

Focus mass testing: how to overcome low test accuracy

EsadeEcPol Insight #25 December 2020

AUTHORS

Miquel Oliu-Barton,
associate professor at
Paris-Dauphine University

Bary Pradelski,
associate professor at the
CNRS, associate member
of the Oxford-Man Institute

EXECUTIVE SUMMARY

A third wave of infections is looming in much of Europe and, although vaccination against Covid-19 will begin in the coming weeks, its implementation will not be as fast as we would like. This forces us to consider other strategies for the coming months.

- One potential relief can come from the rapid antigen tests that are now readily available. These are inexpensive and provide a test result within a few minutes. However, they are imprecise. **We propose focus testing, a simple and highly effective strategy designed to detect and prevent outbreaks.**
- Focus testing aims to **minimise the occurrence of false negatives and false positives** by targeted re-testing:
 - Round 1. Test everyone once: if the test is negative, notify ‘inconclusive’, otherwise require two more tests.
 - Round 2. Re-test twice those who previously tested positive: if both new tests are negative, notify ‘inconclusive’; if they are both positive, notify ‘positive’; otherwise, require a final round.
 - Round 3. Re-test once more those who previously tested positive-negative: if the test is positive notify ‘positive’, otherwise ‘inconclusive’.
- Our case study for Spain finds that **focus testing can increase the positive predictive value 30-fold** (from 2.6% to 90%) in areas with low virus incidence (when, like currently in Spain, one in a thousand is a virus carrier).
- **This improvement is achieved with only 10% more tests than in standard mass testing.** In addition, we consider *double testing*, an alternative strategy that minimises false negatives. This strategy is particularly suited to “red” areas, that is, small areas with high incidence (cities, districts, or provinces).

Research area:

COVID - 19

Directed by Jorge Galindo

Technical vocabulary

	Virus carrier	Virus free
Test positive	True positive <i>tp</i>	False positive <i>fp</i>
Test negative	False negative <i>fn</i>	True negative <i>tn</i>
	sensitivity = $\frac{tp}{tp + fn}$	specificity = $\frac{tn}{fp + tn}$

Preliminary remarks

1. Sensitivity and specificity are inherent to the test, while PPV and FNR depend on the testing strategy.

2. Individuals who have had Covid-19 during the last 6 months (confirmed by a positive PCR or serological test), should be exempt from the mass testing campaign, given their low chance of contracting the virus again.

Prevalence: proportion of virus carriers among the population.

True positive: virus carriers that are tested positive.

False positive: virus free individuals that are tested positive.

True negative: virus free individuals that are tested negative.

False negative: virus carriers tested negative.

Sensitivity: proportion of true positives (among all positives).

Specificity: proportion of true negatives (among all negatives).

PPV: positive predictive value, that is, the probability of carrying the virus when classified as positive.

FNR: false negative rate, that is, the probability of being classified as negative while being infected.

PPV: positive predictive value, that is, the proportion of virus carriers among those tested positive.

FNR: false negative rate, that is, the proportion of false negatives among all virus carriers.

PCR test: molecular testing technology which remains the reference for identifying active infections. These tests are very reliable (high sensitivity and specificity), but capacity constraints and their relatively high cost of PCR tests limit their feasibility on a large scale. Further, they usually require 24h-48h to produce results.

Rapid Antigen Test: alternative testing technology which is simple to use, can be performed at point-of-care, and are cheaper than molecular tests, allowing their use at a very large scale. However, these tests are less reliable than molecular tests – they achieve good specificity but only moderate sensitivity.

Serological test: also known as an antibody test, is a blood test that can detect if a person has antibodies to SARS-CoV-2, the virus that causes COVID-19.

Introduction

The race for a [Covid-19 vaccine](#) has led to several promising candidates in only a few months, some of which have already passed regulatory approval. But unfortunately, vaccination does not simply and instantaneously wipe out the virus. Instead, it requires time and thus patience, before we can declare the defeat of the pandemic.

