Realistic estimate of the cumulative and instantaneous Covid-19 incidence in France

Guillaume MALPUECH

Institut Pascal, CNRS, Université Clermont Auvergne, Sigma Clermont, F-63000 Clermont Ferrand, France

Abstract:

Based on a simple analysis of the reports of Santé Public France, we show that the real incidence during the second wave of the Covid-19 outbreak is around 60% of the mean positivity of virological tests realized week after week in France. The fraction of the French population infected at the beginning of September is about 3% (2 million people). Both intensive care units (ICU) admission and death rate, which dropped by more than one order of magnitude since March, are currently 0.017% and 0.015% respectively. Simulations of the outbreak evolution based on the hypothesis of negligible reinfection probability are performed for France, Ile de France, Puy de Dome and Bouches du Rhone. The incidence peak of 3.5% is expected at week 39 for France. These values are 4.5% at week 37 for Ile de France (for which the peak already passed), 4.2% at weeks 41-42 for Puy de Dome, 7% at week 38 for Bouches du Rhone. The calculated total number of ICU admission and deaths during the second wave are both found around 3000. The cumulative incidence over the two waves is computed close to 50% for France and 80% for Ile de France. We conclude that Covid-19 is much more spread than previously thought, but its severity became very limited since the end of the first wave. The level of spreading raises doubts in the large-scale testing strategy which anyway detects only a few percents of infected people. It also raises doubts in the efficiency of the closing of school classes, where a pupil is tested Covid-positive by chance. Finally, determining the probability of reinfection by making a large-scale testing of persons already found positive several months ago would strongly help to get a better visibility of the future outbreak trajectory.

I Introduction

To predict the Covid-19 outbreak trajectory and adjust public policies, the knowledge of the total (cumulative) infection rate and the instantaneous infection rate (also called incidence) of a given population is crucial. Since May, large-scale virological testing is executed in France. About 300,000 tests a week were performed in May, a number which is reaching 1 million at the end of August and beginning of September. The positivity of these tests was around 2% in May, fell to 1% in June and then continuously rose to reach 5.2% on September 9th. At first glance, it seems plausible that these numbers tell something about the real fraction of the population infected. However, this testing was neither conceived nor currently analyzed in a way that could provide such information. The idea in May was that the outbreak was not spread, and that to control it one needed to find clusters. The hypothesis made by authorities was that most of the infected people are really found by the tests. The quantity which is put forward and used by the French authorities is the total number of positive tests wrongly called incidence. This number of positive tests is still used as a key parameter to determine the status of a department. In that picture, the fraction of positive tests, also called “positivity” had not much meaning, because the tests were targeted. They were centered at people...
likely to have been infected who form the clusters. However, at least since July, the tests performed are not linked with clusters in their wide majority. A majority of tests is performed out of clusters, on asymptomatic people. Moreover, the majority of positive test are issued from this type of tested population: asymptomatic and not belonging to clusters. It is therefore clear that the positivity measured on this population is a good measure of the real incidence, namely the fraction of the population infected at a given time.

II Determination of the real Covid-19 incidence in France.

To be more quantitative, we consider the report of Santé Public France of August 13th [1], concerning the week 32. Page 32, it is indicated that only 18 % of positive tests are linked with clusters. Asymptomatic persons represent 77 % of tests performed and 53 % of positive tests (page 6). The positivity of asymptomatic persons is therefore just 0.53/0.77 of the total measured positivity, a ratio which is 0.69. The positivity of symptomatic people is 0.47/0.23, which is about twice the total positivity of tests. The data provided do not allow to compute the positivity of asymptomatic persons outside of clusters. We assumed, however, that the correction induced by clusters is weak, because they represent only 18 % of positive tests. We therefore slightly correct the 0.69 ratio found above and state that the positivity of asymptomatic persons not linked with clusters is 0.6 times the total positivity of tests. We claim that, regarding the number of tests performed each week, this is a good measure of the real incidence over the French population.

Therefore, by taking again the 5.2 % of positivity measured on week 36 moderated by the above mentioned factor 0.6 allows to deduce that the fraction of the French population infected at the beginning of September is about 3%, which makes 2 million people. This is the first straightforward result of this manuscript. This of course could be more rigorously evaluated by testing even larger random samples, but it is unlikely to find very different conclusions.

