Title: Clinician's Probability Calculator to convert pre-test to post-test probability of COVID-19 based on method validation from each laboratory

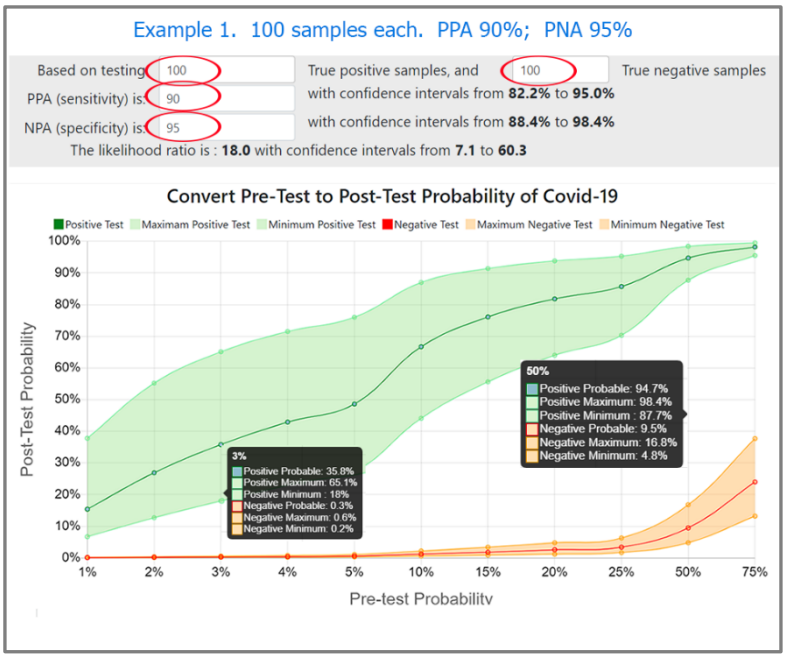
Authors: Zoe Brooks[1], Saswati Das [2], Tom Pliura[3]

1. Awesome Numbers, Worthington, ON, Canada
2. Dr Ram Manohar Lohia Hospital & ABVIMS, New Delhi, India
3. The Medical and Law Offices of Thomas J. Pliura, M.D., J.D., Le Roy, Illinois, United States

Abstract

Identifying the SARS-CoV-2 virus has been a unique challenge for the scientific community. In this paper, we discuss a practical solution to help guide clinicians with interpretation of the probability that a positive, or negative, COVID-19 test result indicates an infected person, based on their clinical estimate of pre-test probability of infection.

The authors conducted a small survey on LinkedIn to confirm that hypothesis that that the clinical pre-test probability of COVID-19 increases relative to local prevalence of disease plus patient age, known contact, and severity of symptoms. We examined results of PPA (Positive Percent Agreement, sensitivity) and NPA (Negative Percent Agreement, specificity) from 73 individual laboratory experiments for molecular tests for SARS-CoV-2 as reported to the FIND database [1], and for selected methods in FDA EUA submissions[2,3]. Authors calculated likelihood ratios to convert pre-test to post-test probability of disease and designed an online calculator to create graphics and text to report results.



Despite best efforts, false positive and false negative Covid-19 test results are unavoidable[4,5]. A positive or negative test result from one laboratory has a different probability for the presence of disease than the same result from another laboratory. Likelihood ratios and confidence intervals can convert the physician or other healthcare professional's

clinical estimate of pre-test probability to post-test probability of disease. Ranges of probabilities differ depending on proven method PPA and NPA in each laboratory. We recommend that laboratories verify PPA and NPA and utilize a the "Clinician's Probability Calculator" to verify acceptable test performance and create reports to help guide clinicians with estimation of post-test probability of COVID-19.

