Lockdown exit and control of the Covid-19 epidemic: group tests can be more effective

Monograph:

Gerschel, Ellie, Gollier, Christian and Gossner, Olivier (2020) Lockdown exit and control of the Covid-19 epidemic: group tests can be more effective. IPP Policy Briefs (54). Institut des Politiques Publiques, Paris, FR.

Institut des Politiques Publiques

LOCKDOWN EXIT AND CONTROL OF THE COVID-19 EPIDEMIC: GROUP TESTS CAN BE MORE EFFECTIVE

IPP Policy Briefs

54

April 2020

Elie Gerschel
Christian Gollier*
Olivier Gossner*

www.ipp.eu

*Authors of the reference study

Abstract

The lack of efficient mass testing tools for SARS-CoV-2 virus that causes Covid-19 has contributed to the accelerated spread of the epidemic. Infected people are unaware that they are spreading the disease during the incubation period as well as in asymptomatic cases or cases with mild symptoms. To limit the number of victims of the epidemic, the strategy adopted by most affected countries is therefore social distancing or complete lockdown, a strategy that can only be beneficial for a limited time, given its economic and social cost. Today, the most feasible way out of the stalemate requires widespread screening of the population. Such screening would make it possible to isolate infected people and allow others to leave the lockdown. However, production capacity for SARS-CoV-2 tests is limited. Although production is increasing, it will not allow for sufficiently systematic and frequent screening to permit the lifting of health restrictions. We here describe how the usefulness of each test can be amplified by applying it to the mixture of samples from several individuals. This technique, called group testing, has already been successfully applied on SARS-CoV-2. We show how the group-test method must be calibrated to maximize the usefulness of each available test.

- To stop the spread of the epidemic, lockdown is now the core strategy of the most affected countries, but it is too costly to be sustainable. Mass screening could break the deadlock.
- The lack of tests currently limits this screening to a few tens of thousands of people per week in France, and the production of tests is not increasing fast enough to envisage an exit from the current lockdown.
- The group-testing technique multiplies the usefulness of each test. A group test, carried out on the mixture of samples taken from n people, when it proves negative, makes it possible to end the lockdown for all members of the group.
- The method requires precise calibration of group sizes to optimize its efficiency. It is most useful when prevalence is low, i.e. the infected proportion of the population at the time of testing is low.
- Under a prevalence scenario of 2%, each test can allow, on average, 18 people to exit lockdown and return to work.
- The usefulness of the tests justifies massive investment in increasing production, particularly as group tests significantly increase their effectiveness.
- Our approach demonstrates the value of the proven method of group testing, which needs to be experimented with on a larger scale.

The Institut des politiques publiques (IPP) has been developed through a scientific partnership between the Paris School of Economics (PSE) and the Centre for Research in Economics and Statistics (CREST). IPP's aim is to promote quantitative analysis and evaluation of public policy using cutting-edge research methods in economics.

CREST CENTER FOR RESEARCH IN ECONOMICS AND STATISTICS

PARIS SCHOOL OF ECONOMICS ECOLE D'ECONOMIE DE PARIS

Introduction

Which strategy to exit lockdown?

In the absence of mass capacities for testing SARS-CoV-2 infections, lockdown (or compulsory social distancing measures in WHO terminology) is the only option available to contain the current Covid-19 epidemic. Symptoms appear in some infected people, and after an incubation period of up to two weeks. During this period, and longer for asymptomatic carriers, the infected person carries the virus without knowing it, and therefore has a high chance of infecting others.

Restrictions on mobility come at a high economic, social and human cost. The financial cost is being mostly absorbed by the public budget deficit, but the economic cost is likely to explode if the situation proves to be long lasting, or if new waves of infections come later. This policy brief focuses on the economic cost, for which it is easier to give quantitative estimates, which should be interpreted as a conservative estimate of the total cost.

The main obstacle on the path out of lockdown is the risk that some individuals carrying the virus will infect others. While waiting for an effective treatment or vaccine, exiting lockdown without sufficiently controlled conditions would lead to a second wave of the epidemic, hitting hospitals and healthcare workers already weakened by the first. In previous epidemics, the second and third waves were even more deadly than the initial outbreak.[1]

There is hope that serological tests will soon be standardized to identify people who have developed immunity. These people developed at least partial immunity against the SARS-CoV-2, and may be protected from getting infected a second time. Even in the best case scenario in which they could all be identified and their immunity was total, releasing only these people from lockdown would be insufficient to restart economic activities as they are too few of them. An estimate by Imperial College (Flaxman, Mishra, Gandy, et al., 2020) tells us that no more than 7% of the French population has been infected.[2] To avoid a new wave of the epidemic, more than 50% needs to be immunized. Given today's estimated mortality rate of approximately 3%, reaching such a proportion would imply almost 7 Million fatalities in the European Union. Identifying those who have developed immunity will therefore not be enough for a safe exit from lockdown.

