

Test, Test, Test; Lessons learned from experience with mass screening programmes.

Advice Note for Independent SAGE 5 June 2020

Angela E. Raffle, Consultant in Public Health, Honorary Senior Lecturer, University of Bristol School of Community and Social Medicine, UK
Sian Taylor-Phillips, Professor of Population Health, Warwick Medical School
Department of Population Evidence and Technologies.

1. Preamble

This advice note draws on our;

- 35 years of developing, researching, delivering, and quality assuring screening programmes, coupled with 30 years' participating in communicable disease control on-call work,
- 13 years experience conducting, analyzing, synthesizing and systematically reviewing evidence related to tests and population programmes, in support of national policy making.

We hope that this paper, which has been quickly prepared, and which feels like 'work in progress', will be of some use to Independent SAGE members. We are well aware that most of you will be familiar with many of the concepts described, but maybe not all members will be familiar with all of them.

Mass testing for COVID-19 has the potential, if delivered to the right people at the right time with the right ongoing action, to bring benefit. However, the current government policies, relying on fragmented services by inexperienced providers without integration into local systems of care, risk poor delivery with consequent harm to tested individuals and to the wider population, as well as loss of trust in Public Health Services. Independent SAGE can help to prevent this harm, by supporting coordinated high quality provision embedded in statutory services.

2. Principles of testing in healthy people

There are common principles applicable both to testing with the aim of communicable disease control (CDC) and testing programmes overseen by the UK National Screening Programmes (NSP), but there are also profound differences. The differences (summarised briefly in Appendix 1) are important, and relate principally to timescale, flexibility, purpose, approach, and ethical/ legal framework.

In practice of course the separation between CDC and NSP is not completely clear-cut. For example the UK NSP oversees infectious disease screening in pregnancy, which has successfully helped bring transmission rate of HIV in infants born to mothers living with HIV down to 0.27%¹.

3. Similarities between communicable disease control (CDC) testing and National Screening Programmes (NSP)

The sheer scale of testing being contemplated for COVID-19 pandemic control, and the length of time that this pandemic (and potentially others) is likely to impact, make it worthwhile to consider from the outset whether there are good practice principles relating to the many mistakes made in the early decades (1950s to 1980s) within screening programmes, which could be useful in informing the current approach.

The UK NSP pays considerable attention to clarity of purpose, systematic delivery, quality assurance, informed participation, equity, ethical oversight, building trust, public understanding, and safeguarding decision making from political and commercial interference. This is because of major adverse incidents, and public and professional outcry, on numerous occasions before 1996 when the UK NSC was established, that stemmed from these matters being overlooked.

We outline briefly below some features that CDC testing and NSP's have in common, and what implications this might have for testing during the pandemic.

3.1 Clarity of purpose

If there is ambiguity and confusion about exactly why, when, and for what purpose a given test is being performed, this leads to problems. COVID-19 testing was initially directed for patients with symptoms, but then shifted to being used to confirm absence of infection in symptomless, i.e. healthy people. When there is confusion about aims, and about how well a test can fulfill that aim, this tends to lead to misinformation, waste of resources, and loss of public and professional confidence in the programme. For example, decades ago when Down's Syndrome screening first began, the professional view was that the aim was to prevent the birth of children with the syndrome. This aim was unethical, and inappropriate, and led to huge difficulties. In relation to COVID-19 testing, the announcement that 100,000 tests were to be done each day led to an unintended shift in the purpose of testing - to a focus on meeting the target rather than on providing benefit. The result was that many tests were done that were irrelevant to meeting the objective of avoiding unnecessary self-isolation in symptomless key workers.

Implication for best practice; it is good to be crystal clear on the ultimate purpose of a given test in a given setting, and on the evidence that demonstrates that the given purpose can be met.