Keywords: COVID-19; SARS-COV-2; False-Positive; False-Negative; Likelihood Ratio; Probability; Calculator; Interpretation

Introduction:

During the coronavirus (COVID-19) outbreak, accurate testing has been a unique and constant challenge for the scientific community. Despite best efforts, false positive and false negative test results are unavoidable [4,5]. At no time in history has the medical community embarked on a diagnostic testing campaign that is not being pursued for clinical reasons, but instead for epidemiological reasons unrelated to the medical aspects of the illness. With the importance of asymptomatic patients, understanding of the probability of true and false results has never been so critical. This article explores a modified application of likelihood ratios to provide practical guidance to the relative probability of true and false results.

What are pre-test and post-test probability?

Pre-test probability and post-test probability are the probabilities of the presence of a disease (such as COVID-19) before (pre) and after (post) a diagnostic test [6]. With COVID-19 and the importance of pre-symptomatic or asymptomatic patients however, tests are often performed even with low pre-test probability. Although sometimes confused with simple prevalence of disease[7], the clinical pre-test probability of disease can be more precisely estimated with clinical information on each patient. In a small survey with 16 respondents on LinkedIn, the authors asked healthcare professionals to "estimate the probability that a 20-to-30-year-old patient, and 60-to-70-year-old patient, actually has, or will soon develop COVID-19 infection?" Symptoms ranged from i) none, to ii) sore throat, and nasal stuffiness, to iii) sore throat, and nasal stuffiness, with reduced taste or smell to iv) sore throat, and nasal stuffiness, with reduced taste or smell, fever, and body ache. We found, as logic and clinical experience would dictate, that pre-test probability increases with local COVID-19 prevalence, patient age, SARS-COV-2 exposure history and clinical symptoms.

Post-test probability is driven by pre-test probability and the likelihood ratio of the test method – as that method was verified in each laboratory. The likelihood ratio is driven by test PPA and NPA. PPA and NPA vary between methods as reported by manufacturers to FDA for EUA evaluation, and between laboratories using the same method.

What are likelihood ratios?

The likelihood ratio is a tool used in evidence-based medicine to assess the value of performing a diagnostic test. It uses PPA and NPA to create a ratio of the probability that a test result is correct to the probability that it is not. A likelihood ratio is the percentage of ill people with a given test result divided by the percentage of well individuals with the same result (true result: false result). Ideally, abnormal test results should be much more typical in ill individuals than in those who are well (high likelihood ratio) and normal test results should be more frequent in well people than in sick people (low likelihood ratio). Likelihood ratios near one have little effect on decision-making; by contrast, high or low ratios can greatly shift the clinician's estimate of the probability of disease. When combined with an accurate clinical diagnosis, likelihood ratios improve diagnostic accuracy in a synergistic manner[10,11].

Tests can be either positive or negative, so there are two ratios:

- Positive LR (LR+): This tells us how much to increase the probability of having a disease, given a positive test result. The ratio is:
  "Probability a person with the condition tests positive (a true positive) /
  probability a person without the condition tests positive (a false positive)[10].
- Negative LR (LR-): This tells us how much to decrease the probability of having a disease, given a negative test result. The ratio is:
  "Probability a person with the condition tests negative (a false negative) /
  probability a person without the condition tests negative (a true negative)[10,11]."
  Likelihood ratios are calculated from PPA and NPA:
  - Positive LR = (PPA / (100 – NPA) (True Positives / False Positives)
  - Negative LR = (100 – PPA) / NPA (False Negatives/True Negatives)

Likelihood ratios are calculated to determine 2 things: i) how useful a diagnostic test is and ii) how likely it is that a patient has a disease[10]. Likelihood ratios range from zero to infinity (9999.9). The higher the value, the more likely the test will indicate that the patient has the condition.

What are Confidence Intervals?

"A confidence interval gives an estimated range of values which is likely to include an unknown population parameter. [10,11]" Confidence intervals provide a range of possible results: minimum, probable and maximum. They tell the end-user how much faith they can have in the value reported.