Testing is a scarce resource

The best strategy possible to safely relax lockdown is to identify a large number of healthy people through mass screening of the population. This can be done with the help of existing "PCR" tests, the results of which are now available in less than 24 hours. This test has already made it possible to isolate people with a positive result so that they do not infect healthy people. Above all, it can clear people who test negative to be released from lockdown, allowing the gradual recovery of the country's economic and social life.

In most countries, testing capacity is not increasing rapidly enough to satisfy the needs for systematic screening at the individual level. On March 28, the French government declared that only 84,000 tests are being conducted per week, with a target of around 350,000 tests per week. The United States plans to reach 1.2 million tests per week for a population of 330 million. Even in Germany, where there are 500,000 tests per week, the order of magnitude is still a long way from that of generalized individual testing, which would allow many people under lockdown to return to work. **Each test is a scarce and valuable resource, and its usefulness must be maximized**. This is precisely what is made possible by the technique of group testing, which has already been experimented with Covid-19; for example, in Israel, the United States, Germany and South-Korea.[3]

Using group tests

The principle of group testing dates back to the work of Dorfman (1943) and syphilis tests for US military recruits. A sample must first be taken from each individual, but instead of performing the test on a single swab, samples from members of the group are mixed. A single test is performed on this mixture, to reveal the presence or absence of the virus in the mixture. This method represents a loss of accuracy if the test is positive because we do not know who carries the virus. But if the test is nega-

[1] The last four pandemics, namely the Spanish flu in 1918, the Asian flu in 1957, the Hong Kong H3N2 flu in 1968, and the A (H1N1) flu of 2009, were all marked by several waves, the first ones being less deadly than those that followed (Miller et al., 2009).

[2] The mean estimated value is 3%, but the error margins are high. Prevalence is probably between 1% and 7%.

[3] At the Technion - Israel Institute of Technology, in hospitals in Nebraska, and at the University Hospital of Goethe University Frankfurt

tive, we know that all members of the group are healthy individuals. A single test, if negative, therefore provides information on the status (infected or not) of several individuals, which multiplies the usefulness of each test, and allows us to optimize the use of this rare resource.

A key issue is the choice of group size. If the groups are too large, the result will too often be positive, providing little information. If the groups are too small, group testing will not deliver its full potential. In this policy brief, we discuss three different strategies for implementing the group-test method, depending on the chosen objective:

1. measuring the proportion of the population that is infected, known as prevalence
2. organizing the exit from lockdown
3. individually identifying the infected people

We will see that in all three cases, group testing gives results that save a significant number of tests.

It is important to note that testing each person once is not a sufficient solution. A person who is not infected at the time of the test can become infected afterwards. This implies that a strategy for mass exit from lockdown must be based both on periodic group tests and on monitoring the social interactions of infected people, in order to very quickly detect and isolate outbreaks as soon as they recur.

Practical application of group tests

Screening consists of a sample collection step and a testing step. Sampling is performed by trained medical staff using a disposable swab to extract material from the upper respiratory tract. The sample must then undergo several treatments: inactivation of the SARS-CoV-2 virus; extraction of genetic material; multiplication of the genetic material and detection of virus genetic markers through PCR (*Polymerase Chain Reaction*).

The sampling stage involves relatively light medical equipment and can be performed by medical personnel trained in a few hours to perform the required procedures. Although there may be occasional strain on the supply of swabs, their technological simplicity means that their availability can be greatly increased. On the other hand, detecting the virus in samples by PCR requires access to reagents and specialized machines, as well as the availability of highly qualified technicians. There is wordlwide shortage of necessary reagents and scaling up production is difficult.

The group test consists of mixing the samples obtained from n individuals, and applying all the following steps only on the mixture of these samples. If the test result is negative, no one in the group is infected. If the test result is positive, then *at least* one of the group members is infected.

PCR procedures are more than 95% reliable in virus detection. However, in some cases, individuals can be ill from Covid-19 while their organisms have already cleared out SARS-CoV-2 from their upper respiratory tract. Non-detection of these cases is not due to a technical limitation of PCR testing, but hinges on the distinction between Covid-19 medical cases and SARS-CoV-2 carriers. Group testing does not exacerbate such cases, as they would return negative whether tested individually or in a group.

