PRELIMINARY REPORT

Title: **Positive and Negative Predictive Values, Specificity, False Positives and their Impact on RT-PCR Mass Testing for SARS-CoV-2**

Author: **M. Lever, PhD**

Date: **11 June, 2021**

Abstract: Positive and negative predictive values (PPVs and NPVs) help researchers assess how reliable a medical test is. It was attempted to estimate PPVs and NPVs for the RT-PCR test used for SARS-CoV-2 based on a range of published values for sensitivity and specificity, and reported numbers of cases and tests. Numbers for Australia, Victoria, New South Wales, Western Australia and the Pfizer and AstraZeneca mRNA vaccine trials were used. In the cases of Australia, Victoria, New South Wales and Western Australia, the ratio of positive test results to total tests was too small to solve for PPV and NPV using plausible values of specificity. In the cases of the vaccine trials, the ratio of positive results to total tests was either too small or the PPV was calculated and was found to be less than 50%, too low for the RT-PCR test to be of any practical value. The findings highlight the unreliability of test results obtained through the use of mass testing.

1. Introduction

Positive and negative predictive values (PPVs and NPVs) indicate the likelihood that a person does or does not have a condition being tested for, based on the results of the test. No test is 100% accurate, so the PPV and NPV can be used to give the diagnosing physician an indication of how trustworthy the results of a test are. The PPV gives the probability that a positive test result is a genuine indication of the underlying condition under test, whereas the NPV gives the probability that a negative test result reflects a genuine absence of the condition.

PPV and NPV depend on the prevalence of the condition in the population, the sensitivity of the test and the specificity of the test.

- Prevalence refers to the number of cases of a disease that are present in a particular population at a given time as a percentage of the population.
- Sensitivity refers to the ability of a test to detect the condition under test, and
- Specificity refers to its ability to discriminate between the condition under test and other conditions.

A test with a low sensitivity will miss a significant number of actual cases resulting in the generating false negative results, whereas a test which has a low specificity will give positive test results in patients who do not have the condition under test, generating false positive results.

PPVs and NPVs are often illustrated in a two-way table as shown below. In this example it is assumed that the prevalence of the condition is 0.2%, the sensitivity of the test is 90%, and the specificity of the test is 95%. It is assumed 10,000 people are tested.

	condition	not condition	tot
+ve	18	499	517
-ve	2	9481	9483
tot	20	9980	10,000

Since the prevalence of the condition is 0.2%, the total number expected to have the condition is 0.2% of 10,000, or 20 people. With a sensitivity of 90% we would expect 18 out of 20 people with the condition to test positive, and two to test negative, generating false negatives.

With a specificity of 95% it would be expected that of the 9980 people without the condition, 95% of 9980, or 9481 people would test negative, whereas 5% or 499 people would test positive, generating false positives. Note in this example, the low prevalence results in the number of false positives generated (499) being significantly higher than the number of true positives (18), in this case approximately 28 times higher. This gives an indication of the potential dangers of using the results of mass testing for public health decision making.

The positive and negative predictive values are calculated below:

PPV = 18/517 x 100 = 3.5%
NPV = 9481/9483 x 100 = 99.98%

This means a person who tests positive has a 3.5% chance of actually having the condition under test, whereas a person who tests negative has a 99.98% chance of not actually having the condition. This example shows how conditions with a low prevalence in the population can have a low PPV even when the sensitivity and specificity may seem high enough for the test to be of use. With a PPV of 3.5% the test would be of no practical use, because the diagnosing physician could not say with any confidence that the patient who tested positive had the condition under test.

The following example shows what happens if the prevalence is increased to 10% without changing the other parameters:

	condition	not condition	tot
+ve	900	450	1350
-ve	100	8550	8650
tot	1000	9000	10000

PPV = 900/1350 x 100 = 66.7%
NPV = 8550/8650 x 100 = 98.8%

The PPV has risen significantly due to the increase in prevalence from 0.2% to 10%, but at 66.7% is arguably still not high enough to be of practical use as it does not give confidence

that the positive test result has a high probability of being correct. A higher sensitivity and specificity would be required to give such confidence at a prevalence of 10%.