- Firstly, the currently available vaccines require two inoculations and maximum efficacy is only reached a couple of weeks after the second shot.
- Secondly, the vaccine rollout will take several months due to production and distribution constraints. In other words, we still need to evaluate how to live with the virus in the months to come.

The idea of mass testing has been discussed since the beginning of the Covid-19 pandemic but was deemed infeasible due to the limited number of PCR tests. The arrival of rapid antigen tests radically changes the situation. Cheap, widely available and analysed at point-of-care, antigen tests give a result within 15 minutes and seem ideal to test massive populations. Slovakia has already used this option, as well as parts of Spain, Austria and the UK; and other countries, such as France, are evaluating these tests.

A more elaborate mass testing strategy can be the key to managing the pandemic in the months to come, and possibly avoid a third wave

There are mainly two reasons for hesitancy among policy makers and researchers before **implementing a mass antigen testing strategy**. Firstly, given the antigen test's low accuracy, the overall benefit for controlling the pandemic is unclear. Secondly, the high number of false negatives and false positives is worrisome as it may trigger mistrust and unintended behavioural changes.

To overcome problems stemming from the low accuracy of antigen tests, alternative test strategies need to be employed. Testing everyone once can be counterproductive due to the large number of erroneous results. But a more elaborate mass testing strategy can be the key to managing the pandemic in the months to come, and possibly avoiding a third wave.

We introduce **focus testing**, an effective mass testing strategy that can be employed at a national level. Designed to minimise the number of erroneous test results – given the current prevalence of the virus in Europe and the relatively low accuracy of available antigen tests –, this strategy requires a very limited number of additional tests.

Focus testing aims to minimise the occurrence of false negatives and false positives by targeted re-testing

In addition, we discuss double testing, an alternative strategy that relies on substantially more tests – nearly double – but may be suitable for red zones, that is, smaller high-prevalence areas (cities, counties or provinces).

Measuring the success of a test

The advantages and drawbacks of mass testing crucially depend on two factors: the prevalence of the virus in the zone, and the accuracy of the test, measured by its sensitivity and specificity.

The **prevalence** is the proportion of virus carriers among the population – usually considerably larger than the proportion of detected cases. The **sensitivity** is the proportion of positive cases that are correctly identified (or true positives), while the **specificity** is the proportion of true negatives. There are several tests currently on the market and sensitivity and specificity vary greatly, both between products and depending on whether they are measured in laboratory conditions or in the field.

No test is perfect, and as a consequence, some virus free individuals will be tested positive (false positives), while some virus carriers will be tested negative (false negatives). For this reason, the careful interpretation of results, negative externalities of faulty test results, and subsequent measures and recommendations are crucial (Watson et al., 2020; Brooks and Das, 2020).

False positives and false negatives: behavioural considerations

The presence of false positives and false negatives has several undesirable consequences. Concerning the former, an erroneous positive test result leads to inappropriate and costly isolation. Even worse, after the isolation period, false positive individuals may believe that they are immune, and thus put themselves and others at increased risk (given their information, their actions are completely rational). Further, false positives undermine the ‘seriousness of the test’: if, for example, more than half of the people tested positive are in fact virus free, then isolation after a positive test may be perceived as unacceptable.



A testing strategy should aim to minimise the occurrence of false negatives and false positives (Photo: Giuseppe Lombardo/Getty Images)

Similarly, false negatives may give an erroneous feeling of security, inducing a temporary disregard of social distancing measures and thus putting others at increased risk (Kumleben *et al.*, 2020). To counter this effect, it is of particular importance to communicate negative results carefully. For instance, the term 'negative' could be replaced by 'inconclusive' with an according explanation (an individual might not carry enough virus cells at the time of the test to be detectable, and still be infected or infect themselves in the future).

Thus, a testing strategy should aim to minimise the occurrence of false negatives and false positives. Standard measures are the positive predictive value (PPV), that is, the proportion of virus carriers among those tested positive, and the false negative rate (FNR), that is, the proportion of virus carriers who falsely tested negative.