Assuming that infected people are positive during one week allows determining that about 280,000 people are infected each day in France. So the testing finds typically 3 % of infected people, because simply it concerns only 2% of people each week. There are many consequences of this simple fact. It means that 1 person over 33 is infected (1 over 15 in Marseille). It means that there is almost one kid or one teenager infected in each class of France, and for sure several ones in each school. It means that closing a class just because one kid was tested positive has a very little meaning. If we would be able to test everybody, we would just close all teaching places. In general, the large scale testing policy could work only by being able to test a much larger proportion of the population. It can work if one can test very regularly all contacts of a group, as done with professional football players, but in the general population, its only value at that stage is for epidemiologic survey.
III Determination of the real mortality rate and the ICU admissions in France during the second wave.

The mortality rate and the ICU admission probability in case of infection show a very strong age dependency. In fact, only people above 65 significantly contribute to deaths and ICU admissions. Nevertheless, since June at least, there is a reasonable correlation between the ICU admissions and the positivity of tests. The number of weekly ICU admissions reached a minimum around 70 in June and rose to almost 300 the first week of September [1], which makes a factor 4-5. During the same period, the positivity of tests rose from 1% to about 4%. We keep here a one week shift for comparing the ICU entrance and the test results. So, it seems that these quantities are correlated and increased by a factor 4 to 5 since June. Then, considering the infection rate computed in the last section, one can estimate 1 ICU entrance for 5600 positive cases during the second wave.

It is probably less precise to determine a mortality rate regarding the delay between infection and death. In practice, the mortality remained quite constant over the summer, around 70-100 deaths a week, and just starts to rise at the beginning of September. We therefore consider that 1% of the French population (67 millions) infected (value achieved at the beginning of August) corresponds to 100 deaths (value achieved at the beginning of September). This makes 1 death for 6700 infections.

This makes a mortality rate of only 0.015%.

Taking this rate and considering the 30,000 people who died during the first wave in France would mean 200 millions infected, which is of course not possible. This allows to conclude that the mortality enormously dropped between the first and second wave, by around one order of magnitude. Several speculative explanations can be put forward to explain this drop. The younger mean age of infected people is one explanation often mentioned. In the absence of any systematic record during the first wave, this hypothesis cannot be verified. This one order of magnitude decay would require a decrease by 20 years of the mean age of infected people. Another possibility is a weaker activity of the virus on infected people, which can be due to unknown factors. One can cite a possible mutation of the virus or a development of immunity in the population, not necessarily against infection, but against the development of symptoms. Influence of the season might also be evoked.

IV What about the future evolution of the outbreak.

In this section, we present a few basic simulations of the outbreak evolution using the numbers found in the previous section and a few hypotheses. One key hypothesis is that re-infection cannot occur in a significant way. This is fully speculative, but could be checked experimentally easily, as commented in the discussion section. The second hypothesis is the estimate of the effective reproduction rate $\tilde{R}$ which depends on many parameters, such as social distancing, restrictions, fraction of susceptible people. The maximum rise of positivity from one week to another reached 40% at the beginning of August. Regarding the ICU admissions, the maximum increase from one week to another was of 30%. So, our simulations are using an effective reproduction rates ($\tilde{R}$) for France of 1.3 and 1.4, but 1.3 is probably more reliable. We would like to stress that $R$ of the order of 3 was found in March, which allowed to deduce a “herd immunity” threshold of $1-1/R$ of 66 % of a susceptible population. Whatever the reason which reduced the $R$ value, $\tilde{R}=1.3$ means a herd immunity threshold of 23 % of the susceptible population. Positivity started to rise at the end of July (week 30), which is the $t=0$ of our simulation for France. We take 0.8 % of real incidence at that time.
which corresponds at week 31 to 95 ICU admissions following the correspondence established in the previous section. Despite several recent publications [1,4], we assume no cross immunity. On the other hand, we consider that a non-zero fraction of the population was previously infected and could not be infected a second time. The computation of this quantity is explained in the next section, but we take 25 % for France at week 30, 45% for Paris, 20 % for Bouches-du-Rhone, 10% for Puy de Dome. We use a homogeneous propagation model, even if in reality the very strong inhomogeneity and the fact that contaminations occurs mainly through super-spreaders are the key elements of this outbreak. The operation to be computed is

