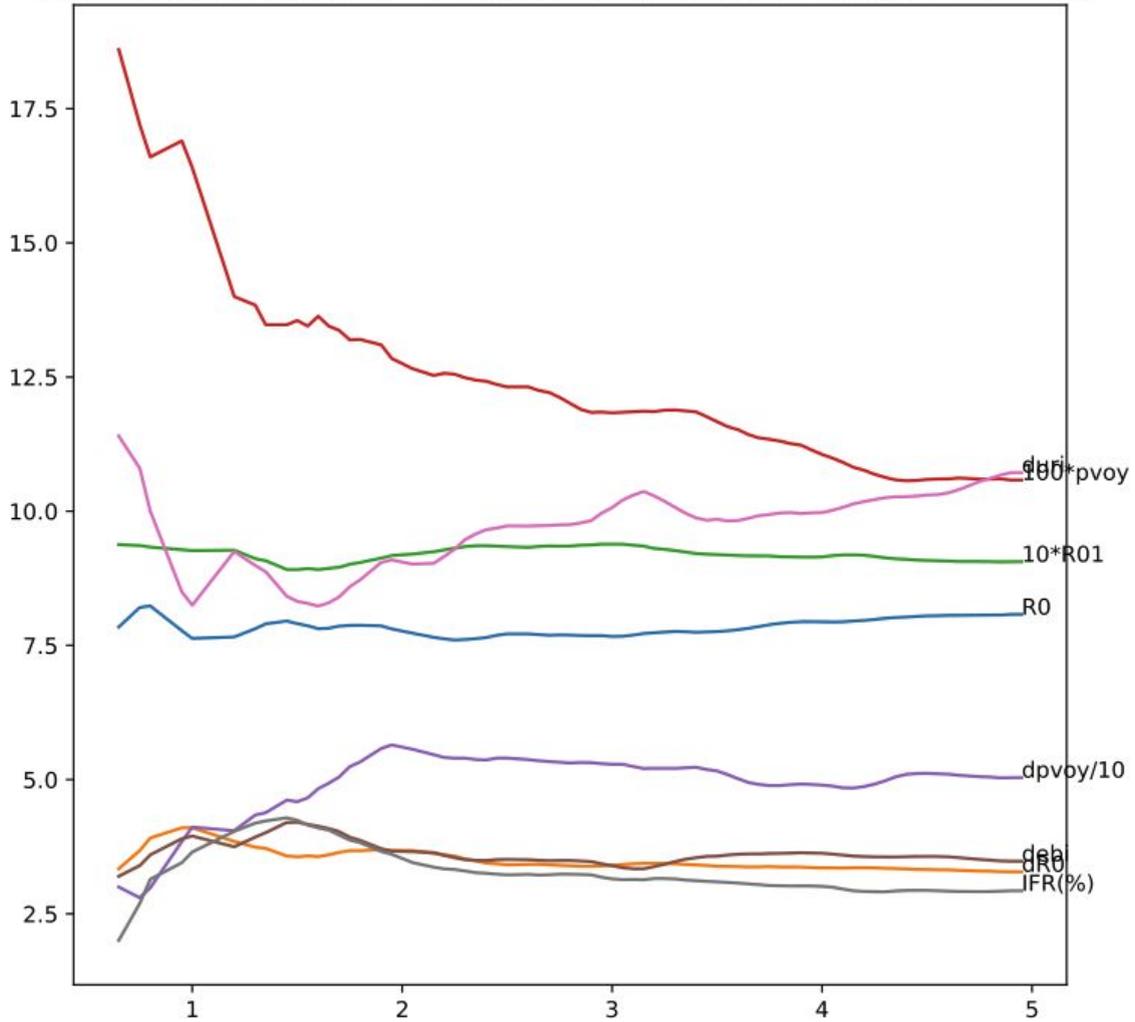
On the data up to and including April 29, 6662 simulations are performed with sets of values of the parameters taken at random in the following intervals:

table 1	min	max
R0	2	10
dR0	1.5	5
R01	0.5	1.2
pvoy	1%	20%
dpvoy	2	100
debi	1	9
duri	3	21
IFR	0.01%	15%

Simulations are aborted as soon as error is more than 20% after March 22.