3.2 It needs a system not just a test

A system is a set of interlinked activities with a collective purpose. A testing programme will not fully meet its objectives unless there is a coordinated pathway; encompassing identifying who to test, communicating the need for a test, taking the test, carrying out the test, communicating the result to the person tested and to others who need to know, and ensuring that appropriate next steps actually happen. It appears at present in the pandemic that parallel systems are being created, with

commercial companies taking on functions that can only be effective if they are delivered by appropriately trained staff, to clear quality standards, and integrated within existing primary care, secondary care and public health services. Accurate and secure recording is needed, with appropriate and timely sharing of information between different agencies, safeguards for confidentiality of personally identifiable data, measurement and reporting on quality standards for the entire system. In the early years of cervical screening for example (1960s-1980) there were no effective systems in place for any part of the pathway, mortality benefit was not realized, harmful treatments were delivered to many who did not need it, and systems for following up people with abnormal test results were incomplete and disjointed². These same issues are appearing in the mass testing programme for COVID-19.

Implication for best practice; it is beneficial to specify the whole system, together with measures of quality relating to the objectives for each step in the pathway.

3.3 The people tested are healthy, and are not patients

Whenever a test is offered or required in someone who has not sought help for signs or symptoms, the nature of the healthworker/patient relationship and of the contract between them changes from traditional diagnosis or disease management. This is significant because the person being tested may derive no benefit, and may suffer harm directly or through unintended consequences of uncertain information. As COVID-19 testing shifts to testing symptomless people, the impacts if the testing is not done well could include diversion of resources, and inadvertent transmission from false reassurance.

Implication for best practice; provision of information to those tested, that explains the consequences, both harmful and beneficial, and gives ethical justification, is important for building public trust, and is a requirement of General Medical Council guidance³. It is good to be explicit about whether a test is primarily for gathering surveillance data, or for research, or for protecting others, or for slowing transmission, or primarily for the direct benefit of the recipient. This is of particular importance with antibody testing done for surveillance, because the recipient may believe the test is for their benefit and put undue reliance on the assumed immunity.

3.4 Test performance, what the test actually measures, and what the results really mean are of critical importance

Decades of harm and confusion have been caused by wrong assumptions about the meaning of screening test results. To give an example, before the UK NSP successfully improved the quality standards for Down's syndrome screening, many terminations of unaffected pregnancies took place on the basis of falsely positive results⁴.

The principle of linking test results, to treatment decisions, to outcomes both for the person tested and for others, is fully established in policy-making for medical tests and screening programmes. It forms part of the evaluation processes of NICE, the UK

National Screening Committee, the US Preventive Services Task force and the Australian MSAC amongst others. The linkage of test results to actions to outcomes is one of the cornerstones of ensuring that medical tests and screening programmes lead to more good than harm. Applying these principles to COVID-19 tests highlights two critical issues. Firstly, PCR is known to miss many cases (low sensitivity), so advising a symptomatic person with a single negative PCR test that they and their contacts no longer need to self-isolate has the potential for considerable harm through onward transmission. WHO guidance is clear that even two consecutive negative PCR tests does not rule out COVID-19⁵. Secondly, we do not yet know whether presence of antibodies infers immunity, and if it does, how that changes over time. If an individual tests positive for antibodies, and mistakenly believes that they are immune to COVID-19 and/or cannot transmit it they may change their behaviour to increase the risk of infection and transmission, including to vulnerable people. Further detail of test performance and evaluations is in Appendix 2.

Implication for best practice; it is important to question any test implementation where it is unclear how receiving the test is of benefit to the individual, their contacts and the population. Further, as health professionals, we tend to use language that is easily understood in health circles, but this can lead to misunderstanding for journalists or members of the public. Words need to be used as carefully as algebraic equations, for example ‘live virus’ testing is a misleading term, since a positive PCR test could result from presence of degraded viral RNA fragment persisting in someone who is not infectious. Numbers, as Gigerenzer has demonstrated⁶, are best expressed simply – not as decimals or percentages but in natural frequencies, and using absolute risks not relative risks. Uncertainty is best acknowledged.

3.5 Public trust, public understanding, and public cooperation

UK experience suggests that an authoritarian or paternalistic approach to testing healthy people will, sooner or later, backfire. The UK population is highly diverse in terms of livelihoods, literacy, beliefs, cultural background, class etc. When testing millions of people, the subtlety and nuance encompassed in ordinary day to day, human to human, interactions with local health service workers gets lost. This makes it difficult to communicate in a way that enables people to feel they are being treated as human beings, and the result can be anger, bitterness, complaints, and legal challenge. For example in the breast screening programme women were initially encouraged to attend without being told that overdiagnosis and overtreatment inevitably occur for some screened women. This led to fierce controversy before the information was changed to include overtreatment⁷.