\[ N_{vt} = N_0 R \left( 1 - \frac{I_0}{f_s} \sum_{k=0}^{i} N_k \right) \]

with \( N_0 = 1 \). \( I_0 \) is the initial incidence (0.8 %), \( f_s \) the susceptible fraction of the population when the simulation starts which is computed as one minus the fraction of previously infected people. The fraction of infected people after \( i \) reproductions is \( I_0/N_i \). Figure 1-a shows the computed incidence for France. The curve computed with \( R = 1.3 \) gives an incidence close to 3% at week 37, which is in agreement with the experimental value. This curve for France shows a peak arising at weak 39 (end of September) with a maximal incidence of 3.5 %. With these parameters, the number of ICU admissions per week should go above 400. It could reach 600 if \( R \) is closer to 1.4. Figure 1-b shows the cumulative incidence and ICU admissions. With \( R = 1.3 \), about 25 % of the population would be infected, and there should be around 3000 ICU admissions and 3000 deaths. The total fraction of the population infected including the first and second wave would be 50 %. We then realized simulations which describe the situation in three different regions of France, which are Ile de France, Bouches du Rhones, and Puy de Dome. Table 1 shows the parameters used.

| Region              | Population (Millions) | \( R \)  | \( t=0 \) | \( I_0 \) | \( f_s \) |
|---------------------|-----------------------|--------|---------|--------|--------|
| France              | 67                    | 1.3-1.4| Week 30 | 0.8 %  | 75 %   |
| Ile de France       | 12                    | 1.3-1.4| Week 30 | 1.1 %  | 55 %   |
| Bouches du Rhones   | 2                     | 1.45   | Week 30 | 0.8 %  | 80 %   |
| Puy de Dome         | 0.65                  | 1.3    | Week 32 | 0.8%   | 90 %   |

Table 1.

Figure 2 shows similar simulations for Ile de France. At week 30, the positivity of test was 1.8 % which gives an incidence of 1.1 %. The positivity in Ile de France over weeks 36-37 is around 7 % which corresponds to an incidence of 4, 4.5 %. This is better reproduced by taking \( R = 1.4 \) instead of 1.3. It is not extremely surprising that the reproduction rate in Paris, even strongly modulated by pre-existing immunity, could be larger than an average for France. The peak is expected to occur for Ile de France at weeks 36-37, which is the period of writing of this manuscript. Interestingly, we notice that the positivity observed at the end of the week 36, beginning of 37 seems to stabilize and even to decay slightly. It is too early to say if this is a fluctuation or a tendency, but it is compatible with the results of the simulations. The cumulative infection rate is above 35%, which combined with the estimated 45 % at week 30, gives a total fraction of population infected of 80 % in Ile de France at the end of October. Figure 3 shows a simulation performed for Puy de Dome. Puy de Dome is probably representative of an average for France in terms of urban density, but it is geographically quite
isolated and has been weakly affected by the first wave. We therefore consider a preexisting cumulative infection rate of only 10% and $R = 1.3$. The second wave of the outbreak really emerged two weeks later than in the rest of France and we take $t=0$ on week 32. As a result, the peak occurs only on weeks 41-42 (mid-October) and the incidence could remain significant till the end of the year. The ICU admission could reach 4 per week, 50 admissions could occur in total. Figure 4 shows the simulation for the last considered area which is Bouches du Rhone, which cumulate a “large” reproduction rate being a urban area (we take $R = 1.45$), but was quite preserved during the first wave. The simulation suggests that the peak could be achieved at week 38, with an incidence of almost 7 % (more than 10 % for test positivity) and 25 ICU admissions a week. One should notice that this value is significantly lower than the one achieved at week 36 which is 45. However regarding various uncertainties, this discrepancy is acceptable.
Figure 1. France. a) Time evolution of incidence. The horizontal dashed line shows the estimated ICU admissions each week for an incidence of 3%. b) Time evolution of cumulative incidence. The horizontal dashed line shows the estimated ICU admission for a cumulative incidence of 30%.
Figure 2. Same as Figure 1 for Ile de France.
Figure 3. Same as Figure 1 for Puy de Dome.
Figure 4. Same as Figure 1 for Bouches du Rhone.
V Cumulative incidence during the first wave.

