

This is the peer-reviewed, manuscript version of an article published in the *Public Health*

The version of record is available from the journal site:

<https://doi.org/10.1016/j.puhe.2020.06.006>

© 2020. This manuscript version is made available under the CC-BY-NC-ND 4.0 license

<http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Access granted due to COVID-19

The full details of the published version of the article are as follows:

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

AUTHORS: N. Kumleben, R. Bhopal, T. Czypionka, L. Gruer, R. Kock, J. Stebbing, F.L. Stigler

JOURNAL: Public Health

PUBLISHER: Elsevier

PUBLICATION DATE: 11 June 2020

DOI: 10.1016/j.puhe.2020.06.006



Short Communication

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



N. Kumleben ^a, R. Bhopal ^b, T. Czepionka ^{c,d}, L. Gruer ^{b,e,*}, R. Kock ^f, J. Stebbing ^g, F.L. Stigler ^h

^a Greenmantle LLC, New York City, NY 10010, USA

^b Usher Institute, University of Edinburgh, Edinburgh, EH8 9AG, UK

^c Institute for Advanced Studies, 1080 Vienna, Austria

^d London School of Economics and Political Science, London, UK

^e Institute of Health and Wellbeing, University of Glasgow, G12 8RZ, UK

^f Department of Pathobiology and Population Science, Royal Veterinary College, Hawkshead Lane, Hatfield, AL9 7TA, UK

^g Imperial College, Hammersmith Hospital, London, W12 0HS, UK

^h Österreichische Gesundheitskasse, 8010 Graz, Austria

ARTICLE INFO

Article history:

Received 3 June 2020

Accepted 3 June 2020

Available online 11 June 2020

Keywords:

COVID-19
Antibody tests
Specificity
Sensitivity

ABSTRACT

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) antibody tests of varying specificity and sensitivity are now available. For informing individuals whether they have had coronavirus disease 2019 (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 on which personal or societal decisions might be based could be harmful for individuals, undermine public confidence and foster further outbreaks.

© 2020 The Royal Society for Public Health. Published by Elsevier Ltd. All rights reserved.

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 reverse transcriptase polymerase chain reaction (RT-PCR) testing to detect current infection has been recently discussed in detail,¹ we focus in this article 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.^{2–4} Whether tests are carried out 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

* Corresponding author. Usher Institute, University of Edinburgh, Edinburgh, EH8 9AG, UK.

E-mail addresses: nicholas@gmantle.com (N. Kumleben), raj.bhopal@ed.ac.uk (R. Bhopal), czepionk@ihs.ac.at (T. Czepionka), gruer.health@gmail.com (L. Gruer), rkock@rvc.ac.uk (R. Kock), j.stebbing@imperial.ac.uk (J. Stebbing), florian.stigler@medunigraz.at (F.L. Stigler).

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.^{5,6}

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 article, we focus on the outcomes of tests of variable accuracy with 5% population prevalence in a hypothetical group of 100,000 people, of whom 5000 have had the infection and 95,000 have not. This is a plausible current prevalence of past COVID-19 in many countries^{7–9} although it could be a lot higher in some areas.

Table 1 shows that if the sensitivity is 90%, 4500 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 9500 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% ($4500/14,000$; $A/A + B$) 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,

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	4500 (A)	9500 (B)	14,000
Negative	500 (C)	85,500 (D)	86,000
Total	5000	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	4000 (A)	950 (B)	4950
Negative	1000 (C)	94,050 (D)	95,050
Total	5000	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	4950 (A)	950 (B)	5900
Negative	50 (C)	94,050 (D)	94,100
Total	5000	95,000	100,000

Predictive value of a positive test: $A/A + B = 83.8\%$.
Predictive value of a negative test: $D/D + C = 99.9\%$.

the power of a positive test would be higher, therefore 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.¹⁰

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, although the true prevalence is 5%, Tables 1–3 give a prevalence in the hypothetical population of 100,000 people of 14% (14,000 positives), 4.95% (4950 positives), and 5.9% (5900 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 the 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 reinfected. Giving false security and reassurance could be harmful for individuals, undermine public confidence and foster further outbreaks.

Author statements

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.

Author contributions

The article was conceived by J.S. during a discussion initiated by R.B. in the COVID-19 researchers Google Group. The manuscript was drafted by N.K. All authors commented on the drafts and agreed the final version. L.G. is the corresponding author and guarantor of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.puhe.2020.06.006>.

References

1. Watson J, Whiting PF, Brush JE. Interpreting a COVID-19 test result. *BMJ* 2020;369:m1808. <https://doi.org/10.1136/bmj.m1808>.
2. 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>.
3. 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>.
4. Bhopal R. The burning building: recruiting recovered COVID-19 patients to the frontline by issuing certificates and offering voluntary registration. *BMJ* 2020;368:m1101.
5. 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.
6. 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.
7. Flaxman S, Mishra S, Gandy A, Unwin J, Coupland H, Mellan T, 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. <https://doi.org/10.25561/77731>.
8. 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, https://www.ciencia.gob.es/stfls/MICINN/Ministerio/FICHEROS/ENECOVID_Informe_preliminar_cierre_primera_ronda_13Mayo2020.pdf; 2020.
9. Salje H, Tran Kiem C, Lefrancq N, Courtejoie N, Bosetti P, Paireau J, et al. Estimating the burden of SARS-CoV-2 in France. *Science* 2020. <https://doi.org/10.1126/science.abc3517>.
10. Okba N, Müller M, Li W, Wang C, Geurts van Kessel C, Corman V, et al. SARS-CoV-2 specific antibody responses in COVID-19 patients. *medRxiv* March 2020;18:20038059. <https://doi.org/10.1101/2020.03.18.20038059>. 2020.03.