Implication for best practice; communication needs to be sufficiently honest, clear and accurate, that it will be trusted irrespective of whether the recipient of the testing programme is someone below living wage caring for a terminally ill relative, a professor of medical ethics, or an asylum seeker whose first language is not English. Primary care services need to be involved and informed so that they can support people who receive tests.

Trust and collaboration between agencies

In both CDC testing and NSP testing there are numerous different agencies conducting different elements of the system. Trust, good relationships, a shared public service ethic, and a universal commitment to quality assurance, is important for achieving best practice.

Implication for best practice; standards relating to confidentiality for personally identifiable data, research governance, and collection/reporting of performance data relating to quality standards, need to apply across all organisations participating in testing programmes for COVID-19. The pandemic has led to many normal rules being relaxed, which has helped clear the way for rapid action, but if all rules of best practice are ignored then harm will result.

3.6 The inverse care law applies

Those who are most able to understand and access tests are generally those who are also healthy, well to do, and less vulnerable. For example before 1988 the cervical screening programme had no call and recall system, with the result that lowest risk women had very frequent tests, and high risk women had none at all. For COVID-19, if testing is haphazardly delivered there is a risk of devoting considerable resource with little beneficial effect.

Implication for best practice; processes are needed to ensure that tests are directed at people and contexts where benefit will ensue, and quality measures need to relate to specific objectives and not just to the number of tests and to turnaround times. Attending to sound evidence that tests can achieve what we want them to achieve, and attention to cost-benefit, can also help. Close working with voluntary sector and civic society in every locality can also be helpful; the relationships between statutory, sector, voluntary sector and civic society appear to have strengthened during the pandemic so far. Strong local participation can help mitigate the divisive influence of the current political climate.

3.7 Political, media, and commercial factors

Where healthy people are concerned, there is no shortage of vociferous opinions about what tests should or not be available. This applies both for screening and for communicable disease control. Public relations activities by companies with vested interests, usually enacted covertly, play a part in influencing the media, politicians, and the public. For example in 2007, journalists and celebrities were flown from across the world to attend what appeared to be a scientific 'Global Summit' on cervical prevention; in truth the event was stage managed entirely on behalf of Sanofi Pasteur with the aim of lobbying for HPV vaccine ahead of mature research evidence⁸. And in 2019 a legal challenge began against Roche for their allegedly fraudulent promotion of Tamiflu for H1N1 flu⁹. There seems little doubt that commercial interests will be at play in the handling of the current pandemic. The

difficulty of making careful and well-founded policy recommendations in this environment is considerable.

Implication for best practice; Independent SAGE can be exemplars, so that the members themselves, and the conduct of the group as a whole, remains above criticism. This will ensure that the group is in a position to ‘call out’ conflicts of interests and poor practice.

3.8 Direct-to-Consumer advertising

COVID-19 PCR tests and antibody tests are being advertised direct to consumers in the UK. Information provided on the websites selling these tests is incomplete, and in some cases dangerously misleading. The marketing of tests direct to the public is governed only by the Advertising Standards Authority and few safeguards exist to prevent misleading information about implied benefits¹⁰. This is a serious concern in the screening world also.

Implication for best practice; clear, impartial, plain English information for the public, based on best evidence, issued by an independent and trusted source, and readily available through social media and other channels, can help to protect ordinary citizens from misleading marketing claims about medical tests and interventions.

Appendix 1. Key differences; timescale, flexibility, purpose, approach, and ethical/legal framework, in relation CDC testing and NSP testing.