Illustration of PPV and FNR as prevalence varies

To familiarise ourselves with the PPV and FNR measures let us consider a rapid antigen test whose accuracy, under laboratory conditions, meets the minimum requirement set by the WHO and agreed by the European Centre for Disease Prevention and Control (ECDC), that is 80% sensitivity and 97% specificity. Suppose every individual in a city of 100,000 is tested once. The FNR is then 20%, notably 100% minus the sensitivity, and does not depend on the prevalence level. In contrast, the PPV critically depends on prevalence. Consider two different scenarios: low prevalence (0.1%), and high prevalence (2%).

In the former, we expect 3,077 positive results, out of which 80 are true positives, and so PPV is 2.6%. In the latter, we expect 4,540 positive results, out of which 1,600 are true positives, so the PPV is 35.2%. Such low PPVs are worrisome, and naturally call for additional testing efforts [1].

Comparative study of several testing strategies

To overcome the problem of a low PPV when testing everyone only once, the natural idea is to test (at least) some individuals multiple times. Using a PCR test for the subsequent round is infeasible for mass testing due to the limited capacity of laboratories for evaluating PCR tests. Further, this technology requires 24h-48h to produce and communicate results, or more in case of laboratory saturation. This is an excessive time lag for breaking the chain of infections. However, fast antigen tests can be performed at a population-wide scale and produce results within a few minutes. Thus, we focus our analysis on antigen testing strategies.

We now introduce several testing strategies and compare them in terms of performance (PPV and FNR) and feasibility (number of tests required). In the sequel, prevalence will range between 0.1% and 2%, as is the case in most European regions [2].

Regarding test accuracies, sensitivity will range from 70% to 100% and specificity from 95% to 100%. Two focal points will be highlighted: (1) 80% sensitivity and 97% specificity – the minimum requirement by the WHO and ECDC (WHO 2020); (2) 86% sensitivity and 99% specificity – an available antigen test by EasyCOV, which was recently authorised for commercialisation by the French health authority (HAS). It is important to note that the levels given for the latter are under lab conditions and may decrease in the field (Surkova *et al.*, 2020).

“ Double testing requires substantially more tests but may be suitable for smaller high-prevalence zones

The strategies will be referred to as ‘multiple-round testing’ and ‘focus testing’. We will also discuss ‘**double testing**’, a strategy that requires substantially more tests but may be suitable for smaller high-prevalence zones, i.e., hotspots.

A **multiple-round strategy** is one where initially everyone is tested once, then those who have tested positive are retested. This can be repeated for multiple rounds [3]. At the end, only those tested positive in every round are notified as ‘positive’, while all others are notified as ‘inconclusive’. For reasons that will become apparent later, we only consider one, two, and three rounds.

Focus testing is a modified multiple-round strategy, and involves re-testing small groups of the population based on the previous outcomes. While most people are tested once, some individuals are tested three or four times. It can be described as follows:

- **Round 1. Test everyone once:** if the test is negative, notify ‘inconclusive’, otherwise require two more tests.
- **Round 2. Test twice those tested positive in Round 1.** If both new tests are negative, notify ‘inconclusive’. If both new tests are positive, notify ‘positive’. Otherwise, require a third round.
- **Round 3. Test once more those tested positive-negative in Round 2.** If the test is positive notify ‘positive’, otherwise ‘inconclusive’.

Finally, **double testing** involves initially testing everyone twice, then retesting those who tested positive-negative in the first round.

The following decision trees illustrate focus and double testing. (Note that the latter corresponds to rounds 2 and 3 of the former, but starting with the entire population.)