2. General Solution

In order to investigate more closely the effects of the parameters affecting the PPV, a general solution is developed. The parameters are defined below:

p = the total number of reported cases in the time frame of interest (p for positive)
S_n = sensitivity
S_p = specificity
T = the tested population
a = the total number in the tested population who have the condition tested for
b = the total number in the tested population who do not have the condition tested for

The table can then be represented in parametric terms:

	condition	not condition	tot
+ve	S_n**a**	$(1–S_p)$b	p
-ve	**a** $(1–S_n)$	b$(1+S_p)$ -1	T-p
tot	**a**	b	T

Summing the top row gives: S_n**a**+ $(1 –S_p)$b = p

Substituting b = T – **a** from the bottom row into the above equation and solving for **a** gives:

$$\mathbf{a} = \frac{p – T\,(1 – Sp)}{Sn + Sp - 1}$$

a cannot be negative, so there is no solution when either the numerator or denominator is negative.

By solving the inequality obtained by making the numerator less than zero, the numerator is negative when $Sp < 1 - p/T$

The denominator is negative when $Sn + Sp < 1$

The case of a negative denominator will be disregarded since with all cases considered the sum of the sensitivity and specificity is greater than 1.

Therefore the criterion for the two-way table to have a (non-negative) solution for **a**, the number of people with the condition under test is:

$S_p > 1 - p/T$

For example, if 10,000 people are tested and 100 cases are diagnosed, then

p/T = 100/10,000 = 0.01 or 1%

1- p/T = 0.99 or 99%

So if we start with a specificity lower than 99%, the value of **a** will turn out to be negative, which means there is no solution. This can be interpreted that the number of cases does not reflect the specificity claimed. Even if they were all false positives, the number is so small that it could only be achieved if the specificity was higher.

If it is not realistic for the specificity to be higher, then the number of reported cases is unrealistically low, because we would expect more people to test positive due to false positives alone, even if there were no true cases in the tested population. In this case the reported numbers of cases cannot be relied upon. In the context of real-world achievable specificities, the reported results are 'too good to be true'.

3. RT-PCR testing for SARS-CoV-2

Nucleic acid testing and in particular RT-PCR has been established as the preferred test for SARS-CoV-2 under public health policies and guidelines.[1] PPV and NPV analyses apply to the diagnosis of SARS-CoV-2. The World Health Organisation's policy is that:

"...disease prevalence alters the predictive value of test results; as disease prevalence decreases, the risk of false positive increases. This means that the probability that a person who has a positive result (SARS-CoV-2 detected) is truly infected with SARS-CoV-2 decreases as prevalence decreases, irrespective of the claimed specificity." [2]

Numerous factors affect the accuracy of the RT-PCR test, in the pre-analysis, RT and PCR steps.[3] Values of sensitivity quoted in the literature for the RT-PCR test for SARS-CoV-2 typically vary from around 75% to 100%,[4,5] but it should be noted that figures at the higher end may be obtained using unusually high cut-off values for cycle threshold value (Ct). The Ct value refers to the number of cycles required to amplify the starting material to an amount which is detectable. Each cycle represents a nominal doubling of the number of target sequences identified, modified by the PCR efficiency and other factors.

Normally Ct values above 35 are not considered valid because of the likelihood that they are due to contamination or non-viable viral material. This is demonstrated in a study of 3,790 samples which tested positive for SARS-CoV-2,[6] which correlated the Ct value with infectivity as measured by the ability to grow out the sample in tissue culture. The results as stated were:

"It can be observed that at Ct = 25, up to 70% of patients remain positive in culture and that at Ct = 30 this value drops to 20%. At Ct = 35, the value we used to report a positive result for PCR, <3% of cultures are positive."