For the 100 simulations with an error less than 5%, we observe the sets of values of the parameters are dispersed, but each parameter converge to a limit when error decreases to 0:

100 simulations with error < 5.0, average of parameters function of max error



We use linear and quadratic regression to calculate these limits, and obtain the following results, where cri (credibel interval) denotes intervals for standard deviation, limit1(resp. limit2) is the limit obtained by linear(resp. quadratic) regression, r21(resp. r22) is the coefficient of determination for linear(resp. quadratic) regression:

table 2 100 sim. err<5.0	average	cri	limit1 (r21)	cri1	limit2 (r22)	cri2	common interval	final average
err	3.3	2.2-4.4	0.0 (0.0)	0.0-0.0	0.0 (0.0)	0.0-0.0	2.2-0.0	1.1
R0	8.1	6.8-9.3	7.7 (0.1)	7.5-7.9	8.1 (0.3)	7.9-8.3	7.9-7.9	7.9

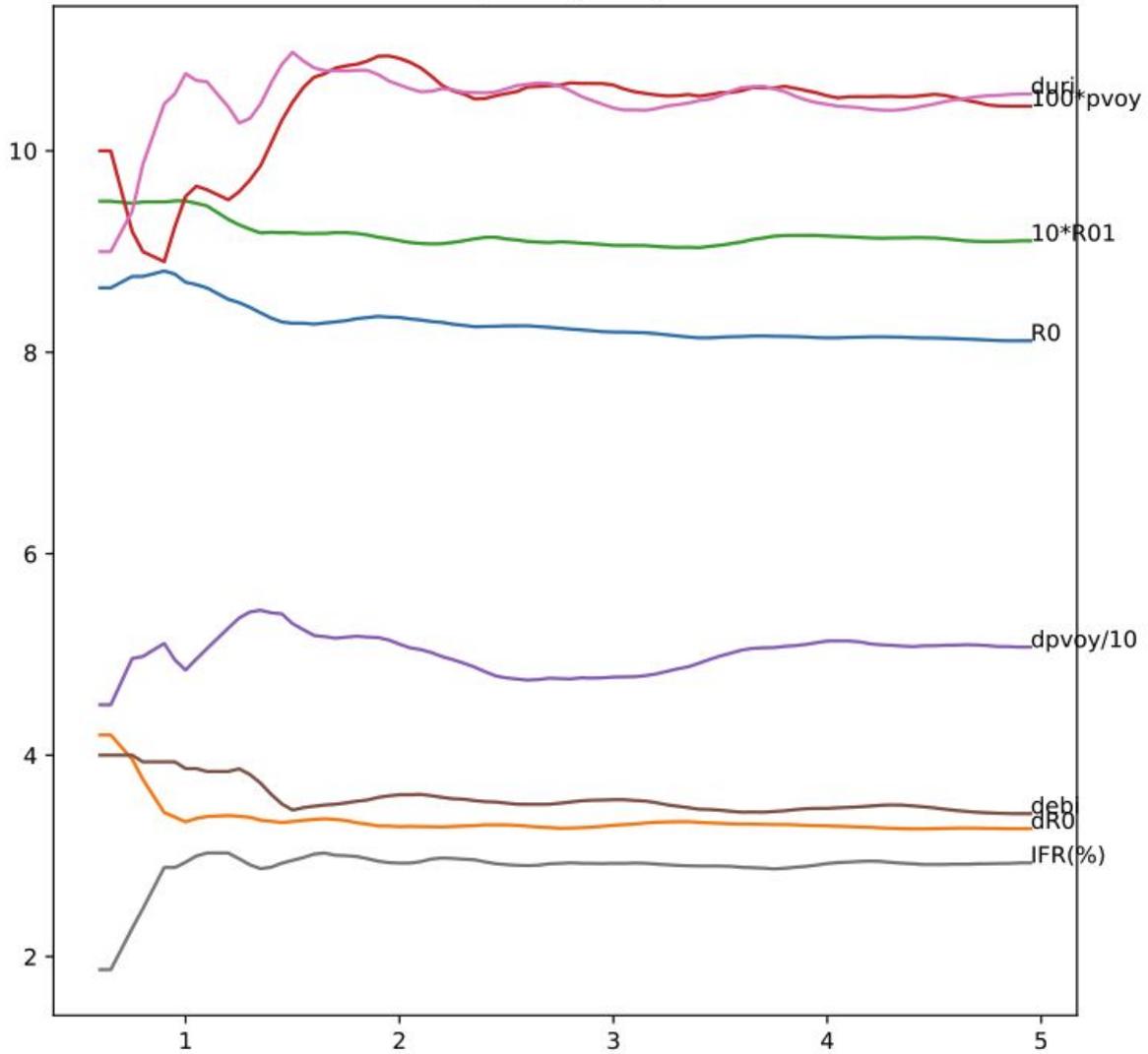
dR0	3.3	2.3-4.2	3.8 (0.4)	3.6-4.0	3.8 (0.4)	3.6-4.0	3.6-4.0	3.8
R01	0.91	0.74-1.07	0.93 (0.1)	0.91-0.94	0.91 (0.1)	0.90-0.93	0.91-0.93	0.92
pvoy	0.11	0.06-0.15	0.17 (0.7)	0.16-0.18	0.20 (0.9)	0.20-0.21	0.20-0.15	0.17
dpvoy	50	21-80	40 (0.3)	34-46	19 (0.7)	15-23	34-23	28
debi	3.5	1.8-5.2	3.8 (0.1)	3.5-4.0	3.7 (0.1)	3.5-3.9	3.5-3.9	3.7
duri	10.7	5.9-15.6	9.0 (0.2)	8.2-9.7	10.3 (0.3)	9.6-11.0	9.6-9.7	9.7
mor	0.029	0.013 -0.046	0.036 (0.1)	0.031 -0.042	0.028 (0.2)	0.023 -0.033	0.031-0.03 3	0.032

