



# THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Test, test, test for COVID-19 antibodies

**Citation for published version:**

Kumleben, N, Bhopal, R, Cypionka, T, Gruer, L, Kock, R, Stebbing, J & Stigler, FL 2020, 'Test, test, test for COVID-19 antibodies: the importance of sensitivity, specificity and predictive powers', *Public Health*, vol. 185, pp. 88-90. <https://doi.org/10.1016/j.puhe.2020.06.006>

**Digital Object Identifier (DOI):**

[10.1016/j.puhe.2020.06.006](https://doi.org/10.1016/j.puhe.2020.06.006)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Public Health

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



## **Test, test, test for COVID-19 antibodies: the importance of sensitivity, specificity and predictive powers**

Nicholas Kumleben<sup>1</sup>, Raj Bhopal<sup>2</sup>, Thomas Czypionka<sup>3,4</sup>, Laurence Gruer<sup>5,6</sup>, Richard Kock<sup>7</sup>, Justin Stebbing<sup>8</sup>, Florian L. Stigler<sup>9</sup>

1. Research Analyst, Greenmantle LLC, New York City, NY 10010, USA. [nicholas@gmantle.com](mailto:nicholas@gmantle.com)
2. Emeritus Professor of Public Health, Usher Institute, University of Edinburgh, Edinburgh, EH8 9AG, UK. [raj.bhopal@ed.ac.uk](mailto:raj.bhopal@ed.ac.uk)
3. Head of Health Economics and Health Policy, Institute for Advanced Studies, 1080 Vienna, Austria. [czypionk@ihs.ac.at](mailto:czypionk@ihs.ac.at)
4. Visiting Senior Research Fellow, London School of Economics and Political Science, London, UK.
5. Honorary Professor of Public Health, Usher Institute, University of Edinburgh, Edinburgh, EH8 9AG, UK
6. Honorary Professor of Public Health, Institute of Health and Wellbeing, University of Glasgow, G12 8RZ, UK. [gruer.health@gmail.com](mailto:gruer.health@gmail.com)
7. Professor, Department of Pathobiology and Population Science, Royal Veterinary College, Hawkshead Lane, Hatfield, AL9 7TA, UK. [rkock@rvc.ac.uk](mailto:rkock@rvc.ac.uk)
8. NIHR Research Translational Professor, Imperial College, Hammersmith Hospital, London, W12 0HS, UK. [j.stebbing@imperial.ac.uk](mailto:j.stebbing@imperial.ac.uk)
9. GP and Public Health Researcher, Österreichische Gesundheitskasse, 8010 Graz, Austria. [florian.stigler@medunigraz.at](mailto:florian.stigler@medunigraz.at)

Corresponding author: Laurence Gruer [gruer.health@gmail.com](mailto:gruer.health@gmail.com)

## **Abstract**

SARS-CoV-2 antibody tests of varying specificity and sensitivity are now available. For informing individuals whether they have had COVID-19, they need to be very accurate. For measuring population prevalence of past infection, the numbers of false positives and negatives need to be roughly equal.

With a series of worked examples for a notional population of 100,000 people, we show that even test systems with a high specificity can yield a large number of false positive results, especially where the population prevalence is low. For example, at a true population prevalence of 5%, using a test with 99% sensitivity and specificity, 16% of positive results will be false and thus 950 people will be incorrectly informed they have had the infection. Further confirmatory testing may be needed.

Giving false reassurance upon which personal or societal decisions might be based could be harmful for individuals, undermine public confidence and foster further outbreaks.

