Group Testing against COVID-19

Olivier Gossner*

March 29, 2020

Abstract

We show how group testing can be used in three applications to multiply the efficiency of tests against COVID-19: estimating virus prevalence, releasing group to the work force, and testing for individual infectious status. For an infection level around 2%, group testing could potentially allow to save 94% of tests in the first application, 95% in the second, and 85% in the third one.

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 that 1.3 Billion residents, the total world population under lockdown is now around 3 Billion.

The French government recently estimated the economic slowdown due to lockdown measures to 33% of total production value [10]. Based on this estimate, each month under lockdown amounts to an economic loss of about 2.5% of yearly GDP. This is not sustainable for long, and all governments will have to develop exit strategies very quickly.

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", and the 21st century has already witnessed 2009 "Swine Flu". These four pandemics came in waves, with subsequent waves being more deadly than the first [7].

*CNRS – CREST, Ecole polytechnique, and London School of Economics. The author is grateful to Fabien Ferrage, Christian Gollier, Stephane Straub as well as participants on the USC workshop "The Economics of the Covid-19 Crisis" for useful comments.

A successful exit strategy will have to come as quickly as possible to restart economic production, but must contain the current epidemic wave and should detect foyers of future waves as quickly and efficiently as possible. For this, large scale testing of populations is a necessary component. Massive testing is necessary both to monitor the prevalence of the virus in the population in different period of times and geographical areas. It is also a necessary component to detect infected individuals, quarantine them and provide medical treatment whenever necessary.

Testing for the SARS-COV-2 virus responsible for the COVID-19 is costly and scaling up production of tests takes time. The USA 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 (pop 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.

Each test allows to detect presence of the virus in one sample. This sample can be coming from one individual (individual testing), or from a group of individuals (group testing). In this later case, individual samples are combined to create a group sample, and the test will detect presence of the virus in the group sample. While individual testing allows to determine 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 or the virus ; 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.

2 Applications of Group Testing

Group testing can be used for the same purposes as individual testing is. Only the protocol needs to be adapted to the situation. We detail below two practical applications of group testing, and discuss its efficiency in comparison with individual testing. Group testing is not a new idea, it originated in [3] in the context of syphilis detection. A more advanced mathematical theory of group testing can be found for instance in [5], and [1] provide a recent survey on the topic.

2.1 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 to closely monitor the spread of the disease, it also allows to estimate the ratio of critical cases over total number of cases, as well as the fatality rate, and it allows to identify 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 lesser number of tests (cf. also [8]).

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

Individual Testing. Assume that 2% of people in the sample are infected, returning 240 positive tests[1]. A standard binomial test returns the following 95% confidence interval on the infected population:

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

Group Testing. Assume here too that 2% of individuals in the sampled population are infected, and that individuals are allocated to groups randomly for testing. Each group of 35 has a probability $1-(1-0.02)^{35} \sim 50.07\%$ to contain one infected person, hence to return positive. This corresponds to 253 group tests returning positive, and 250 returning negative. With such data, the 95% confidence interval on the proportion of groups of 35 in the population containing at least one infected person is: $[45.5\%, 54.5\%]$. The corresponding confidence interval on the underlying proportion of infected people in the population is[2]:

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

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

[2] The confidence interval on proportion of infected people is given by $[1-(1-.455)^{\frac{1}{35}}, 1-(1-.545)^{\frac{1}{35}}]$.

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 if 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 24.

Note that group size 35 is optimised so that each group test positive with probability circa .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.

3 Releasing individuals to the workforce

As the economy could tank we cannot resume production soon, it is important to release as many people as quickly as possible to the workforce. Doing so without risking of spreading the disease through the release of infectious people means that only individuals who are known not to be virus carriers should be released. Our objective function should then to detect as many non-carriers as possible, using as few tests as possible. Here again, we compare individual testing with group testing, assuming that 2% of the population is infected. A more detailed of group testing for releasing individuals was developed simultaneously in [4].

Individual Testing. Each individual has 98% chances of not being infected and released after testing. Each test allows the release of .98 people on average.

Group Testing. Consider testing groups of 50 people. Each test returns negative if everyone in the group is healthy, which has probability $.98^{50} \sim 36\%$. The average number of people each test allows to release is then $.36 \times 50 \sim 18.2$.

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 more than 18.

The paper [4] presents an optimisation of group size in this context, and finds the optimal group size is $1/p$, where p is the prevalence of the disease. In practice, group size has to be tailored according to

available information on risk prevalence. Also, groups of people may be correlated in their risks of being infected.

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.

4 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 of in the population will be market as positive ("+"), negative ("-"), or unknown ("?"). Initially everyone is marked as "?".

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:

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

so the number of people returning with known status is on average $32 - 31/4 = 97/4$.

Each test therefore allows to return the status of $\frac{97}{4}/\frac{7}{2}$ 6.9.

Applying the protocol is tantamount to an increase of test production by a factor almost 7. Even a factor 3 would mean a huge scaling up in world testing capabilities.

Note that the test may require that several swabs are used for a given individual. Given the cost of collecting swabs is much smaller compared to 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 2 stages of testing allowed only, an optimal algorithm has been proposed by [6].

5 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 ([9] [2]) 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,
- is the entropy per individual in the population. In the case of an iid 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 $1/.112 \sim 8.9$, showing that the protocol suggested above achieves near-optimality.

6 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 practicalities 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 3 or more.

References

[1] Matthew Aldridge, Oliver Johnson, Jonathan Scarlett, et al. Group testing: an information theory perspective. *Foundations and Trends® in Communications and Information Theory*, 15(3-4):196–392, 2019.

[2] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley Interscience, New York, 2nd edition, 2006.

[3] R. Dorfman. The detection of defective members of large populations. *The Annals of Mathematical Statistics*, 14(4):436–440, 1943.

[4] Christian Gollier. Optimal group testing to exit the covid confinement. Technical report, Toulouse School of Economics, March 2020.

[5] M. Mézard, M. Tarzia, and C. Toninelli. Group testing with random pools: Phase transitions and optimal strategy. Technical Report 0711.2242v1, Arxiv, 2007.

[6] Marc Mézard and Cristina Toninelli. Group testing with random pools: Optimal two-stage algorithms. *IEEE Transactions on Information Theory*, 57(3):1736–1745, 2011.

[7] Mark A Miller, Cecile Viboud, Marta Balinska, and Lone Simonsen. The signature features of influenza pandemics—implications for policy. *New England Journal of Medicine*, 360(25):2595–2598, 2009.

[8] Nicholas A Pritchard and Joshua M Tebbs. Estimating disease prevalence using inverse binomial pooled testing. *Journal of agricultural, biological, and environmental statistics*, 16(1):70–87, 2011.

[9] Claude Elwood Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:3–55, 1948.

[10] Jean-Luc Tavernier. Point de conjoncture du 26 mars 2020. Technical report, INSEE, 2020.