Since group testing involves dilution of samples, it may push PCR to its detection limits when applied to large groups. Research conducted at the Technion Israël Institute of Technology showed that PCR is able to identify a positive sample combined in a group of 64. These results will need to be reproduced and validated if the method is to be generalized at a large scale.

What objectives can the group-testing method achieve?

The value of group tests in estimating viral prevalence

The first objective - to establish the proportion of people carrying the virus - is essential in the fight against the epidemic. Knowing better the prevalence rate allows to control the spread of infections, to estimate the proportion of serious cases, or to estimate the lethality of the virus. It is also essential for identifying which geographical areas and which age or socio-economic categories are most affected.

At the moment, several countries rely on proxies such as hospital admissions or death numbers to monitor the progression of the epidemics. However, not all serious cases reach the hospital, and this proportion probably fluctuates as hospitals become overcrowded and guidelines change. Deaths due to Covid-19 are also counted approximately only. In addition, the length of the period of incubation and worsening of symptoms means that this information comes late in relation to the extent of the spread of the epidemic. For example, estimates obtained as of 30 March 2020 by the Imperial College epidemiology team (Flaxman, Mishra, Gandy, et al., 2020) are imprecise, ranging from 1 to 7%.

The simplest method would be to test a randomly selected sample from a national population (or from the population of a region or department, to obtain geographic estimates). The amount of individual tests needed to obtain a sufficiently precise estimate would be very high whereas the group-test method would allow equivalent or better precision with far fewer tests.

Assuming a prevalence of 2% (i.e. 2% of the population is infected), a set of 12,000 tests on randomly selected individuals yields a margin of error of 0.25% (see Box 1), which is small enough for the results to be used to guide public policy decisions.

A similar error margin can be obtained with the group-test method, by creating instead 600 groups of 20 people (same number of people), and thus using only 600 tests instead of 12,000. This represents a drastic saving in the number of tests to be performed. The group size, set here at 20 people, could be further increased to optimize performance with a fixed number of tests.[4] In practice, the exact prevalence is unknown, and accuracy may be somewhat lower than this estimate.[5] The savings in the number of tests, however, are still considerable.

Austria recently conducted a national testing campaign, in which 5 people out of 1544 returned positive. This gives an estimate of the virus prevalence of 0.33%, with a confidence interval ranging from 0.12% to 0.76%. Testing 500 groups of 48 people would have returned a similar estimate with a confidence interval would range from 0,26% to 0,41%. In this practical application, group testing would allow to cut the number of PCR tests by a factor 3 while being much more accurate.

Optimizing group size to allow exit from lockdown

We are seeking to maximize the number of people allowed to return to work. The objective is that each test should allow release of a maximum number of healthy people. Each individual test releases a maximum of one person. Conversely, a group test, if it is negative, allows the release of all members of the group. We will see that it is possible to choose a group size such that each PCR test authorizes, on average, the release of a number of people well above one.

We can therefore design the following procedure. Let p be the prevalence of the virus (the proportion of the population carrying the virus), and n the size of each tested group. For the moment, we assume that individuals are randomly assigned to groups, and that each individual in a group that tests positive remains under lockdown.

If n is too large, it becomes very likely that a group will contain a carrier of the virus. A large proportion of groups will test positive, and most tests will not be very informative. If n is too small, each negative result will only allow a small number of individuals to be released from lockdown. The challenge is therefore to choose n optimally.

The probability of an individual carrying the virus is p (for example, 2%). The probability of being healthy is, symmetrically, $1-p$ (in this example, 98%). A group gets a negative result if no individual is infected. The probability of this event is $(1-p)^n$ (or 54.5% in this example, with a group of 30 people). In case of a negative result, a test will release n people. On average, each test on a randomly constructed group frees $N = n \times (1-p)^n$ people (in our example, about 16 people).

The objective is to choose n, for a given prevalence p, in order to maximize the effect of each test. A mathematical calculation (first-order condition) yields that n must satisfy the following condition:

$$n = \frac{-1}{log(1-p)} \approx \frac{1}{p}$$

For low prevalences, the optimal group size corresponds to the inverse of the prevalence. The higher the prevalence of the virus, the smaller groups should be. Thus, it is possible to recalculate, on average, how many individuals each test allows to be released (column (3) in Table 1). For a prevalence of 2%, it is in fact optimal to form groups of about 50 individuals. In comparison, an individual test releases on average $1-p$ individuals (the probability that the test is negative). The ratio of the average number of individuals released with a group test to the average number released with an individual test is shown in column (4) of Table 1.