Evaluation

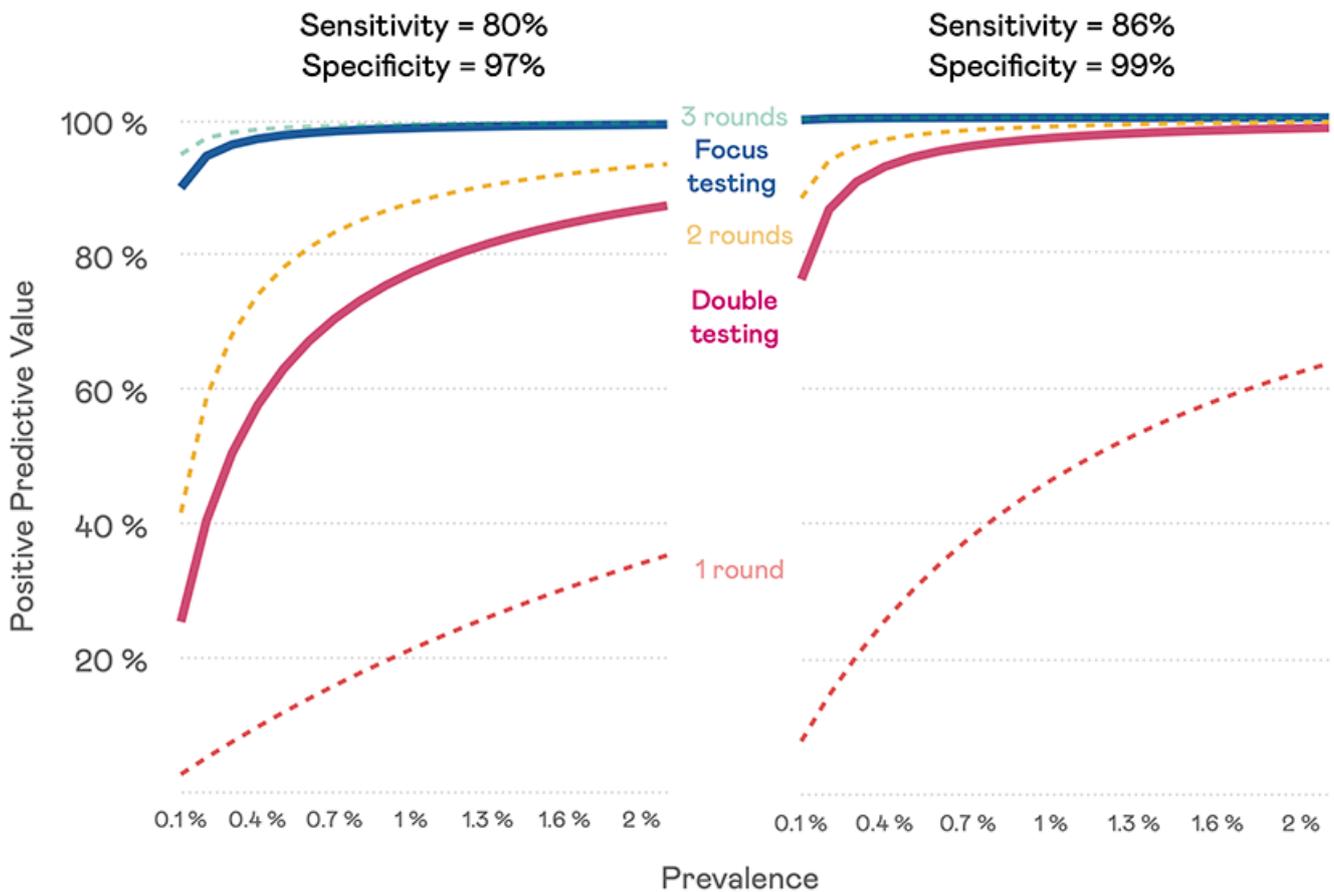
Testing everyone once would outperform all other strategies in an idealised world, where tests have no error (that is, when sensitivity and specificity are 100%). However, as we will see, an imperfect accuracy produces too many erroneous tests, especially in zones with low prevalence.

From a theoretical point of view the obvious way to increase the reliability of the test (that is, increase PPV, and lower FNR at the same time) is to administer multiple tests to each individual. This idea was voiced early on in the Covid-19 pandemic (Ramdas *et al.*, 2020). (See also Lau (1989) for a discussion on how to ensure any desired level of PPV and FNR by this ‘brute force strategy’.)