## Step 2

We then continue the random exploration with credible intervals of table 2.

With 2707 simulations, we get 58 simulations with error < 5%

125 simulations with error < 5.0, average of parameters function of max error



And values:

table 3 125 sim. err<5.0	average	cri	limit1 (r21)	cri1	limit2 (r22)	cri2	common interval	final average
err	3.1	2.0-4.2	0.0 (0.0)	0.0-0.0	0.0 (0.0)	0.0-0.0	2.0-0.0	1.0
R0	8.1	7.4-8.8	8.7 (0.7)	8.6-8.8	9.0 (0.8)	8.9-9.0	8.9-8.8	8.8
dR0	3.3	2.7-3.8	3.7 (0.3)	3.5-3.8	4.1 (0.5)	4.0-4.3	4.0-3.8	3.9

R01	0.91	0.82-1.00	0.94 (0.5)	0.93-0.95	0.98 (0.8)	0.97-0.98	0.97-0.95	0.96
pvoy	0.10	0.08-0.13	0.10 (0.2)	0.09-0.10	0.09 (0.5)	0.08-0.09	0.09-0.09	0.09
dpvoy	51	32-69	49 (0.0)	47-52	50 (0.0)	48-53	48-52	50
debi	3.4	2.4-4.4	3.9 (0.6)	3.8-4.0	4.2 (0.7)	4.1-4.3	4.1-4.0	4.1
duri	10.6	7.9-13.3	10.2 (0.1)	9.7-10.6	9.2 (0.3)	8.9-9.6	9.7-9.6	9.7
mor	0.029	0.020 -0.039	0.026 (0.1)	0.024 -0.029	0.021 (0.4)	0.019 -0.023	0.024-0.02 3	0.023

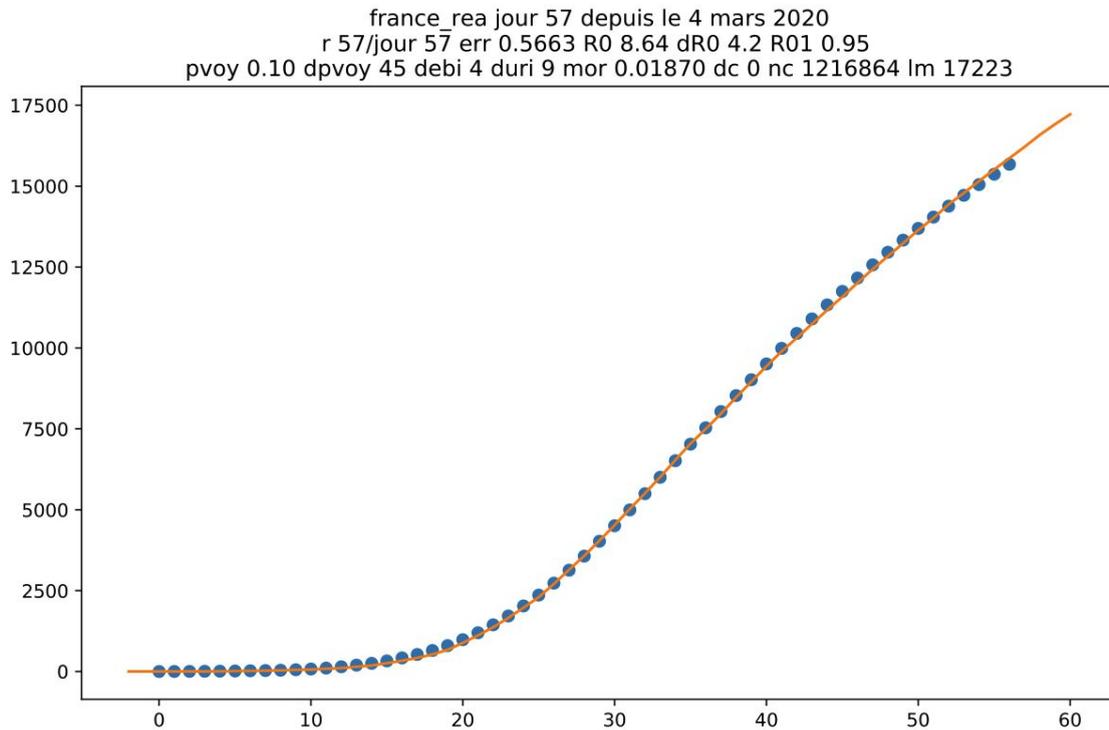
In particular:

- $R_0 = 8.8$ , which shows a contagion much higher than that commonly accepted,
- $IFR = 2.3\%$ ,
- $R_0$  during lockdown is 0.96, which shows that the current lockdown will not stop the epidemic for a long time.

The other parameters indicate that

- the social distancing divided  $R_0$  by 3.9, so  $R_0 = 2.2$  between March 9 and March 17.
- an infected person becomes contagious after 4 days, and then is for 10 days.
- the lockdown divided by 50 the number of travelling people.

Here is the simulation with the lowest error (0.5%):



With the values in last column of table 3 and 820 simulations over 400 days, we got 160 simulations with error < 5%.

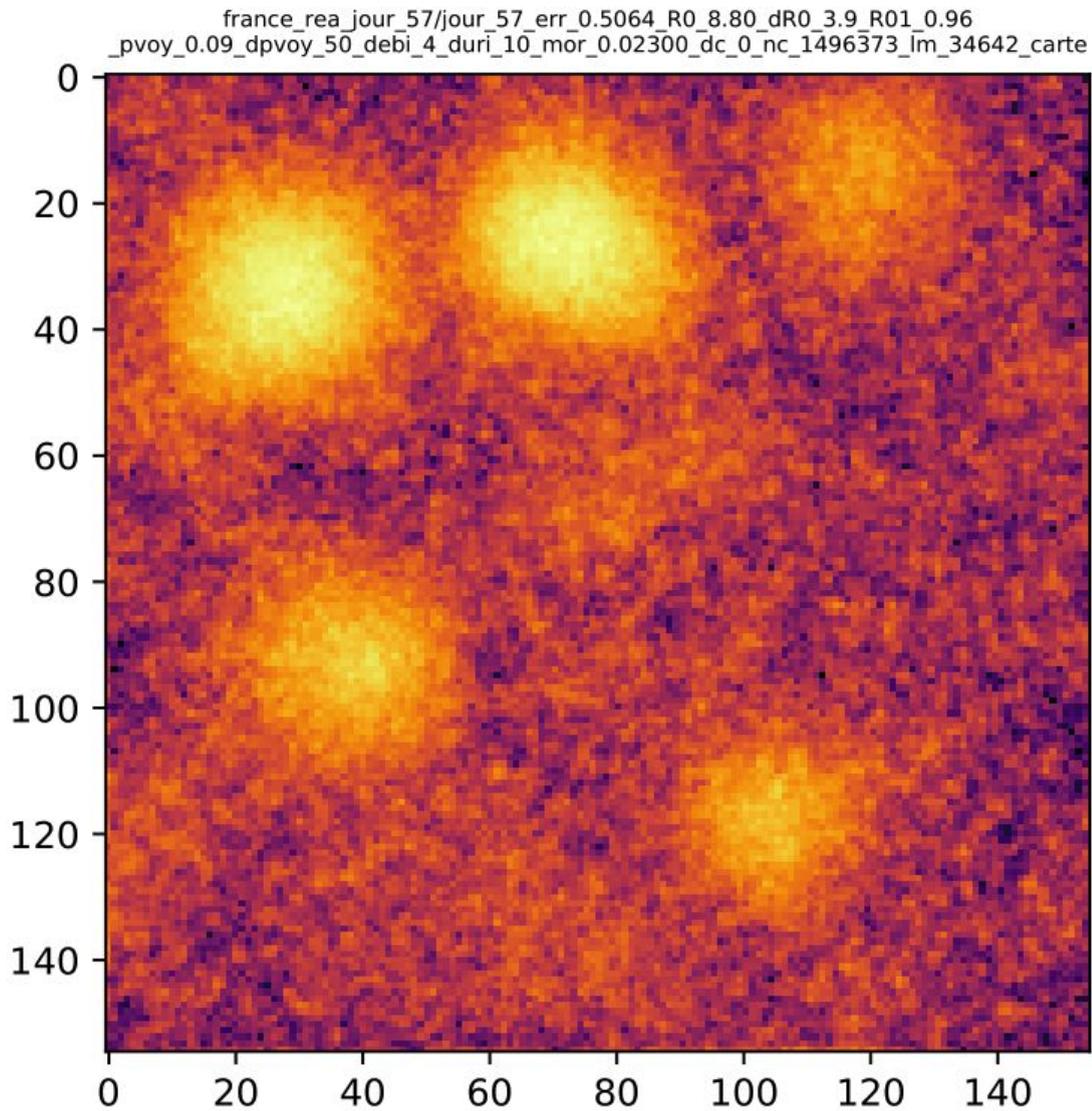
With the best 50, the limit on the number of infected is approximately 1.5 million people, and the limit on number of deaths is 35,000, limits reached in about 8 months.