Three further references are given in the study in support of this finding, which is also supported by Dr Fauci, who has stated, "If you get a cycle threshold of 35 or more, the chances of it being replication-competent are miniscule".[7] At present the maximum Ct value

set by the CDC for submitting samples for investigating cases of SARS-CoV-2 in people who have been vaccinated (referred to as 'vaccine breakthrough cases') is only 28.[8]

The fact that excessively high Ct values have been contributing to the generation of false positives since early 2020 is supported by key studies and documents. The original study used by the WHO to set worldwide protocols for RT-PCR testing for SARS-CoV-2 utilised a Ct value of 45.[9] Although this study was discredited by a team of international scientists who identified 10 critical flaws with the methods used,[10] the protocols taken from it were not revised.

The Australian Public Health Laboratory Network guidance on nucleic acid test result interpretation for SARS-CoV-2 states that usually 35 to 45 cycles are taken, and it is noted that increasing the Ct value beyond 40 may increase the risk of false positives due to non-specific amplification.[11] The Doherty Institute post-market validation of the Beijing Genomics Institute (BGI) SARS-CoV-2 Real Time PCR platform states that a positive result was interpreted for Ct values less than 38.[12]

The effect of Ct cut-off value on sensitivity can be observed by comparing the sensitivities of seven commercial assays which varied from 77% to 100% based on a cut-off Ct value of 42.5.[4] If the cut-off Ct value is lowered to 35, then the sensitivities drop to between 54% and 77%. In general public health reporting Ct values are not included in the data reported so it is not possible to include them in the analyses carried out in Section 4.

Values reported for specificity tend to be typically in the range of 95% to 100%.[12,13,14,15] However, in some of these studies it is clear there are problems with reproducibility, suggesting the high specificities reported are not easily achieved in practice. For example, in one study, of 45 samples which had previously tested positive using RT-PCR, only 8 (18%) tested positive in the initial test. Of 52 samples which tested positive in the initial test, only 29 (56%) tested positive in the second test.[13]

It could be interpreted that non-confirmation in the second test either represents a loss of sensitivity from the first to the second test, or a lack of specificity in the first test, i.e. the samples which did not confirm the second time had originally tested positive due to something else. Depending on which assumption you make, the sensitivity or specificity reduces to 56% (29/52). The authors tested for loss of sensitivity due to storage but did not obtain statistically significant results, suggesting the second interpretation, a lack of specificity, is more likely.

These findings are consistent with other studies demonstrating the low reliability of RT-PCR in general. In one study of 8,240 patients, 42 types of virus were initially identified, but on further investigation it was found that 23 were due to contamination in laboratory reagents or from the surrounding environment.[16] In another study, a survey of PCR users showed that 65% of respondents had experienced failure to repeat their own results.[3] While highly experienced PCR users working under ideal laboratory conditions may consistently be able

to obtain the high sensitivities and specificities reported, it is unlikely these high values would be widely obtained in real-world practice.

4. PPV and NPV Impacts in Various Situations

(a) Australia

According to Worldometer on 8 March 2020 the total number of cases reported in Australia was 29,046 and the total number of tests carried out was 14,663,141.[17] In order to estimate the tested population from the number of tests it is assumed the average number of tests per person is 1.5 (i.e. every second person is tested twice). On this basis the tested population is 9,775,427.

So p = 29,046, T = 9,775,427, 1 – p/T = 0.997 or 99.7%

From section 2, the criteria for the two-way table to have a solution is $S_p > 1 - p/T$, which means the specificity must be greater than 99.7%. Given all the uncertainties and sources of error outlined in the previous section, it is not realistic to be able to achieve these high levels of specificity in practice. A significantly higher number of cases would be expected based on false positives alone, which suggests the reported results are not reliable.