“An imperfect accuracy leads to too many erroneous tests, especially in zones with low prevalence

Nevertheless, this route does not seem feasible as testing the entire population several times would require substantially more tests than available. The multiple-round testing strategies take into account the scarcity of antigen tests and personnel to carry them out, as these strategies require only a small additional testing effort (less than a 10% increase, whereas double testing would entail a 107% increase). As we will see, the problem with multiple rounds is the stark increase in the FNR in each round.

The following figure depicts the PPV as a function of the prevalence of the virus (from 0.1% to 2.0%) for the different strategies. The left-hand panel is for the minimal accuracy required by the WHO and ECDC, and the right-hand panel for the EasyCOV.

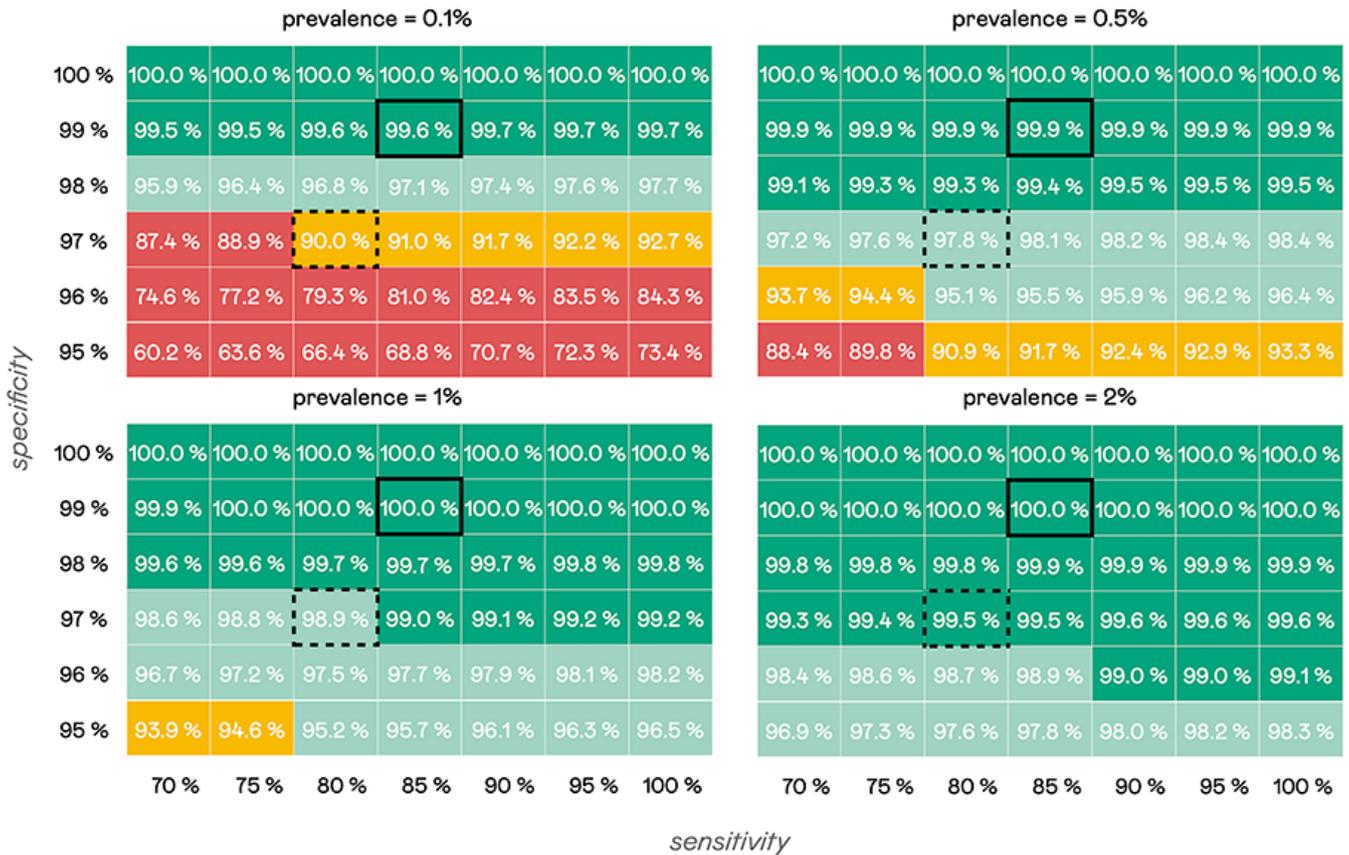


Among all the considered strategies, testing everyone once requires the least number of tests – one per individual. Further, its FNR is relatively low, namely 100% minus the sensitivity. However, this strategy is problematic in terms of its PPV, most notably when prevalence is low. The multiple-round strategies increase the PPV at each round, reaching excellent values after only three rounds. Further, this is achieved with little additional testing effort. However, as the following table shows, the problem with these strategies is the FNR, which increases considerably with each round.

The table also shows the **number of tests required by each strategy**: the range corresponds to different prevalence levels (at a higher prevalence, more people are tested positive, and thus more tests are required).

		One round	Two rounds	Three rounds	Focus testing	Double testing
WHO reference test	FNR	20.0%	36.0%	48.8%	28.3%	10.4%
	# tests/ population	100%	103-105%	103-106%	107-110%	206-207%
EasyCov test	FNR	14.0%	26.0%	36.4%	18.6%	5.3%
	# tests/ population	100%	101-103%	101-104%	102-106%	202-203%

Focus testing offers a very high PPV (comparable to the three-round strategy), while requiring a limited number of additional tests (no more than 10%). Further, **by performing two tests in the second round, the proportion of false negatives is considerably reduced compared to multiple-round strategies.** The following figure illustrates the robustness of the focus strategy, as it remains efficient for different levels of test accuracy and prevalence. (The highlighted boxes in each plot correspond to the two references tests.)