Methodology

1. We created a LinkedIn survey asking for "your estimate that a 20-to-30-year-old patient (or a 60-to-70-year-old patient) actually has, or will soon develop, COVID-19 infection.
2. We used the calculations and definitions in Table 1 to examine results of individual laboratory experiments for molecular tests for SARS-CoV-2as reported to the FIND database[1], and for selected methods in FDA EUA submissions[2,3].
3. The authors created an Excel spreadsheet and then designed an online application to graph confidence intervals of post-test probability of infection on with a positive or negative test result (on the y-axis) against the clinician's estimate of pre-test probability (on the x-axis.) Confidence intervals for the graph and each of the following indicators are driven by PPA and PNA reported, plus the number of samples tested[14]:
   a) Post-test probability of COVID-19 with positive and negative test result.
   b) Number of true positive and negative tests in every ten positive or negative results seen.
   c) Number of false positive and negative tests in every 10 positive or negative results seen
   d) One in 'x' positive tests is true, and one in 'x' negative tests is True.

Table 1. Common definitions		
PPA	Positive Percent Agreement (sensitivity) = True positive results / All positive results Drives the True Positive and False Negative rates	
NPA	Negative Percent Agreement (specificity) = True negative results / All negative results Drives the True Negative and False Positive rates	
Pre-test probability	Clinical probability that a person being tested has COVID-19, before the test is performed. Based on prevalence, contact, symptoms and age.	
Pre-test odds	= Pre-test probability/(1-Pre-test probability) = Probability person is infected/ Probability they are not	
Calculations	**Positive Test**	**Negative Test**
Likelihood Ratio	Positive (LR+) = PPA / (1 - NPA) = True Pos Rate/False Pos Rate	Negative (LR-) = (1-PPA)/NPA = False Neg Rate/True Neg Rate
Post-test odds	= Pre-test odds x LR+ = Probability person is infected x (True Pos Rate/False Pos Rate)	= Pre-test odds x LR- = Probability person is infected x (False Neg Rate/True Neg Rate)

Post-test probability	= Post test odds Pos Test / (Post-Test Odds Pos +1) The probability that a person with a positive test is infected	= Post test odds Neg Test / (Post-Test Odds Neg Test +1) The probability that a person with a negative test is infected

Results:

1. LinkedIn Survey:

Seventeen people responded to a LinkedIn survey asking for "your estimate that a 20-to-30-year-old patient (or a 60-to-70-year-old patient) actually has, or will soon develop Covid19 infection" - with escalating Covid19 symptoms, with and without KNOWN contact." The local prevalence was given as 3%. (Questionnaire attached as supplementary material).

Table 2 presents the survey results, showing a clear pattern that pre-test probability increases with patient age, known contact and presence of typical COVID-19 symptoms.

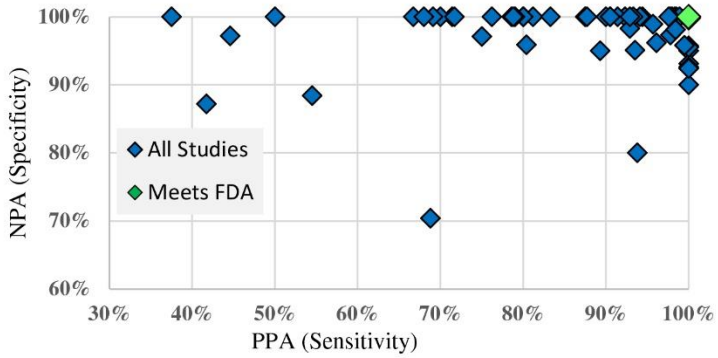
Table 2. Median Votes of Pre-Test Probability from LinkedIn Survey	No Symptoms	sore throat, and nasal stuffiness	sore throat, and nasal stuffiness, with reduced taste or smell	sore throat, and nasal stuffiness, with reduced taste or smell, fever, and body ache
Without Contact - Young	3%	35%	79%	85%
Without Contact - Old	5%	50%	80%	90%
WITH Contact - Young	40%	77%	87%	96%
WITH Contact - Old	45%	60%	94%	98%