## **Key words**

COVID-19; antibody tests; specificity; sensitivity

## **Highlights**

- COVID-19 antibody tests need to be very reliable to inform individual decisions
- Even with high specificity, antibody tests can produce many false positives
- Giving false reassurance could undermine public confidence and foster new outbreaks

## **Main text**

To help reverse the current lockdowns while suppressing COVID-19 rates, we need to identify who currently has the infection and who has had it and recovered. As RT-PCR testing to detect current infection has been recently discussed in detail,<sup>1</sup> we focus in this paper on antibody tests. The presence or absence of antibodies can inform individuals if they have had the infection or not and guide personal and societal decisions about if and when they can return to normal activities. Antibody testing thus needs to be particularly accurate. It can also be used to provide an estimate of the population prevalence of previous infection. We demonstrate that for this

purpose high accuracy is not required, but the numbers of false positives and false negatives need to be approximately equal.

Antibody tests are increasingly available but with variable accuracy. It is hoped they can be used to identify people who are at least partially immune. Immunity certificates, a more appropriate phrase than immunity passports that promises too much, for individuals thought to have recovered from COVID-19, are being discussed internationally.<sup>2 3 4</sup> Whether tests are done for clinical diagnosis, screening or immunity certificates, we need to have sufficient confidence they are accurate.

A sensitive test will detect the presence of antibodies to SARS-CoV-2 (the virus that causes COVID-19), and a specific test will not react to other antibodies e.g. to other coronaviruses. No diagnostic or screening test is perfect and incorrect results are inevitable, not least because the timing of the test is critical. Seroconversion takes time, with IgM, IgG and IgA antibodies usually developing in that order, and can be variable and dependent upon the severity of the illness and the individual's immune system. Antibody levels subsequently decline with time. Antibody test systems may perform less well than the manufacturers' results suggest. For example, both Roche and Abbott reported their antibody test had 100% sensitivity for samples taken 14 days or more after the onset of symptoms, yet Public Health England found sensitivity at 14 or more days of only 87% and 93.4% respectively.<sup>5 6</sup>

We show here how to measure the test's accuracy and how this changes along with the prevalence of disease (12 tables showing the results with varying sensitivity, specificity and population prevalence of 1%, 5%, 10%, and 20% are available in the Supplementary File). The two key measures of its accuracy are sensitivity and specificity, set out in Table 1, with the cells identified as A (true positives), B (false positives), C (false negatives), and D (true negatives). Sensitivity ( $A/A+C$ ) is the proportion of people with a disease who, when tested, receive a positive test result. It is also known as the true positive rate. Specificity ( $D/D+B$ ) is the proportion of individuals without a disease who, when tested, receive a negative test result. It is also known as the true negative rate.

To establish sensitivity and specificity, we could test a sample of patients with proven disease (in this case laboratory detection of SARS-CoV-2), and a sample of people known to be free of disease (for example, using stored blood samples taken before COVID-19 existed in humans). In practice, a test's performance will usually be poorer than the values established due, for example, to problems in storing or transporting specimens or the variable time lag from the onset of infection until antibodies appear in the blood (seroconversion) and then decline. The proportion of test results that are false partly depends on the prevalence of the disease in the population. With a low prevalence, even a test with high sensitivity and specificity will produce a high proportion of false positives. In this paper, we focus on the outcomes of tests of variable accuracy with 5% population prevalence in a hypothetical group of 100,000 people, of whom 5,000 have had the infection and 95,000 have not. This is a plausible current prevalence of past COVID-19 in many countries<sup>7 8 9</sup> although it could be a lot higher in some areas.

Table 1 shows that if the sensitivity is 90%, 4,500 people will correctly test positive, but 500 will incorrectly test negative and be wrongly told they have no antibody evidence of the disease. If the specificity is 90%, 85,500 people will correctly test negative, but 9,500 will incorrectly test positive and be wrongly told they have antibody evidence of previous infection. Thus, of the 14,000 people who received positive test results, only 32% ( $4,500/14,000$ ;  $A/A+B$ ) had the disease. This is referred to as the predictive value (or power) of a positive test. The other 68% would be given wrong information. Of the 86,000 people who received negative tests, 99% ( $85,500/86,000$ ;  $D/C+D$ ) would receive a correct result. This is called the predictive value (or power) of a negative test.