Note that within the ranges considered – under which most of the antigen tests currently available fall – the PPV value is more reactive to changes in specificity than changes in sensitivity. Further, in zones with low virus prevalence (0.1%) it is unlikely that current tests can provide satisfactory PPV levels. On the other hand, in zones where the virus circulates actively the resulting PPV is above 99%.

“ Focus testing offers a very high PPV, while requiring a limited number of additional tests

A possible critique of focus testing is that it may be too complicated to be effectively communicated and implemented. If this is the case, a slightly simpler alternative could be considered: **initially test everyone once, then retest three times those who tested positive**. Notify as ‘positive’ those with overall three or four positive tests, otherwise ‘inconclusive’. This strategy, which is obtained by merely merging rounds 2 and 3 together, yields the same PPV and FNR as focus testing, at very little additional cost.

Another issue is the FNR. Here, double testing is preferable, since by performing an initial round with two tests, this strategy obtains a huge reduction in false negatives. More precisely, with 107% additional tests, the FNR is reduced by 13-18 percent points compared to focus testing. This strategy is particularly suited for zones with a high prevalence to minimise the number of false negatives.



By performing an initial round with two tests, double testing obtains a huge reduction of false negatives

Finally, small quantities of a more accurate point of care test (e.g., Covid-19 PCR tests with a 90-minute turn-around [4]) could provide a comparable performance to focus testing: test initially everyone once, then retest with the more accurate test those who tested positive. If the two tests are positive notify as 'positive', otherwise as 'inconclusive'.

Case study: mass testing at a large scale (country or province)

In Spain (population 47 million), the incidence over the past 7 days is 0.1% as of 8 December 2020. Prevalence, however, is generally higher than the incidence, as the latter accounts for detected cases only. Thus, considering a range between 0.1% (lower bound) and 2% (upper bound) accounts for undetected cases.

Suppose that everybody above the age of 11 is asked to take part in a mass testing campaign, and that around 75% of the eligible population opts in [5]. In this case, around 50 million people will participate in the campaign. The following table illustrates the results for a low and a high prevalence, using a test with 80% sensitivity and 97% specificity.