We remind that very few virological tests were available in March-April during the first wave of the outbreak, and the real incidence during this period is not very well known. Two influential sources in France estimating this value are a numerical simulation of outbreak evolution published in Science \cite{1}, which found a total infection rate of 5.3 %, and, more recently, a serological measurement published by Santé Public France on July 9\textsuperscript{th}, reporting a rate of 6.7 % \cite{2} with a large uncertainty. This last testing was performed during the week of April 10\textsuperscript{th}. One should notice that the antibody rate decays versus time after the infection, as now well documented. Testing performed late after the outbreak peak will probably show comparable or even lower values as it happened for serial measurements performed in other countries \cite{3}. This decay explains why everywhere these tests are showing relatively low values. More precisely, they rise at the heart of the outbreak then peak and probably decay. They do not follow the dynamics of other signals such as hospitalizations and ICU admissions and should rather be understood as a kind of delayed instantaneous signal rather than a measurement of cumulative incidence. If we take the case of France, the measurement performed on April 10\textsuperscript{th} is a minimum value of the fraction of people infected before March 26\textsuperscript{th} (to let time to the antibody response to grow). Then, we analyze the number of hospitalizations and ICU admissions related to the first wave, so occurring before June 1\textsuperscript{st}. The quantity of interest for us is the number of hospitalization and ICU admission which occurred in April and May, compared to ones happening in March. The number of ICU admissions is the same for the two periods. The number of hospitalizations is 1.5 times larger after April 1\textsuperscript{st}, compared to March \cite{4}. It is therefore logical to assume that the real cumulative incidence is between 2 and 2.5 times 6.7 %, which is already a lower bound. This yields cumulative incidence between 13.4 and 16.75 %, which can be approximated to 15% for simplicity, which makes 10 million people. This is moreover assuming that the severity of infection remained constant, whereas as discussed previously, this severity appears to have strongly decayed with time. Then, comparing mortality rates, one gets that the Ile de France region gets an infection rate 2.25 times larger than the mean value in France, which makes a cumulative infection rate of 33 %. Finally, for the period June-July, we estimate the cumulative incidence using the positivity of tests, which allows to estimate a cumulative incidence around 7% for this period for France, and a quite larger for Ile de France (12 %). One could also notice that the mortality during the first wave can be evaluated to 0.33 % (30.000 deaths, 10 millions infections) which is 20 times more than the one we estimate for the second wave.

VI Discussion-Conclusion

This work is essentially based on the analysis of Santé Public reports. It enunciates an evidence. Testing a large population of asymptomatic people randomly in a large population allows to determine the real incidence of this population. As a consequence, the number of infected people in France, at a given time, at the beginning of september is not 50.000 but rather 2 millions. This counting also leads to another evidence, the mortality rate, the probability of severe forms decayed since march-april by more than a factor 10 and is now about 0.015%. Simulations assuming that re-infections are not likely show that the second wave peak will occur in the second half of September. The total number of infections during this wave is comparable, or even larger than the one (not fully known) related to the first wave, which was controlled by a severe lockdown. However, the expected
number of death is 10 times smaller. The validity of these simulations also rely on keeping the bare transmission rate as it is now, namely limited by social distancing and restrictions.

**Acknowledgments:** The author thank Dmitry Solnyshkov for discussions, corrections of the manuscript and technical support. He thanks Anne Tournadre for discussions, corrections of the manuscript. I acknowledge the support of the ANR program "Investissements d'Avenir" through the IDEX-ISITE initiative 16-IDEX-0001 (CAP 20-25).
References

1 https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-13-aout-2020

2 https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-3-septembre-2020

3 J Braun et al, Nature https://doi.org/10.1038/s41586-020-2598-9 (2020).
4 J. Mateus et al, Science, DOI: 10.1126/science.abd3871 (2020).
5 H. Salje et al., Science 369, 208, (2020).
6 In French: https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-9-juillet-2020
7 Miguel A. Hernan et al, The lancet, DOI:https://doi.org/10.1016/S0140-6736(20)31483-5.
8 These data can be found in various Santé public France report. The easier source remains probably the dedicated Wikipedia page https://fr.wikipedia.org/wiki/Pand%C3%A9mie_de_Covid-19_en_France or data gov. https://dashboard.covid19.data.gouv.fr/vue-d-ensemble?location=FRA . 40.000 people were hospitalized before april 2nd included. The total number of hospitalization on august 7\textsuperscript{th} is 107.000 which makes a ratio 2.6.