(b) Victorian "snap" lockdown

During February 2021, Victoria was placed on a snap lockdown based on a reported five new cases of SARS-CoV-2, taking the total number of cases to 20,465.[18] The total number of tests carried out at the time was 4,652,545. Assuming, as above, the average number of tests per person was 1.5, the tested population is 3,101,697. The parameters are then:

p = 20,465, T = 3,101,697, 1 – p/T = 0.993 or 99.3%

In this case a specificity equal to or greater than 99.3% must be assumed in order for the two-way table to have non-negative solutions. This is considered too high to be achievable in practice. Even if it was achieved, the PPV would be zero, because all the reported cases would have to be false positives (if any of the 20,465 cases were true positives then that would lower the number of false positives and increase the required specificity to achieve it). These finding suggest the reported numbers of tests and cases are not reliable.

(c) New South Wales

As of 11 May 2021 the number of cases reported in New South Wales was 5357 and the total number of tests carried out was 5,803,433.[19] Assuming an average number of tests per person of 1.5, the tested population is 3,868,955. The parameters are:

p = 5357, T = 3,868,955, p/T = 0.0014 or 0.14%, 1 – p/T = 0.9986 or 99.86%

A specificity of 99.86% is too high to be consistently achievable in practice, suggesting the numbers reported are not reliable.

(d) Western Australian lockdowns

Western Australia has been locked down twice in 2021. The first lockdown was for 5 days, commencing on 1 February and was based on only a single reported case. Subsequently, it was reported that after more than 28,000 tests no-one else had tested positive.[20,21] The second lockdown commenced on 24 April and lasted for three days. It was also based on only one reported case. It was reported after the first day of lockdown that 12,000 tests had been carried out and only one additional positive case had been reported.[22] These numbers convert to specificities of 100% and 99.99% which are too high to be plausible in practice, suggesting the numbers being reported are not reliable.

(e) Pfizer safety and efficacy study

According to the Pfizer mRNA Covid-19 Vaccine trial published in the New England Journal of Medicine on 31 December 2020, the efficacy calculations were based on a total of 170 cases.[23] Eight cases came from the vaccinated group and 162 from the control group. The total number of at-risk participants were 17,411 and 17,511 respectively. It is assumed that all the participants are tested.

The relevant parameters for the vaccinated group are:

$p = 8$, $T = 17{,}411$, $1 - p/T = 0.9995$ or 99.95%

This specificity is too high to be plausible suggesting the reported results are unreliable.

The relevant parameters for the control group are:

$p = 162$, $T = 17{,}511$, $1 - p/T = 0.991$ or 99.1%

It is considered unlikely this specificity would be achieved in practice, but even if it was the PPV would be zero because all the positive cases would be false positives (if any of the 162 cases were true positives then that would lower the false positive rate and hence raise the specificity). This suggests the reported results for the control group are not reliable.

(f) AstraZeneca safety and efficacy study

The AstraZeneca safety and efficacy study was published in the Lancet on 8 December 2020.[24] For all low dose and high dose participants in the vaccinated groups 30 positive cases were reported out of a total of 5,807 participants. In the control group 101 positive cases were reported out of 5,829 participants.

The relevant parameters for the vaccinated group are:

$p = 30$, $T = 5807$, $1 - p/T = 99.5\%$

This is considered to high a specificity to be achieved in practice suggesting the reported results for the vaccinated group are not reliable.

The relevant parameters for the control group are:
$p = 101$, $T = 5829$, $1 - p/T = 0.983$ or 98.3%

Although it is unlikely to be achieved in practice, even if we assume a specificity as high 99.0% is achievable, this results in a PPV of 42.7% which is too low for the test to be of any practical value. The NPV corresponding to a specificity of 99% is 99.99%.

5. Conclusion

Published values for sensitivity and specificity for RT-PCR tests for SARS-CoV-2 have been obtained, and it has been attempted to calculate the PPV and NPV for Australia, Victoria, New South Wales, Western Australia, and the Pfizer and AstraZeneca trial cohorts. In the cases of Australia, Victoria, New South Wales and Western Australia, the ratio of positive results to total tests was too small to solve for PPV and NPV using plausible values of specificity.