How much should we, collectively, be willing to pay for each test performed? Aside from the social and human cost of lockdown and of the epidemic itself, we can give an approximation of the economic cost (which therefore rep-

[4] For a fixed number of tests, performance is optimal when about 80% of the groups contain at least one infected person. However, this does increase the number of samples required.

[5] If prevalence is actually 3%, using 600 groups of 20 people, the confidence interval obtained is somewhat wider, between 2.66% and 3.37%.

Panel 1: Calculation of confidence intervals

The aim is to estimate the infected proportion of the population (prevalence) and the accuracy of the result, with both individual and group tests. Accuracy depends on the actual level of prevalence. A prevalence of p is assumed, e.g. 2% of the population is infected. A randomly selected individual therefore has a probability of $1-p$ (in this case, 98%) of not being a carrier of the virus.

Individual tests. If the individuals are chosen randomly, then the proportion of infected people that will be measured in the sample, which is assumed to follow a binomial law, has a 95% chance of falling within the range of 1.76% to 2.27%. This range is due to the fact that the sample is an imperfect representation of the population.

Group tests. The probability that no member of a group of 20 people is infected, yielding a negative test for that group, is $(1-p)^{20} = (1-2\%)^{20} = 66.8\%$. The probability that a group will test positive is therefore 33.24% (i.e. 100 - 66.8). Assume we have 600 groups, and the proportion with a positive group test is assumed to follow a binomial law. The proportion of groups testing positive therefore has a 95% probability of being in the range between 29.5% and 37.0%. By doing the previous steps in the opposite direction, we can calculate the level of prevalence to which these thresholds correspond: respectively 1.73% and 2.28% (because $100-(1-1.73\%)^{20} = 29.5\%$ and $100-(1-2.28\%)^{20} = 37.0\%$).

Both methods therefore provide very close intervals, using far fewer tests for the second.

resents a lower bound for the social cost). If putting a person under lockdown costs society q euros, then each test saves about qN euros (the individual cost multiplied by the average number of people released from lockdown).

Although the economic cost remains difficult to measure, as a first approximation we can estimate the cost of putting an individual under lockdown in terms of GDP per capita. Suppose that, in the absence of a test, universal lockdown is required for two months. The cost to society of putting an individual under lockdown is then at least equal to two months of GDP per capita[6]. For the European Union, with a GDP per capita of about 31,000 euros per year, this represents a cost $q =$ 5,167 euros.

Let us try to measure the savings achieved by each individual test and each group test, using the same example of a 2% virus prevalence. In individual testing, the result has a 98% chance of being negative. Each individual test therefore frees an average of 0.98 people, avoiding a cost of 5,063 euros.

Now consider the case of group testing. With such a prevalence, the optimal value of n is 50 people. The probability that the test result is negative is then only 36% (0.98^{50}), but since each group is relatively large, a test releases an average of 18.2 people (0.36×50 =18.2). A single group test avoids, on average, a cost of 94,083 euros.

The last column of Table 1 reveals the economic cost avoided by each group test, clearly indicating that the production of tests is of considerable value to society and justifies substantial investment. The usefulness of the group-test method decreases as the level of prevalence increases.

Table 1: Optimal strategy for group testing, depending on virus prevalence

(1) Prevalence (p)	(2) Optimal size (n)	(3) Average number of "released" (N)	(4) Relative power of group test	(5) Avoided cost in euros (qN)
1%	99	36,60	36,97	189 129
2%	49	18,21	18,58	94 083
5%	19	7,17	7,55	37 046
10%	9	3,49	3,87	18 016
20%	4	1,64	2,05	8 466
30%	3	1,03	1,47	5 317
40%	2	0,72	1,20	3 720

Note: We assume that putting an individual under lockdown costs society the equivalent of the EU's GDP per capita in relation to the duration of lockdown.
Interpretation : For a prevalence level of 1%, the optimal size of the test groups is 99 individuals. On average, each test releases 36.6 individuals, i.e. 36.97 times more individuals than an individual test. This represents an average avoided economic cost of 189,129 euros.
Source : Authors' calculations.

Note that it is not necessary for groups to be selected randomly. Testing all members of a production unit (e.g., a factory or assembly line) at the same time would significantly improve the power of the group-testing method. This is because if any member of a group is positive and has frequent and necessary contacts, it is likely that the other members are also infected. This is also true

[6] INSEE provides an estimate of the effects of lockdown for the first month at 33% of GDP in its last point de conjoncture du 26 mars (INSEE, 2020), taking into account working from home and the fact that some people are allowed to return to their workplace. Our figures are therefore more pessimistic for the economic cost, but can be easily adjusted.

for households.[7] Creating groups in this way can also facilitate the organization of sampling.