	Communicable disease control (CDC) testing	National Screening Programmes (NSP)
Timescale	Notifications and surveillance systems have evolved over centuries, but for a new infectious agent, time is of the essence, and testing programmes need to be delivered extremely rapidly in response to a threat. Preparedness can help.	Screening programmes are major undertakings, testing millions of people (12 million screened by the England NSP in 2017/18). It needs a decade or more to research and implement a high quality programme.
Flexibility	Testing programmes during an epidemic or outbreak are introduced and then scaled down. Rushed implementation can cause an expensive and damaging mess, but will not be ‘set in stone’ in the same way as for a national screening programme.	It is extremely hard to modify or cease a screening programme once it is started. Rushing into screening without prior evaluation and careful implementation (as happened with cervical screening in the 1960s) creates a damaging, expensive, long-lasting mess.
Purpose	Tests are done for multiple purposes , including; SURVEILLANCE (prevalence and incidence monitoring of agent and disease as part of the CDC programme) RESEARCH (into transmission, morbidity, mortality, treatment, immunity, prevention is part of the CDC programme) DIAGNOSIS (to guide treatment during illness) PROTECTION OF OTHERS (e.g testing aimed at ensuring vulnerable people are not exposed to infectious individuals) PREVENTION OF TRANSMISSION (e.g. testing those travelling into ports of entry), CASE FINDING in order to offer prompt treatment, and CONTACT TRACING – although the ‘test’ in this case may consist of questions about contact rather than a medical test.	A screening programme generally has a single, explicit purpose – to reduce chance of a specific unwanted future outcome for the screened individual (e.g. reducing risk of death from ruptured aortic aneurysm). Research to first evaluate the consequences of screening is done separately, and covered by research governance. Epidemiological research such as morbidity surveys and disease surveillance are also carried out separately. In the early days, poor quality screening was often ‘nodded through’ because of a desire to gather morbidity statistics relating to symptomless people. This led to problems due to public expectations that screening was proven, harm through inconsequential findings etc.
Approach	A variety of approaches to testing are practiced, and evolve, across	National programmes are highly uniform, and universally

	<p>time and place, according to a) level of circulating infection b) nature of population e.g. social and ethnic mix, mobility, vulnerability etc. c) configuration and capacity of local services.</p> <p>Testing is only one element, with many other control measures, pharmaceutical and non-pharmaceutical, also applying.</p> <p>Also, considerable reliance is placed on a single test result, with no confirmatory step, Also infections arise all the time therefore someone can become infected immediately after having a test, with misleading conclusions.</p>	<p>provided. This practice was adopted because early haphazard screening programmes led to criticisms of unfairness, and to escalation of overdiagnosis and harmful overtreatment. Without uniform standards clinicians had to 'play safe' or risk being criticized for inevitable undetectable cases. Also, in screening programmes there is usually a further diagnostic phase to examine in more detail those whose initial screen result is positive.</p>
<p>Ethical/ legal framework</p>	<p>Testing in CDC programmes is generally a recommendation rather than an offer. Some tests are required, for example for certain employees, visas etc. The Public Health (Control of Disease) Act 1984 gives certain powers to Local Authorities. In practice, these powers are exercised through persuasion, appeal to duty, power to exclude from work, payment vouchers, assistance with accessing tests etc. and enacted by Environmental Health Officers, health service staff or social care workers, rather than through forceful means involving police officers.</p>	<p>For most of the National Screening Programmes the screening is an 'offer' with strong emphasis on informed choice given that screening leads to harm as well as to potential benefit. For infectious disease screening in pregnancy (HIV, Hep B, syphilis) screening is 'recommended', and for newborn bloodspot 'strongly recommended'. Up until 2001 a more persuasive approach was practiced, and in the 1960s-1990s people were not given real choice. It took decades of challenge from lawyers, ethicists, NHS staff, and citizens before the principle of informed choice was fully recognised.</p>

Appendix 2: Test accuracy and implications of results

The World Health Organisation and the Cochrane Collaboration are working together to systematically review the evidence about the accuracy of COVID-19 molecular tests, mostly polymerase chain reaction (PCR) and antibody tests (led by Jon Deeks, Birmingham University). These will be published in the coming weeks.

The testing pathways using PCR and antibody tests in the UK have not been evaluated by NICE or the UK National Screening Committee. However, NICE have produced a best practice guide for undertaking COVID-19 test evaluation studies (ref). In this they recommend evaluating test accuracy in real-world practice (single gate studies) rather than just in the laboratory on known samples and pre-pandemic sera (two gate studies). This is because two gate studies tend to include extreme cases so overestimate accuracy. These evaluations of the antibody test are underway in the UK currently. The UK roll-out of the antibody test was based on two gate laboratory studies.

