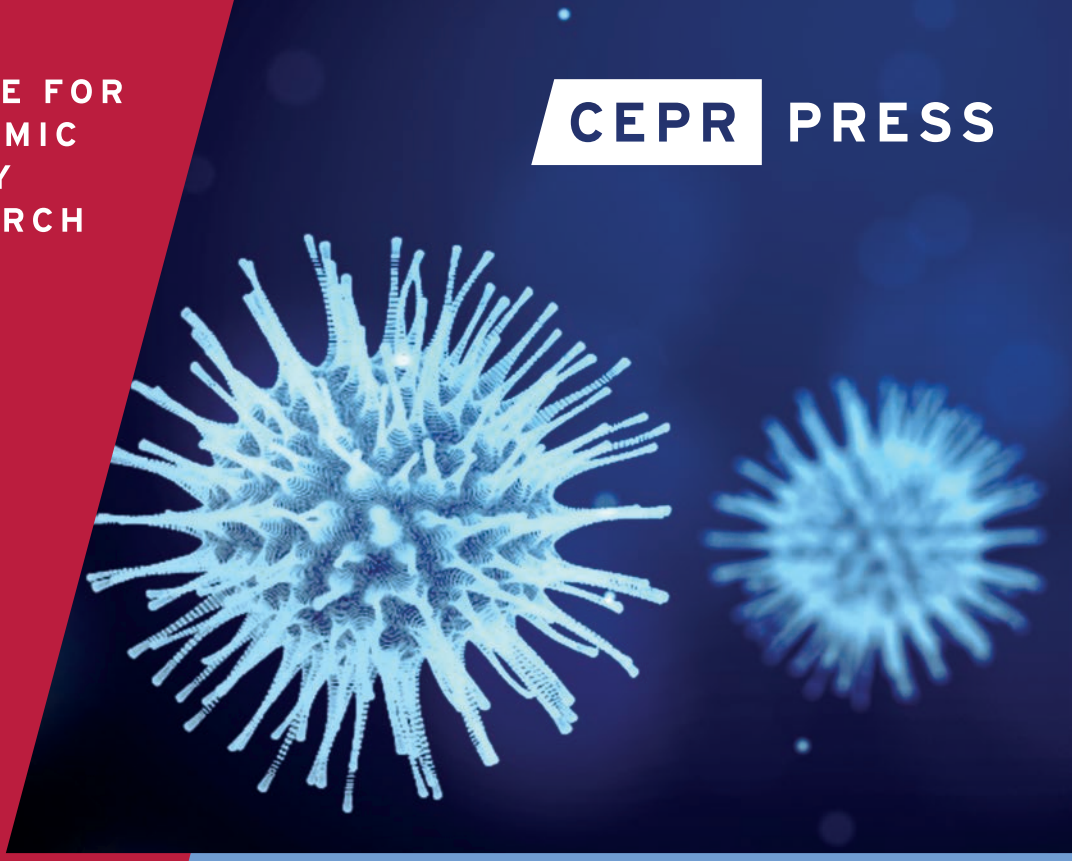
CENTRE FOR ECONOMIC POLICY RESEARCH

CEPR PRESS



COVID ECONOMICS

VETTED AND REAL-TIME PAPERS

ISSUE 2
8 APRIL 2020

SWAP LINES
Saleem Bahaj and Ricardo Reis

SOCIAL DISTANCING AND BUSINESS
Miklós Koren and Rita Peto

GROUP TESTING
Christian Gollier and Olivier Gossner

LOCKDOWNS AND SUPPLY CHAINS
Hiroyasu Inoue and Yasuyuki Todo

WORK VERSUS SAFETY
Tito Boeri, Alessandro Caiumi, and Marco Paccagnella

FISCAL POLICY INSTRUMENTS
Miguel Faria-e-Castro

Group testing against Covid-19[1]

Christian Gollier[2] and Olivier Gossner[3]

Date submitted: 29 March 2020; **Date accepted:** 1 April 2020

It is well-known that group testing is an efficient strategy to screen for the presence of a virus. It consists of pooling n individual samples with a single test using RT-PCR. If no individual in the group is infected, the group test is negative. Thus, a single test may reveal this crucial information. We show how group testing can be optimised in three applications to multiply the power of tests against Covid-19: Estimating virus prevalence to measure the evolution of the pandemic, bringing negative groups back to work to exit the current lockdown, and testing for individual infectious status to treat sick people. For an infection level around 2%, group testing could multiply the power of testing by a factor of 20. The implementation of this strategy in the short run requires limited investments and could bypass the current immense shortage of testing capacity.