By the end of the epidemic, 2.5% of the French population will have been in contact with the virus.

On April 29, the number of people infected is around 930,000, or 1.5% of the population.



and, for this simulation, the geographic location of the epidemic on April 29:



We see the 3 initial clusters, and many areas with less infection.

**Simulations on April 24 with the parameters given by [1] (Pasteur-02548181: Henrik Salje, Cécile Tran Kiem, Noémie Lefrancq, Noémie Courtejoie, Paolo Bosetti, et al.)**

The values of R0 and IFR given by [1] relate to a period ending April 14:

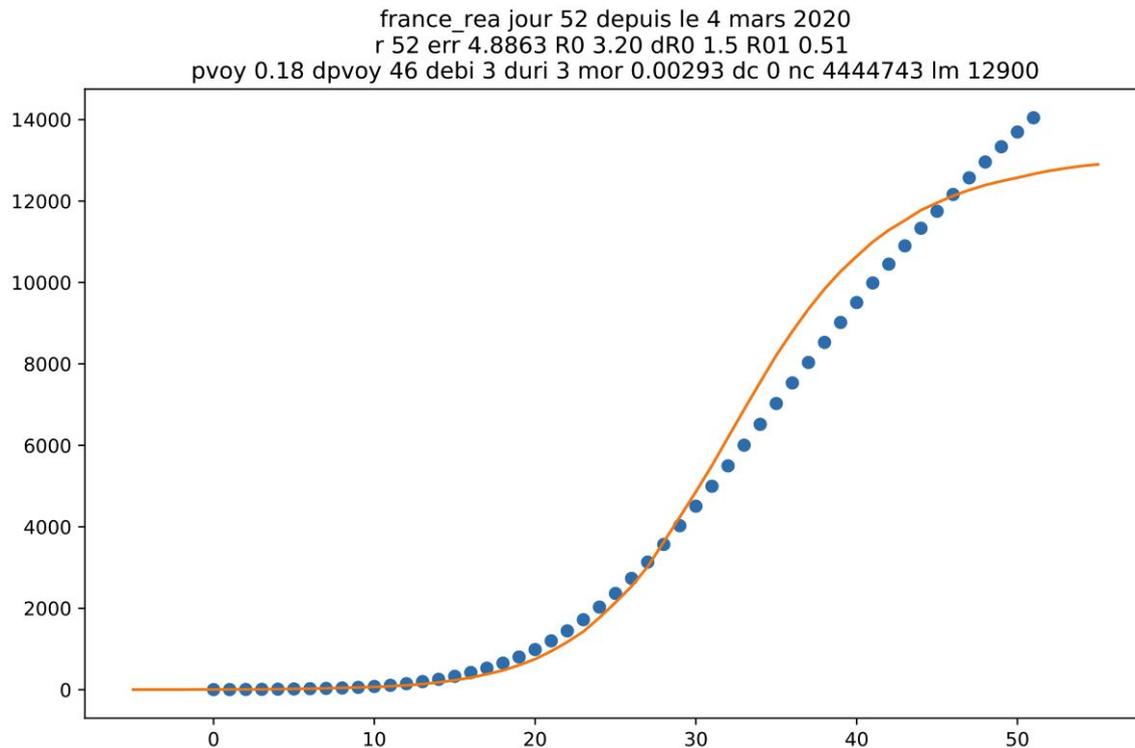
- R0 = 3.31 (95% CrI: 3.18-3.43)
- R01 = 0.52 (95% CrI: 0.50-0.55)
- IFR = 0.53% (95% CrI: 0.28-0.88)