2. Results reported to FIND and FDA- with calculations:

Ninety-two laboratories reported both PPA and NPA to the FIND database[1] as of October 17, 2020. We removed 19 results from two laboratories in 1 country that reported NPA of 100% based on only one negative. **Figure 1** shows the number of known positive and negative samples reported in each laboratory's study. Fifteen of the laboratories (21%) tested fewer than

Figure 1. Known Samples Tested

Figure 2. PPA and NPA by Study



five known negative samples.

Figure 2 shows the PPA and NPA for 73 studies reported to FIND. Only 15 studies (20.5%) met the FDA recommendations "FDA defines the acceptance criteria for the performance as 95% agreement at 1x-2x LoD, and 100% agreement at all other concentrations and for negative specimens.[12]"

The authors compared the Information For Use (IFU) documents provided to FDA for two manufacturers[2,3], to five FIND[1] laboratory studies for manufacturer 1 (Mfg-1) and six FIND laboratory studies for manufacturer 2 (Mfg-2.) To calculate PPA, Positive Percent Agreement, Mfg-1 tested "30 contrived clinical nasopharyngeal (NP) swabs prepared by spiking clinical NP swab matrix with purified viral RNA containing target sequences from the SARS-CoV-2 genome at concentrations approximately 2x LOD (20 samples) and 5x LOD (10 samples)." Mfg-2 tested "45 patient samples collected during COVID-19 pandemic in the US that had previously been characterized as positive for SARS-CoV-2by an EUA RT-PCR test." Mfg-1 reported PPA of 100% (30/30). Mfg-2 reported PPA of 97.8% (44/45).

To prove NPA, Negative Percent Agreement, Mfg-1 reported that "Thirty Negative NP swab samples were also tested in this study." Mfg-2' IFU reported testing 45 samples, saying "Fifteen of the 45 SARS-CoV-2negative NP swab specimens were collected before December 2019 and are expected to be negative for SARS-CoV-2. The others had previously been characterized as negative for SARS-CoV-2by an EUA RT-PCR test." Mfg-1 reported NPA of 100% (30/30). Mfg-2 reported NPA of 95.6% (43/45).

These results were driven by each manufacturer's test methods, the type and number of samples tested plus the competency of staff and instrument performance at the manufacturers' sites.

Figure 3 shows the reported PPA on the x-axis, and NPA on the y-axis, by Manufacturers 1 and 2, and in the FIND studies for all labs reporting studies from the same manufacturers. The circles representing 'Manufacturer 1' labs are coloured blue; the manufacturer is shown as the clear circle. The diamonds represent 'Manufacturer 2'; labs are coloured yellow; the manufacturer is shown as the clear diamond. Labs A, B, C and D are examined in greater detail in Tables 3, 4, 5 and 6.

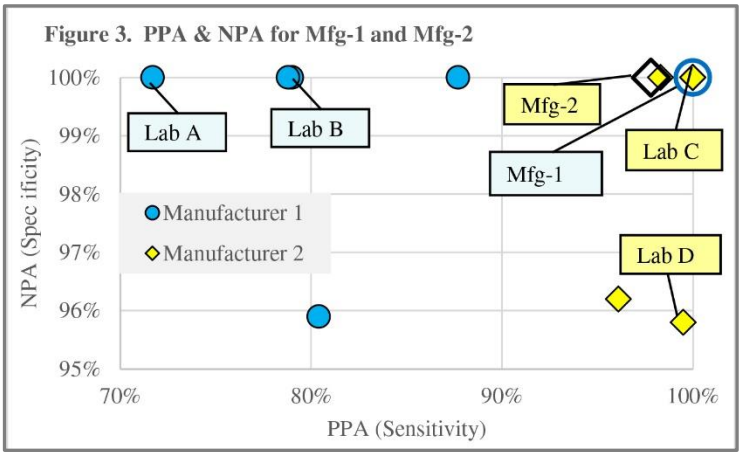


Figure 4 shows the number of known samples tested by Manufacturers 1 and 2, and in the FIND studies for the labs reporting studies from the same manufacturers.