	Prevalence = 0.1%				Prevalence = 2%			
	One round	Two rounds	Three rounds	Focus testing	One round	Two rounds	Three rounds	Focus testing
True positive (tp)	24,000	19,200	15,360	21,504	480,000	384,000	307,200	430,080
True negative (tn)	29,070,900	29,943,027	29,969,191	29,967,858	28,518,000	29,373,540	29,399,206	29,397,679
False positive (fp)	899,100	26,973	809	2,142	882,000	26,460	794	2,321
False negative (fn)	6,000	10,800	14,640	8,496	120,000	216,000	292,800	169,920
PPV (=tp/[tp+fp])	2.6%	41.6%	95.5%	90.9%	35.2%	93.6%	99.7%	99.5%
FNF (=fn/[tp+fn])	20.0%	36.0%	48.8%	28.3%	20.0%	36.0%	48.8%	28.3%
# tests/ population	100%	103%	103%	107%	100%	105%	106%	110%

Focus testing is preferable to three rounds, as it identifies significantly more virus carriers. Note that the number of false positives critically depends on the testing strategy used, but it is little altered by prevalence. One round gives around 900,000 false positives: while two rounds, three rounds, and focus testing yield, respectively, around 26,500,800 and 2,200 false positives (regardless of the prevalence). We shall thus focus on three rounds and focus testing strategies. (Recall that these strategies require at most 10% additional tests compared to testing everyone once, as shown in the previous table.)

Comparing the FNRs, it becomes clear that focus testing is preferable to three rounds, as it identifies significantly more virus carriers (+40%). Indeed, with a prevalence of 0.1%, the three-round strategy finds 15,360 true positives (and misses 14,640 false negatives), while focus testing finds 21,504 (and misses 8,496 false negatives). Similarly, with a prevalence of 2%, the three-round strategy finds 307,200 true positives (and misses 292,800 false negatives), while focus testing finds 430,080 (and misses 169,920).

Case study: mass testing for a high prevalence city, county or province

Consider a city with a particularly high prevalence, say 2%, and where 100,000 are eligible and opt in for mass testing. The following table illustrates the results using a test with 80% sensitivity and 97% specificity for testing everyone once (or 1 round), focus testing, and double testing.

	One round	Focus testing	Double testing
True positive (tp)	1,600	1,434	1,792
True negative (tn)	95,060	97,992	97,741
False positive (fp)	2,940	8	259
False negative (fn)	400	566	208
PPV (=tp/[tp+fp])	35.2%	99.5%	87.4%
FNR (=fn/[tp+fn])	20.0%	28.3%	10.4%
# tests/ population	100%	110%	207%

Focus testing results in a much higher PPV than testing everyone once with only 10% more tests (8 false positives versus nearly 3,000). It also outperforms double testing in terms of PPV, with nearly half as many tests. However, the situation is reversed in terms of FNR. Double testing gives a much lower FNR (10.4% compared to 28.3% for focus testing, or 20% for testing everyone once), but of course this comes at a higher cost of testing as everyone is tested at least twice. This case study illustrates the stark trade-off between PPV and FNR.

Additional considerations

We briefly discuss here additional considerations when designing a mass testing strategy, but which are beyond the scope of this study.

- **Incubation period.** Any test can only detect the virus in an individual's nose or saliva once the body has reproduced the virus in sufficient loads. For the coronavirus the incubation period before an antigen test can detect the virus is at least three days. Thus, for an effective mass test, the whole population twice within a period of 5-7 days should be considered.
- **Test availability.** There may not be sufficient tests for country-wide mass testing, especially in larger countries. In this case, mass-testing could be carried out sequentially within regions or provinces of the country. To render this effective, the [zoning strategy](#) now used in many countries should be in place. Furthermore, as already discussed, different testing strategies might be appropriate for zones with different prevalence levels.
- **Participation incentives.** Participation in mass testing will differ within varying population groups and depend on the incentive structure. For example, to increase participation, a conditional obligation could be considered. That is, testing could be made obligatory for school attendance or using public transport, for example.