1. Introduction

As the coronavirus pandemic develops, governments around the world have now reacted and imposed lockdowns in many countries. Since India imposed strict lockdown restrictions on more than 1.3 billion residents, the total world population under lockdown is now around three billion. By stopping many production processes, the economic cost of the lockdown is very large. For example, Thunstrom et al. (2020) estimate the cost of the lockdown in the US at $7.2 trillion. Finding a way forward is a critical issue. No doubt that the decision to unlock people in the next few weeks or months will be a complex political,

1 The authors are grateful to Marija Backovic, John Cochrane, Romain Gérémi, Mélanie Gollier, Julie Harou, Margarita Kirneva, Larry Kotlikoff, Michael Kotlikoff, Marc M´ezard, Vincent Rollet, David Sraer, Stephane Straub, and Charlotte Wiatroweski as well as participants to the USC workshop "The Economics of the Covid-19 Crisis" for useful comments.

2 Director, Toulouse School of Economics and Professor of Economics, University of Toulouse-Capitole.

3 Director of Research, CNRS — CREST and Professor, Ecole polytechnique and London School of Economics.

health, social, and economic issue. A major risk exists that once the pandemic slows down or appears to be under control and lockdown measures are lifted, new waves of Covid-19 reappear. The 20th century has known three influenza pandemics: the 1918 'Spanish flu', the 1957 'Asian flu', and the 1968 H3N2 'Hong Kong flu'. The 21st century has already witnessed the 2009 'Swine Flu'. These four pandemics came in waves, with subsequent waves being more deadly than the first (Miller et al. 2009).

Therefore, a key element to reduce the economic consequences of Covid-19 is the ability to test individuals, given the large prevalence of asymptomatic but highly contagious people in the population. Massive testing is necessary to monitor the prevalence of the virus in the population in different times and geographical areas. It is also a necessary component to detect infected individuals, quarantine them, and provide medical treatment whenever necessary. Moreover, mass reliable testing would bring back people who have tested negative to work in strategic sectors of the economy, without risking a second wave of contagion. As shown by the experience of South Korea, mass testing is crucial to control the pandemic. As stated by Dewatripont et al. (2020), "restarting production in the economy requires the reliable identification of individuals who will not contract the virus or transmit it to others, whether they have previously displayed the associated symptoms or not."

The standard method for testing for the presence of Covid-19 in a sample is called Real-Time Polymerase Chain Reaction (RT-PCR), which involves a chemical reaction that produces fluorescent light if viral DNA is present. Testing involves two steps - first taking samples from individuals, then amplifying parts of the virus DNA known as markers through a PCR machine. The first step is relatively cheap, but the second one is the bottleneck that limits our testing capacities. Scaling up the capacity of RT-PCR testing for the SARS-COV-2 virus responsible for Covid-19 will take time. It reduces our expectation of a rapid exit from the current lockdown strategy. The US is currently scaling up production up to 1.2 million per week (for a population of 330 million), Germany is producing 500,000 tests per week (population 84 million) and France is producing a mere 84,000 tests per week, scaling up to 210,000 per week in April (population 65 million). Current test production levels are insufficient for mass testing in these countries, not to mention the huge need for tests in developing countries. Each Covid-19 test has to be viewed as a precious resource, to be utilised as efficiently as possible.

In this paper, we exploit a standard testing methodology in which individual samples are pooled.[4] This pooled sample is then tested with a single test. If the test of the combined sample is negative, then all individuals in the group are known to be virus-free, a highly valuable information if the size of the group is large. The implementation of this methodology at the Technion University for Covid-19 suggests that the dilution effect of pooling individual samples is very limited.[5] While individual testing determines whether a given person is a carrier of the virus, group testing will determine whether the virus is present in the group sample or not. Therefore, group testing will be able to reach one of two conclusions: a negative outcome will indicate that none of the individuals of the group is a carrier of the virus, while a positive outcome will indicate that at least one individual in the group is a virus carrier, without any further information on the identity of this person. The optimisation of the group testing strategy depends upon the objective pursued by the test. In this paper, we examine three highly relevant objectives in the context of the Covid-19 pandemic, and we characterise efficient detection strategies to attain them.