Sensitivity and specificity vary with different tests but, for any particular antibody test, these can be adjusted by altering the level of antibody required to determine a positive result. Requiring a higher level of antibody for a positive result would increase the specificity but lower the sensitivity. This would reduce the false positives (C) but increase the false negatives (B). Choosing a test that has 80% sensitivity and 99% specificity, as shown in Table 2, 81% of people who test positive have had the disease, an increase from 32%. Now, about one in five people who test positive will not have had the disease. This shows that when the prevalence of the

disease is low, antibody testing, even with a specificity as high as 99%, still produces many false positives so the predictive power of a positive test is far from 100%.

If a test is extremely accurate, as is claimed for the Roche and Abbott systems, say 99% sensitivity and specificity, the results are shown in table 3. Even now, the predictive power of a positive test has only risen from 81% with a sensitivity of 90%, to 83.8%. If the prevalence rises to 20% then the predictive power of a positive test is 96.1% and of a negative test 99.7% (Supplementary File Table A12).

If immunity certificates, or personal or societal decisions about returning to normality, are based on these results, a significant proportion will be incorrect. Where the disease has become highly prevalent, for example among health care and care home workers, the power of a positive test would be higher, so more reliance could be placed on it. Even with a prevalence of 20% and 99% sensitivity and specificity the test itself does not give a guarantee at the individual level and personal and clinical judgements are required in applying the findings. A major hope of antibody testing is that those who test positive can resume work and social activities more fully and confidently than those who test negative. The presence of antibodies should signify the same illness will not recur, the person is not contagious and there is at least partial immunity to future COVID-19 infections. We need to establish whether this is true.<sup>10</sup>

If the purpose of antibody testing is to assess the prevalence of COVID-19 in a representative sample of the population, these clinical issues do not apply. The veracity of the prevalence derived by such measurements will depend upon achieving equal false positives and false negatives. For example, while the true prevalence is 5%, Tables 1, 2 and 3 give a prevalence in the hypothetical population of 100,000 people of 14% (14,000 positives), 4.95% (4,950 positives) and 5.9% (5,900 positives) respectively. Perhaps surprisingly, the test with 80% sensitivity and 99% specificity (Table 2) gives the most accurate estimate at this level of population prevalence.

In conclusion, at currently reasonable estimates of general population prevalence, even high sensitivity and specificity will produce an important number of false positives. People testing positive, especially those without indicative case histories, may need further testing to confirm

the result. Given the current uncertainty about the level of immunity signalled by antibodies, all those testing positive for antibodies would be well advised to maintain protective measures. More information is also urgently needed to ascertain the strength and duration of immunity in people who have recovered from COVID-19 and whether some can still be infectious or become re-infected. Giving false security and reassurance could be harmful for individuals, undermine public confidence and foster further outbreaks.

## **Acknowledgements**

### *Contributors*

The article was conceived by JS during a discussion initiated by RB in the COVID-19 researchers Google Group. The manuscript was drafted by NK. All authors commented on the drafts and agreed the final version. LG is the corresponding author and guarantor of the manuscript.

### *Ethical approval*

Not required

### *Funding*

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

### *Competing interests*

None declared

Table 1 Predictive powers of a test with 90% sensitivity and specificity (5% prevalence)

Test result (90% sensitivity and 90% specificity)	People truly with disease	People truly without disease	Totals
Positive	4,500 (A)	9,500 (B)	14,000
Negative	500 (C)	85,500 (D)	86,000
Total	5,000	95,000	100,000

Predictive value of a positive test:  $A/A+B = 32.1\%$

Predictive value of a negative test:  $D/D+C = 99.4\%$

Table 2 Predictive powers of a test with 80% sensitivity and 99% specificity (5% prevalence)