The evidence to date is that laboratory based antibody tests using venous blood have good (but not perfect) accuracy to detect previous infection. This accuracy is heavily dependent on time since infection. The implications of this accuracy for positive predictive value are heavily dependent on disease prevalence.

Below are some pathways by which use of the PCR test and antibody test may benefit and harm the people tested and others.

1. **PCR positive:** The intended use of the PCR test is to detect people who currently have active COVID-19 virus, who may infect others. The action in this case is to isolate the person who tests positive, to prevent infection of others, and this reduction of transmission in turn reduces COVID-19 related mortality and morbidity. A further action which could beneficially reduce mortality and morbidity is tracking and tracing the contacts of the confirmed case. This is a clear pathway to benefit from the test.
2. **PCR negative:** The PCR test has low sensitivity, so there are many false negative results (undetected cases). If symptomatic people who test negative are allowed to end self-isolation this is very likely to result in significant transmission, and the associated mortality and morbidity harm. The World Health Organisation laboratory guidelines¹¹ state COVID-19 cannot be ruled out even after two consecutive negative PCR tests. On that basis the policy in New Zealand is a requirement for three consecutive negative PCR tests before COVID-19 is ruled out. To prevent unnecessary transmission a symptomatic person who tests negative once on PCR should continue to self-isolate. If symptomatic people who test PCR negative end self-isolation this will increase transmission and may undermine the benefits of introducing the test.
3. **Antibody test positive:** We do not know whether the presence of antibodies confers immunity from re-infection with COVID-19. If an individual tests positive for antibodies, and misunderstands that result, they may put

themselves at greater risk of infection and transmission to others based on the mistaken assumption that they are immune. There is currently no obvious pathway to benefit from the antibody testing programme. There is potential for a large national experiment, where infection rates in people with previous positive and negative antibody test results are recorded, to help us understand the relationship between antibodies and immunity. This however would need to be clearly explained to the research participants taking the test.

REFERENCES

- ¹ UK National Screening Programme Key Performance Indicators for 2017/18
- ² Raffle, A.E. Mackie, A. Gray, J.A.M. (2019) Case Study 1.1 (p 15-20) in *Screening; Evidence and Practice*. 2nd edition. Oxford University Press. June 2019. ISBN 9780198805984
- ³ General Medical Council (2008) *Consent: patients and doctors making decisions together*. London: General Medical Council.
- ⁴ Interview with Sir Muir Gray
<https://www.healthknowledge.org.uk/interactive-learning/screening/chapter9>
[accessed 5 June 2020]
- ⁵ World Health Organization. (2020). Laboratory testing for coronavirus disease (COVID-19) in suspected human cases: interim guidance, 19 March 2020. World Health Organization. <https://apps.who.int/iris/handle/10665/331501> . License: CC BY-NC-SA 3.0 IGO
- ⁶ Gigerenzer, G. (2002) *Reckoning with risk*. London: Penguin
- ⁷ Raffle, A.E. Mackie, A. Gray, J.A.M. (2019) Case Study 3.1 (p59-60) in *Screening; Evidence and Practice*. 2nd edition. Oxford University Press. June 2019. ISBN 9780198805984
- ⁸ Boseley, S. (2007) Vaccination campaign funded by drug firm. *The Guardian* 26 March 2007
<https://www.theguardian.com/society/2007/mar/26/cancercare.health2>
[accessed 5 June 2020]
- ⁹ Dyer, O. (2020) Researcher sues Roche over Tamiflu claims. *BMJ* 2020;368:m314 https://www.bmj.com/bmj/section-pdf/1019683?path=/bmj/368/8231/This_Week.full.pdf [accessed 5 June 2020]
- ¹⁰ Raffle A E Advertising private tests for well people. Guest Editorial. *Clinical Evidence* 2, June 2010
- ¹¹ World Health Organization. (2020). Laboratory testing for coronavirus disease (COVID-19) in suspected human cases: interim guidance, 19 March 2020. World Health Organization. <https://apps.who.int/iris/handle/10665/331501> . License: CC BY-NC-SA 3.0 IGO