2. Applications of group testing

Group testing is not a new idea. It originated in Dorfman (1943) in the context of syphilis detection, but it has also been applied in the case of hepatitis B, avian pneumovirus, and HIV (see for example May et al. 2010). A more advanced mathematical theory of group testing can be found for instance in Mézard et al. (2007) and Mézard et al. (2011). A recent survey is Aldridge et al. (2019). Our paper illustrates three applications of this theory to the problem of fighting Covid-19 in the coming weeks. Group testing can be used for the same purposes as individual testing. However, the protocol needs to be adapted to the situation. We detail below practical applications of group testing and discuss its efficiency in comparison with individual testing.

As we write this article, group testing for Covid-19 has already been implemented in Nebraska[6] and in Israel.

4 See also Jain and Jain (2020).

5 PCR was able to detect the presence of the virus in a pooled sample from 64 individuals with a single infected person. See https://www.technion.ac.il/en/2020/03/pooling-method-for-accelerated-testing-of-Covid-19/ . A team at the University of Frankfurt came to a similar conclusion: https://aktuelles.uni-frankfurt.de/englisch/pool-testing-of-sars-cov-02-samples-increases-worldwide-test-capacities-many-times-over/.

6 https://www.3newsnow.com/news/coronavirus/live-gov-ricketts-provides-coronavirus-briefing-3-24-20

3. Prevalence estimation

There is widespread discussion about the prevalence of the virus in different populations. This information is of crucial importance and will impact policy in many cases. In particular, it allows close monitoring of the spread of the disease. It becomes possible to estimate the ratio of critical cases over total number of cases, as well as the fatality rate, and it allows identification of geographical zones with high infection levels.

The main reason why the information is not well known is the limited availability of tests. Typically, a testing method would involve randomly sampling and testing a group in the population. Relying on hospital admissions is not satisfactory as many cases are either asymptomatic or symptoms are mild enough to recommend prolonged confinement without testing. Here we show how group testing leads to more accurate results with a fewer number of tests (see also Pritchard and Tebbs 2011).

We compare two methods for estimating the prevalence of the virus in the population: (i) individual testing, in which a sample of 12,000 people are tested for the virus, and a standard binomial test is applied to derive a 95% confidence interval, and (ii) group testing, in which 500 groups of 35 people are tested (total population involved 17,500).

3.1 Individual testing

Assume that 2% of people in the sample are infected, returning 240 positive tests.[7] A standard binomial test re- turns the following 95% confidence interval on the infected population:

$$\mathrm{CI}_{IT} = [1.76\%, 2.27\%].$$

3.2 Group testing

Assume again that 2% of individuals in the sampled population are infected, and that individuals are allocated to groups randomly for testing. For each group of 35, there is a probability of $1 - (1 - 0.02)^{35} \sim 50.7\%$ that it contains at least one infected person, meaning the test returns positive. This corresponds to 253 group tests returning positive, and 247 returning negative. With such data, the 95% confidence interval on the proportion of groups of 35 in the population

7 For simplification, the tests are assumed in these applications to return no false positives or negatives.

containing at least one infected person is: [46.1%, 55.1%]. The corresponding confidence interval on the underlying proportion of infected people in the population is:[8]

$$\mathrm{CI}_{GT} = [1.75\%, 2.26\%].$$

3.3 Comparison of results

Both group testing and individual testing return the same point estimate on the proportion of infected individuals (2%). They return slightly different confidence intervals due to a non-linearity in the formulas involved. Both confidence intervals have the same size of 0.5%, which is a reasonable size on which policy making decisions can be based. However, the cost in terms of number of tests is drastically lower for group testing (500) compared to individual testing (12,000). In this application, group testing allows to economise on tests by a factor of 24.

Note that group size 35 is optimised so that each group test positive with probability circa 0.5 for 2% prevalence. In principle, prevalence is not known, so group size may not be chosen optimally. This will lead to a slightly degraded performance of group testing. In this application, group testing allows to economise on tests by a factor 24 while keeping groups of reasonable size.