The specificities required for the two-way tables to have non-negative solutions were greater than 99%, too high to be consistently achievable in practice. Even if the genuine incidence of SARS-CoV-2 was zero, with mass testing we would expect to see significantly higher numbers of positive test results based on false positives alone. In the cases of the vaccine trials, the ratio of positive results to total tests was either too small or the PPV was calculated and was found to be less than 50%, too low for the RT-PCR test to be of any practical value. These results highlight the unreliability of test results obtained from mass testing.

PPV, sensitivity, and specificity are important parameters in testing for medical conditions as they give an indication of the reliability of test results, yet they appear to have been overlooked when evaluating the data arising from the implementation of SARS-CoV-2 mass testing. When the numbers of positive results obtained from mass testing are less than the expected number of false positives based on real-world achievable specificities, the two-way table for calculating PPV has no solution, and the data cannot be considered reliable. Decisions to implement lockdowns and other measures based on data obtained from mass testing are ill-informed and irresponsible.

References

1. CDNA National guidelines for public health units, *Coronavirus Disease 2019 (COVID-19)*
2. WHO Information Notice for IVD Users 2020/05, "Nucleic acid testing (NAT) technologies that use polymerase chain reaction (PCR) for detection of SARS-CoV-2", 20/1/2021
3. Bustin S, Nolan T. Talking the talk, but not walking the walk: RT-qPCR as a paradigm for the lack of reproducibility in molecular research. European Journal of Clinical Investigation 2017; 47
4. Van Kasterena PB, Van der Veera B, Van den Brinka S, Wijsmana L, de Jongea J , Van den Brandta A, Molenkampb R , Reuskena CBEM, Adam Meijera A. Comparison of seven commercial RT-PCR diagnostic kits for COVID-19. *Journal of Clinical Virology* 2020; 128
5. Lieberman JA, Pepper G, Naccache SN, Huang ML, Jerome KR, Greningera AL. Comparison of commercially available and laboratory developed assays for in vitro detection of SARS-CoV-2 in clinical laboratories. *Journal of Clinical Microbiology* 2020; 58
6. Jaafar R, Aherfi S, Wurtz N, Grimaldier C, Hoang VT, Colson P, Raoult D, La Scola B. Correlation Between 3790 Quantitative Polymerase Chain Reaction–Positives Samples and Positive Cell Cultures, Including 1941 Severe Acute Respiratory Syndrome Coronavirus 2 Isolates 2020. *Clinical Infectious Diseases2020;* DOI: 10.1093/cid/ciaa1491
7. World Freedom Alliance, Copenhagen Wake Up Call. (Interview - quote is at 29 min 45 sec) https://www.youtube.com/watch?app=desktop&v=SEsjvkIJlZs&feature=share
8. https://www.cdc.gov/vaccines/covid-19/health-departments/breakthrough-cases.html Accessed 11 June 2021
9. Corman V, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DKW, Bleicker T, Brunink S, Schneider J, Schmidt ML, Daphne Mulders DGJC, Haagmans BL, Van der Veer B, Van den Brink S, Wijsman L, Goderski G, Romette J-L, Ellis J, Zambon M, Peiris M, Goossens H, Reusken C, Koopmans MPG, Drosten C. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveillance* 2020; 25: 23-30. https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045
10. Borger P, Malhotra BR , Yeadon M, Craig C, McKernan K , Steger K, McSheehy P, Angelova L, Franchi F, Binder T, Ullrich H, Ohashi M, Scoglio S, Doesburg-van Kleffens M, Gilbert D, Klement R, Schruefer R, Pieksma BW, Bonte J, Carbonare BHD, Corbett KP, Kämmerer U. External peer review of the RTPCR test to detect SARS-CoV-2 reveals 10 major scientific flaws at the molecular and methodological level: consequences for false positive results. https://cormandrostenreview.com/report/
11. Australian Public Health Laboratory Network guidance on nucleic acid test result interpretation for SARS-CoV-2. https://www1.health.gov.au/internet/main/publishing.nsf/Content/Publications-13
12. Doherty Institute. Post-market validation of the Beijing Genomics Institute (BGI) SARS-CoV-2 Real Time PCR platform. https://www.health.gov.au/sites/default/files/documents/2020/06/post-market-validation-of-the-beijing-genomics-institute-bgi-sars-cov-2-real-time-pcr-platform.pdf
13. Skittrall JP, Wilson M, Smielewska AA, Parmar S, Fortune MD, Sparkes D, Curran MD, Zhang H, Jalal H. Specificity and positive predictive value of SARS-CoV-2 nucleic acid amplification testing in a low-prevalence setting. *Clinical Microbiology and Infection* 2021; 27
14. Tran T, Papadakis G, Mackenzie C, Druce, J, Catton M. Post-market validation of the Beijing Genomics Institute (BGI) SARS-CoV-2 Real Time PCR platform. Report prepared for: Office of Health Protection, Commonwealth Government of Australia The Therapeutics Goods Administration (TGA) of Australia 10 June 2020.
15. Watson J, Whiting PF, Brush JE. Interpreting a covid-19 test result. *BMJ* 2020; 369
16. Moustafa A, Xie C, Kirkness E, Biggs W, Wong E, Turpaz Y, Bloom K, Delwart E, Nelson KE , Venter JC, Telenti A. The blood DNA virome in 8,000 humans. *PLOS Pathogens* 2017
17. https://www.worldometers.info/coronavirus/ (accessed 8 Mar 2021)
18. https://www.dhhs.vic.gov.au/coronavirus-update-victoria-12-february-2021
19. https://www.nsw.gov.au/covid-19/find-the-facts-about-covid-19
20. https://www.abc.net.au/news/2021-02-02/no-new-coronavirus-cases-in-wa-overnight-as-lockdown-continues/13112234
21. https://www.abc.net.au/news/2021-02-03/wa-lockdown-to-end-on-friday-night-if-no-coronavirus-covid-cases/13116272
22. https://www.sbs.com.au/news/western-australia-records-no-new-locally-acquired-covid-19-cases-amid-perth-and-peel-lockdown

23. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, Perez JL, Marc GP, Moreira ED, Zerbini C, Bailey R, Swanson KA, Roychoudhury S, Koury K, Li P, Kalina WV, Cooper D, Frenck Jr RW, Hammitt LL, Tureci O, Nell H, M.D, Schaefer A, Unal S, Tresnan DB, Mather S, Dormitzer PR, Şahin U, Jansen KU, Gruber WC. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine* 2020; 383
24. Voysey M, Costa Clemens SA, Madhi SA, Weckx LW, Folegatti PM, Aley PK, Angus B, Baillie VL,Barnabas SL, Bhorat QE, Bibi S, Briner C, Cicconi P, Collins AM, Colin-Jones R, Cutland CL, Darton TC, Dheda K, Duncan CJA, Emary KRW, Ewer KJ, Fairlie L, Faust SN, Feng S, Ferreira DM, Finn F, Goodman AL, Green CM, Green CA, Heath PT, Hill C, Hill H, Hirsch I, Hodgson SHC, Izu A, Jackson S, Jenkin D, Joe CCD, Kerridge S, Koen A, Kwatra G, Lazarus R, Lawrie AM, Lelliott A, Libri V, Lillie PJ, Mallory R, Mendes AVA, Milan EP, Minassian AM, McGregor A, Morrison H, Mujadidi YF, Nana A, O'Reilly PJ, Padayachee SD, Pittella A, Plested E, Pollock KM, Ramasamy MN, Rhead S, Schwarzbold AV, Singh N, Smith A, Song R, Snape MD, Sprinz E, Sutherland RK, Tarrant R, Thomson EC, Török ME, Toshner M, Turner DPJ, Vekemans J, Villafana TL, Watson MEE, Williams CJ, Douglas AD, Hill AVS, Lambe T, Gilbert SC, Pollard AJ. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet* 2020. https://doi.org/10.1016/S0140-6736(20)32661-1