Using group tests to obtain individual results

Even when it comes to obtaining individual results, the group-testing method can limit the number of tests required. Testing each individual requires as many tests as there are individuals. However, if a group of size n is chosen, in the event of a negative result, it is possible to give a result for all its members simultaneously. Dorfman's original protocol, already implemented by the Technion Institute of Technology, Frankfurt University Hospital, and Nebraska hospitals, consists of two steps. The first is a group test. If the test is negative, it can be directly deduced that none of the group members is infected. If the initial group tests positive, the individuals are all tested separately. Assuming that the two steps are done consecutively, each individual is therefore tested with a maximum of two samples (only one if the group is tested negative).

We can then calculate the average total number of tests used per individual. If the group tests negative (which happens with probability $(1-p)^n$ as explained above), we use $1/n$ test per individual. If the initial group tests positive (which happens with probability $1-(1-p)^n$), we use $1/n$ test per individual in the first round, and one test per individual in the second round. The average total number of tests used per individual, denoted as $T(n)$, can therefore be expressed as:

$$T(n) = \frac{1+(1-(1-p)^n)n}{n}$$

Here, with the objective of obtaining individual results, the optimal group size should minimize $T(n)$. For a prevalence of 2%, the average number of tests performed per individual is minimized with groups of 8 individuals ($n = 8$). Each test provides on average the status (infected or not) of 3.7 individuals, a clear productivity gain compared to using only individual tests.

Other methods have been proposed, for instance by Sinnott-Armstrong, Klein, and Hickey (2020). In general, it is possible to further increase the power of group testing beyond the above factor by using more sophisticated protocols involving a higher number of samples. Our objective here is to show the usefulness of group testing, even when it comes to obtaining individual results.

[7] For example, suppose a prevalence of 2%, and assume that an infected member of a couple has an 80% chance of infecting the other member. It is then optimal to make groups of 42 couples or 84 people. Each test then makes it possible, on average, to release 31 people - a much higher performance than with randomly constructed groups.

Conclusion

Mass screening is essential for solving today's health crisis. Its deployment is currently restricted by our production capacities, but these can be boosted tenfold by the group-testing method. The method needs to be optimized and tested on a large scale to reveal its full potential. It is nevertheless promising and justifies substantial investment. The method will then have to be adjusted according to the objective (identifying healthy carriers of the virus, obtaining individual results...) and the target population (prevalence being higher, for example, in certain regions and population groups).

On its own, group testing will not be enough to allow an end to lockdown. It must be supplemented with an adequate combination of personal protective equipment, social distancing and contact tracing. Testing will probably need to be organized on a regular basis, and would probably benefit from being combined with monitoring of interactions, in order to trace contamination chains as quickly as possible.

Reference study

This policy brief is based on the article: "Group Testing against Covid-19", by Christian Gollier and Olivier Gossner, *Covid-Economics*, **2**: 32-42, Avril 2020 .

Authors

Elie Gerschel, Research Assistant at CREST-École Polytechnique, affiliated to Institut des Politiques Publiques.
Christian Gollier, Professor at Toulouse School of Economics, University of Toulouse 1 Capitole.
Olivier Gossner, Directeur de Recherche at CNRS-CREST, Professor at Ecole polytechnique.

Acknowledgments

The authors are grateful to Dr. Marija Backovic from Institut Pasteur for precious expertise on the fields of biochemistry and virology.

References

Dorfman, Robert (Dec. 1943). "The Detection of Defective Members of Large Populations." In: *Ann. Math. Statist.* 14.4, pp. 436–440.

Flaxman, S., S. Mishra, A. Gandy, et al. (Mar. 2020). "Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries." In: *Imperial College preprint*.

INSEE, Institut National de la Statistique et des Études Économiques (Mar. 2020). *Point de conjoncture du 26 mars*. Tech. rep.

Miller, Mark, Cecile Viboud, Marta Balinska, and Lone Simonsen (June 2009). "The Signature Features of Influenza Pandemics - Implications for Policy." In: *The New England journal of medicine* 360, pp. 2595–8.

Sinnott-Armstrong, Nasa, Daniel Klein, and Brendan Hickey (2020). "Evaluation of Group Testing for SARS-CoV-2 RNA." In: *medRxiv*.