3.4 Optimal group size

Given a prevalence level p and a number of groups, the variance estimator is minimized for a group size such that the probability q that a group of size n is tested positive satisfies $q \sim -\ln(1-q)/2$, which gives $q \sim 0.80$, and $n \sim \frac{\ln .2}{\ln (1-p)}$.[9] For a prevalence of 2%, groups of size 80 are optimal from the statistical point of view. In practice, technical limitations as well as the cost of collection of individual samples put a downwards pressure on group size.

4. A plan to exit the lockdown

Building testing capacity will take time, even with a wartime mobilisation of means. We therefore propose to complement this investment plan with an immediate expansion of the testing capacity by using group testing. Contrary to Dorfman (1943), we don't attempt in this section to identify infected individuals. We rather determine the size of group testing that maximises the number of individuals whose testing demonstrates they are not infected. The scarcity of tests obviously means that it is better to use a test to detect the virus in another

8 The confidence interval on proportion of infected people is given by $[1-(1-.455)^{\frac{1}{35}}, [1-(1-.455)^{\frac{1}{35}}]$.
9 The authors are grateful to Xavier d'Hautfeuille for this insight.

untested group than to try to discover who is infected in a positive group. This is because the value of information from the test does not come from the treatment of infected people in the absence of an efficient drug to do that. In the context of Covid-19, the value of the test rather comes from sending healthy people back to work as soon as possible, without risking infection.

Suppose that the prevalence rate of the virus in the target population is p. The testing capacity is assumed to be very limited in the sense that even group testing will not allow for testing the entire population. We assume that when a group is detected with the virus, their members remain confined. Let n denote the size of the groups to be tested. If n is too large, too many groups will be detected with the virus, and that will reduce the expected number of people who will be allowed to get back to work. Technically, the frequency of groups tested negative is equal to $(1 - p)^n$, so that the expected number of people freed from confinement with a single test is equal to $n(1 - p)^n$. The optimal size of group testing maximises this function of n. It satisfies the following first-order condition:

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

The optimal size of the group is decreasing with the prevalence ratio. It is optimal that the group size be approximately equal to the inverse of the prevalence ratio. The above equation gives us the following expected number N of people back to work with a single test:

$$N = \frac{(1-p)^{\frac{-1}{\log(1-p)}}}{-\log(1-p)}. \tag{2}$$

The expected number of people freed from confinement with a single test is decreasing in the prevalence ratio. The individual testing strategy with one test allows for freeing an expected number of people equalling $1 - p$. We obtain that the power of the group testing strategy over the individual testing strategy is equal to

$$P = \frac{(1-p)^{\frac{-1}{\log(1-p)}-1}}{-\log(1-p)}. \tag{3}$$

This means that the optimal group testing strategy frees in expectation P times more people from the lockdown than when using the individual testing strategy.

We can also value the benefit of increasing the testing capacity. To do this, we need to measure the social cost q of individual confinement. Suppose that the optimal confinement strategy in the absence of testing is to remain idle for two months. Therefore, we can assume that this social cost equals two months of

GDP per capita. For the EU whose GDP per capita is approximately €31,000 per annum, this corresponds to q = €5,167. The social value of each test is thus equal qN.

4.1 Individual testing

Suppose for example that the prevalence ratio is 2%. Each individual has 98% chances of not being infected and released after testing. Each test allows the release of 0.98 people on average. The value of a single test is thus equal to €5,063.

4.2 Group testing

Consider testing groups of n = 50 people. Each test returns negative if everyone in the group is healthy, which has probability $0.9850 \sim 36\%$. The average number of people each test allows to release is then $N = .36 \times 50 \sim 18.2$. The value of a single test is thus equal to €94,077. Although fewer tests are negative with group testing, each of them allows to release 50 people back to work. Group testing is more efficient than individual testing by a factor P = 18.6.

In Table 1, we describe the characteristics of the optimal strategy for different values of the prevalence ratio, taking account of the integer nature of n. We assumed that the health status is i.i.d. in the target population. In practice, group size must be tailored according to available information on risk prevalence. Also, groups of people may be correlated in their risks of being infected.