Conclusion

Mass testing can provide relief from the pandemic, but only if performed with care. There are several important considerations: Firstly, the accuracy of the test, meaning its sensitivity and specificity, is crucial. Secondly, even with a widely available accurate test, a well-designed strategy needs to be employed to keep the proportion of false positives and false negatives low (that is, high PPV and a low FNR). Indeed, testing everybody once can be counterproductive as it would yield a PPV of between 3% and 35% (with the WHO benchmark test), meaning that the vast majority of individuals tested positive are not infected with the virus.

“Mass testing can provide relief from the pandemic, but only if performed with care

On a large scale, **focus testing is promising as it provides a substantially higher PPV of over 90%**, that is, more than nine in ten people with a positive test are indeed virus carriers. In addition, focus testing does not compromise the number of false negatives (as multi-round testing does), it requires no more than 10% additional tests, and it works well for different levels of prevalence.

For smaller zones with a particularly high prevalence, **double testing can be preferable and feasible**. It reduces the number of undetected cases, while reaching acceptable levels of false positives.

Finally, the combination of focus testing in zones with low prevalence, and double testing in high-prevalence zones could be considered, as this could further reduce the occurrence of false positive and false negatives.

To conclude, any testing strategy should consider the behavioural change triggered by a positive or inconclusive test – as this may lead to unintended negative consequences.

NOTAS

- [1] It is worth noting, however, that the situation radically changes if only people with symptoms are tested. Suppose the prevalence among this group is 20%, then the PPV increases to 87%.
- [2] The incidence over the past 7 days lies between 0.05% and 0.5% in most European regions according to the ECDC (as of 8 December 2020). Thus, by considering the range between 0.1% and 2% accounts for undetected cases.
- [3] When performing several tests on a single individual, we assume that the results of each test are independent.
- [4] Recent results for Covid-19 suggest that such tests may have a sensitivity of 94% and a specificity of close to 100% (Mahase 2020).
- [5] In Slovakia, three mass testing campaigns have already taken place, and the participation rate has been of 85% (Pavelka et al., 2020). The incentives to participate were pretty strong, as only people presenting a negative result were allowed to go out.



ACKNOWLEDGEMENTS

The authors thank Jaume Izquierdo for his suggestions regarding the testing strategy, and Luis Garicano and Toni Roldan for insightful policy considerations. This note will be concurrently published in France by Terra Nova.

REFERENCES

- Brooks, ZC and S Das (2020). [Covid-19 testing: Impact of prevalence, sensitivity, and specificity on patient risk and cost](#), *American Journal of Clinical Pathology*, 154(5), 575-584.
- Kumleben, N, R Bhopal, T Czypionka, L Gruer, R Kock, J Stebbing, FL Stigler (2020). [Test, test, test for Covid-19 antibodies: the importance of sensitivity, specificity and predictive powers](#), *Public Health*, 185, 88-90.
- Lau, TS (1989). [On repeated screening tests](#), *Biometrics*, 45, 891-898.
- Mahase E (2020). [Covid-19: Point of care test reports 94% sensitivity and 100% specificity compared with laboratory test](#), *BMJ*, 370:m3682.
- Pavelka M, K Van-Zandvoort, S Abbott, K Sherratt, M Majdan, CMMID Covid-19 working group, Inštitút Zdravotných Analýz, P Jarčuška, M Krajčí, S Flasche, S Funk (2020). [The effectiveness of population-wide, rapid antigen test based screening in reducing SARS-CoV-2 infection prevalence in Slovakia](#), *medRxiv*.
- Ramdas, K, A Darzi and S Jain (2020). [‘Test, re-test, re-test’: using inaccurate tests to greatly increase the accuracy of Covid-19 testing](#), *Nature Medicine*, 26, 810-811.
- Surkova, E, V Nikolayesky, F Drobniewski (2020). False-positive Covid-19 results: hidden problems and costs, *The Lancet Respiratory Medicine*, 8-12, 1167-1168.
- Watson, J, PF Whiting and JE Brush (2020). [Interpreting a Covid-19 test result](#), *BMJ*, 369:m1808.
- WHO (2020). [Antigen-detection in the diagnosis of SARS-CoV-2 infection using rapid immunoassays](#), *Interim guidance*, WHO/2019-nCoV/Antigen_Detection/2020.1.