We then perform a random exploration of sets of parameter values with R0, R01 and IFR in these confidence intervals, and the values of the other parameters in the intervals of the previous paragraph:

	min	max
R0	3.18	3.43

dR0	1.5	5
R01	0.50	0.55
pvoy	1%	20%
dpvoy	1	100
debi	1	9
duri	3	15
IFR	0.28%	0.88%

Among 1156 simulations, here is the best simulation (4.8% error):



Until April 14, the simulation is fairly close to the actual data, but moves away from it later. The same is true for other simulations. As R0 and IFR are consistent with recent studies on the Chinese epidemic, one might think that R01 is too low.

Keeping R0 and IFR and exploring values of R01 between 0.5 and 0.99, out of 1100 simulations, we obtain only 37 error simulations lower than 20%, with an average value of R0 = 0.84 (CrI: 0.73-0.95) after lockdown.

In conclusion, it therefore seems unlikely, with the ICU data until April 24, and under the hypotheses of [1], that R0 during lockdown is between 0.5 and 0.55. It is rather higher than 0.7.

## Conclusion

With the described modeling (60 million individuals simulated for around 100 days), we see the following facts.

For values of the parameters taken in wide intervals (random exploration), and for 10,000 simulations:

- for different values of  $R_0$  and IFR, we can obtain simulations with similar small errors (<2% ).
- $R_0 = 8.8$ , which shows a contagion much higher than that commonly accepted,
- IFR = 2.3%, which is not contradictory with some recent studies ([7] gives 1.29% (CrI 0.89-2.01) in Lombardy)
- $R_0 = 0.96$  during lockdown, which shows that the current lockdown will not stop the epidemic for a long time.
- Social distancing and the ban on large meetings from March 9 to March 17 reduced  $R_0$  to 2.2.
- an infected person becomes contagious after 4 days, and then is contagious for 10 days.

Assuming, following [1],  $R_0$  between 3.18 and 3.43 and mortality (IFR) between 0.28% and 0.88%, the 2000 simulations carried out suggest that  $R_0$  is not between 0.5 and 0.55 during lockdown, but rather between 0.73 and 0.95.

In both sets of values of  $R_0$  and IFR,  $R_0$  during lockdown is around 0.9, which is quite high: the epidemic will continue for a long time, a reduction in lockdown (therefore  $R_0 > 1$ ) will cause it to flare up again.

Finally, our study suggests that  $R_0 = 8.8$ , IFR = 2.3%,  $R_0 = 0.96$  during lockdown.

Could it be that  $R_0$  and, to a lesser extent IFR, have been underestimated in most studies on the epidemic?

## Bibliography

[1] Henrik Salje, Cécile Tran Kiem, Noémie Lefrancq, Noémie Courtejoie, Paolo Bosetti, et al. Estimating the burden of SARS-CoV-2 in France. 2020. ? Pasteur-02548181

[2]

[https://geodes.santepubliquefrance.fr/#c=indicator&f=0&i=covid\\_hospit.rea&s=2020-04-22&t=a01&view=map2](https://geodes.santepubliquefrance.fr/#c=indicator&f=0&i=covid_hospit.rea&s=2020-04-22&t=a01&view=map2)

[3] <https://www.pasteur.en/fr/medical-center/disease-sheets/illness-covid-19-new-coronavirus>

[4] [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)30566-3/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30566-3/fulltext)

[5] [https://wwwnc.cdc.gov/eid/article/26/7/20-0282\\_article](https://wwwnc.cdc.gov/eid/article/26/7/20-0282_article)

[6] The OCaml language <https://ocaml.org/index.fr.html>

[7] <https://www.medrxiv.org/content/10.1101/2020.04.15.20067074v2>