Prevalence ratio (p)	**Optimal size (n)**	**Expected number deconfined (N)**	**Power of group testing (P)**	**Expected benefit (qN, in euros)**
0.01	99	36.60	36.97	189 129
0.02	49	18.21	18.58	94 083
0.05	19	7.17	7.55	37 046
0.1	9	3.49	3.87	18 016
0.2	4	1.64	2.05	8 466
0.3	3	1.03	1.47	5 317
0.4	2	0.72	1.20	3 720

Table 1 Optimal group testing strategy

Notes: Optimal group testing strategy as a function of the prevalence rate in the target population. We assume that q = €5167.

Testing positively correlated groups and adjusting group size adequately would increase performance of the system. People working in the same production units, such as production lines or offices, have a high degree of correlation in their infectious statuses. Individual workers also have a high degree of complementarity. In such situations, it is efficient to test a whole production unit as a group and close it when the test returns positive.

5. Testing individuals with group testing

One of the most important applications of testing is to know whether an individual is infected. Group testing can allow for a much more efficient way of testing each individual in a population than individual testing.

Here we present a protocol for testing whether individuals in a population carry the virus, based on sequential group tests. Each individual in the population will be marked as positive ('+'), negative ('-'), or unknown ('?'). Initially everyone is marked as '?'.

Box 1 Testing protocol

T32 Test a group of 32 individuals.

1. If the test is negative, mark all 32 individuals as '-' and the protocol stops
2. If the test is positive, form two subgroups of 16, tagged 16A and 16B T16

Test the group 16A

1. If 16A is positive, mark everyone in 16B as '?', from 16A create two subgroups of 8 individuals, tagged 8A and 8B
2. If 16A is negative, mark everyone in 16A as '-', from 16B create two subgroups of 8 individuals, tagged 8A and 8B

T8 Test the group 8A

1. If 8A is positive, mark everyone in 8B as '?', from 8A create two subgroups of 4 individuals, tagged 4A and 4B
2. If 8A is negative, mark everyone in 8A as '-', from 8B create two subgroups of 8 individuals tagged 4A and 4B

Proceed until a group of 2 individuals is known to hold at least one virus holder.

T1 Test one of the two individuals

1. If the test returns positive, mark this individual '+', the other as '?'.
2. If the test returns negative, mark this individual '-', the other as '+'.

The protocol returns the infectious status of individuals marked '+' or '-'. No information is known about those marked '?' and these individuals re-enter the protocol in newly formed groups of 32.

Estimation of the protocol efficiency

We estimate the average number of tests for each run of the protocol, as well as the average number of individuals for whom the infection status returns as known. For simplification we make the approximation that a group of 32 individuals has probability 50% to contain at least one infected person.

In case the first group is negative, the protocol ends. In case it is positive, it runs tests T32, T16, T8, T4, T2, and T1, hence 6 tests. So on average the protocol runs 7/2 tests.

If the first test is negative, all 32 people's status is returned as known. If the first test is positive, each test TX (X = 16, 8, 4, 2,1) returns either positive or negative with probabilities approximately 1/2. If it returns positive, X people exit the protocol with unknown status at this stage; if it returns negative none exit with unknown status at this stage. Therefore, the average number of people who exit with unknown status is:

$$\frac{1}{2}(\frac{1}{2}16 + \frac{1}{2}8 + \frac{1}{2}4 + \frac{1}{2}2 + \frac{1}{2}1) = \frac{31}{4},$$

So the number of people returning with known status is on average 32 – 31/4 = 97/4 .

Each test therefore returns the status of on average $\frac{97}{4}/\frac{7}{2} \sim 6.9$.

Applying the protocol is tantamount to an increase of test production by a factor of almost seven. Even a factor of three would mean a huge scaling up in world testing capabilities.

5.1 Two-stage protocols

Note that the sequential protocol may require several swabs for a given individual. Given the cost of collecting a swab, including its labour cost, is much smaller than the cost of testing a sample, we find this point essentially non-problematic. In practice, one should probably amend the protocol in order to have a reasonable upper bound on the number of swabs each individual is required to provide.