Test result (80% sensitivity and 99% specificity)	People truly with disease	People truly without disease	Totals
Positive	4,000 (A)	950 (B)	4,950
Negative	1,000 (C)	94,050 (D)	95,050
Total	5,000	95,000	100,000

Predictive value of a positive test:  $A/A+B = 80.8\%$

Predictive value of a negative test:  $D/D+C = 98.9\%$

Table 3 Predictive powers of a test with 99% sensitivity and 99% specificity (5% prevalence)

Test result (99% sensitivity and 99% specificity)	People truly with disease	People truly without disease	Totals
Positive	4,950 (A)	950 (B)	5,900
Negative	50 (C)	94,050 (D)	94,100
Total	5,000	95,000	100,000

Predictive value of a positive test:  $A/A+B = 83.8\%$

Predictive value of a negative test:  $D/D+C = 99.9\%$



## References

---

<sup>1</sup> Watson J, Whiting PF, Brush JE. Interpreting a COVID-19 test result. *BMJ* 2020;**369**:m1808 doi: 10.1136/bmj.m1808.

<sup>2</sup> Horowitz J. In Italy, Going back to work may depend on having the right antibodies. *New York Times* <https://www.nytimes.com/2020/04/04/world/europe/italy-coronavirus-antibodies.html>

<sup>3</sup> Forgey Q. Fauci: Coronavirus immunity cards for Americans are 'being discussed' *Politico* <https://www.politico.com/news/2020/04/10/fauci-coronavirus-immunity-cards-for-americans-are-being-discussed-178784>.

<sup>4</sup> Bhopal R. The burning building: recruiting recovered COVID-19 patients to the frontline by issuing certificates and offering voluntary registration. *BMJ* 2020;**368**:m1101.

<sup>5</sup> Public Health England. Evaluation of Roche Elecsys Anti- SARS-CoV-2 serology assay for the detection of anti-SARS-CoV-2 antibodies. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/886159/PHE\\_Evaluation\\_of\\_Roche\\_Elecsys\\_anti\\_SARS\\_CoV\\_2.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/886159/PHE_Evaluation_of_Roche_Elecsys_anti_SARS_CoV_2.pdf).

<sup>6</sup> Public Health England. Evaluation of the Abbott SARS-CoV-2 IgG for the detection of anti-SARS- CoV-2 antibodies. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/886158/PHE\\_Evaluation\\_of\\_Abbott\\_SARS\\_CoV\\_2\\_IgG.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/886158/PHE_Evaluation_of_Abbott_SARS_CoV_2_IgG.pdf).

<sup>7</sup> Flaxman S, Mishra S, Gandy A et al. Estimating the number of infections and the impact of nonpharmaceutical interventions on COVID-19 in 11 European countries. Imperial College London (30-03-2020) doi: <https://doi.org/10.25561/77731>.

<sup>8</sup> Ministerio de Ciencia y Información. Estudio ENE-COVID-19: Primera Ronda. Estudio Nacional de Sero-epidemiología de la Infección por SARS-CoV-2 en España. Informe preliminar de 13 de mayo 2020. [https://www.ciencia.gob.es/stfls/MICINN/Ministerio/FICHEROS/ENECOVID\\_Informe\\_preliminar\\_cierre\\_primera\\_ronda\\_13Mayo2020.pdf](https://www.ciencia.gob.es/stfls/MICINN/Ministerio/FICHEROS/ENECOVID_Informe_preliminar_cierre_primera_ronda_13Mayo2020.pdf).

<sup>9</sup> Salje H, Tran Kiem C, Lefrancq N, Courtejoie N et al. Estimating the burden of SARS-CoV-2 in France. *Science* 10.1126/science.abc3517 (2020).

<sup>10</sup> Okba NMA, Muller MA, Li W, et al. SARS-CoV-2 specific antibody responses in COVID-19 patients. *medRxiv*. March 2020:2020.03.18.20038059. doi:10.1101/2020.03.18.20038059.