With only two swabs, both Technion Institute of Technology and Nebraska hospitals have started implementing the original algorithm of Dorfman (1943), which goes as follows:

- Test a group of n individuals
  - If the test is negative, all n individuals are negative
  - If the test is positive, test each individual separately

With a probability p of each individual of being infected, the average number of tests per individual is

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

Given p, we must adjust group size to minimize $T_p(n)$. For $p = 2\%$, we find that $n = 8$ is optimal, using 0.27 tests per individual, thus allowing to find out about 3.65 individual conditions per test used. For $p = 1\%$, $n = 11$ is optimal, allowing to find out about 5.11 individual conditions per test.

In practice, such a simple algorithm is not optimal, but already allows for very significant savings in the number of tests used.

6. Errors and information theory

Abstracting from virus detection, sequential group testing can be viewed as a coding problem. The list of infectious status of all individuals in the population consists of a message, and a sequence of test results read should be enough to recover this message. Information Theory (Shannon 1948, Cover and Thomas 2006) tells us that a lower bound on the number of tests required per individual in the population is:

$$h/C$$

where

- C is known as the capacity of the channel and depends on the test accuracy. A perfect test returning the infectious status of the patient (positive or negative) with no errors has a capacity of 1. Tests with lower accuracy also have lower capacities, and
- h is the entropy per individual in the population. In the case of an i.i.d. population with prevalence p, $h = H(p) = -p \log_2(p) (1 - p) \log_2(1 - p)$. When $p = 2\%$, $h \sim 0.112$. Assuming a test with no errors, the theoretical bound on the number of tests per individual is then $1/0.1414 \sim 7.1$, showing that the protocol suggested above achieves near-optimality.

7. Conclusion

Testing for Covid-19 is a bottleneck that we face in front of the pandemic. Test production is currently much below what is necessary for mass testing strategies which are required in order to control the pandemic while letting people go back to work. Adequate use of group testing can save many tests, between 85% and 95% depending on the applications. Although this work is of theoretical nature and does not account for many technical details of group testing such as maximal group sizes and error types, a very conservative assessment of the tests that can be saved in this application is about two-thirds, which means that use of group testing is equivalent to a scaling up of test production by a factor of three or more.

In this paper, we focused our attention to RT-PCR tests that are able to detect infection. Alternatively, serological tests are used to detect the presence of antibodies, thus the immunity of the individual. In the absence of a vaccine, it is an urgent strategic issue to detect immunity in the most essential professions, and group testing should also be used for this purpose.

References

Aldridge, M, O Johnson and J Scarlett (2019), "Group Testing: An Information Theory Perspective", *Foundations and Trends® in Communications and Information Theory* 15(3-4): 196-392.

Cover, T M and J A Thomas (2006), *Elements of Information Theory*, New York: Wiley Interscience, 2nd edition.

Dewatripont, M, M Goldman, E Muraille and J P Platteau (2020), "Rapidly identifying workers who are immune to covid-19 and virus-free is a priority for restarting the economy", VoxEU.org, 23 March.

Dorfman, R (1943), "The detection of defective members of large populations", *The Annals of Mathematical Statistics* 14(4): 436-440.

Jain, T and B N Jain (2020), "Accelerated infection testing at scale: A proposal for inference with single test on multiple patients", Working paper.

May, S, A Gamst, R Haubrich, C Benson and D M Smith, (2010), "Pooled nucleic acid testing to identify antiretroviral treatment failure during HIV infection", *Journal of acquired immune deficiency syndromes* 53(2): 194-201.

M´ezard, M, M Tarzia and C Toninelli (2007), "Group testing with random pools: Phase transitions and optimal strategy", Technical Report 0711.2242v1, Arxiv.

M´ezard, M and C Toninelli (2011), "Group testing with random pools: Optimal two-stage algorithms", *IEEE Transaction on Information Theory* 57(3): 1736-1745.

Miller, M, C Viboud, M Balinska and L Simonsen (2009), "The signature features of influenza pandemics – Implications for policy", *New England Journal of Medicine* 360(25): 2595-2598.

Pritchard, N and J Tebbs (2011), "Estimating disease prevalence using inverse binomial pooled testing", *Journal of agricultural, biological, and environmental statistics* 16(1):70–87.

Shannon, C E (1948), "A mathematical theory of communication", *The Bell System Technical Journal* 27:3–55.

Thunstrom, L, S Newbold, D Finnoff, M Ashworth, and J F Shogren (2020), "The benefits and costs of flattening the curve for covid-19", Working paper.