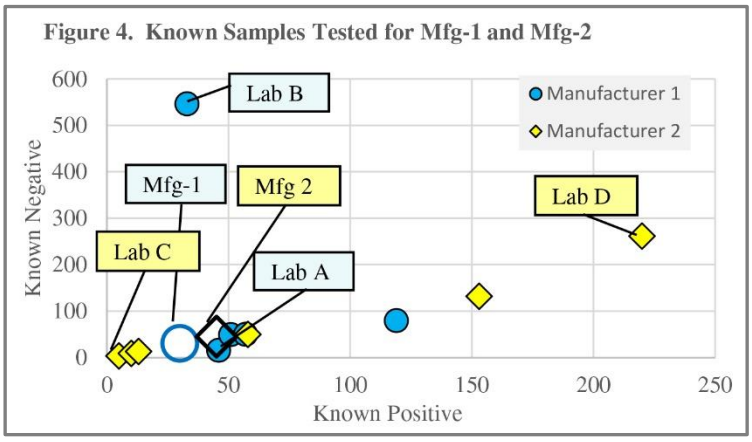


Table 3 presents the data reported by the selected manufacturers and two laboratories using each test method. The number of known samples that each laboratory tested drives the confidence intervals around PPA and NPA which, in turn, drive the confidence intervals around likelihood ratios, which drive post-test probability and other indicators. Notice the wide variation in the number of known samples tested in rows 1 and 2. In row 3, notice that the labs A and B reported PPA values below Mfg-1 while Labs C and D reported higher PPA than Mfg-2.

Table 3. Reported values		Mfg-1	Lab A	Lab B	Mfg-2	Lab C	Lab D
1	Known Positive Samples	30	46	33	45	5	220
2	Known Negative Samples	30	15	546	45	3	261
3	PPA Reported	100%	71.7%	78.8%	97.8%	100%	99.5%
4	NPA Reported	100%	100%	100%	95.6%	100%	95.8%
5	Pos Likelihood Ratio (LR+)	9999.9	9999.9	9999.9	22.2	9999.9	23.7
6	Neg Likelihood Ratio (LR-)	0.00	0.28	0.21	0.023	0.000	0.005

Table 4 presents the probable post-test interpretation of results for patients with pre-test probability of 3% and for pre-test probability of 50%. Red numbers in the table indicate variation from others and/or less-than-ideal test performance. Notice in Row 4 that a positive test result from Mfg-2 and Lab D indicates less than a 50% post-test probability of disease, when pre-test probability is only 3%. On row 7, notice that only 4 of 10 positive tests are true. This relates in Row 10 to the fact that only one of every 2.5 positive tests seen is true. When pre-test probability raises to 50%, (rows 12-21) post-test probability rises to over 90% for these labs. Even with a negative test, post-test probability is approximately 20% in Labs A and B (row 15). In row 18, with pre-test probability of 50%, only 8 of 10 negative tests are true for Labs A and B.

	Table 4. Probable Post-Test Interpretation of Results						
1	**Post-Test Projections: With 3% Pre-Test Probability**						
2	Post-test Probability of COVID-19 (Ideal is 100% with positive test; 0% with negative test)						
3	Reported values	**Mfg-1**	**Lab A**	**Lab B**	**Mfg-2**	**Lab C**	**Lab D**
4	With Positive Test	100%	100%	100%	41%	100%	42%
5	With Negative Test	0.0%	0.9%	0.7%	0.1%	0.0%	0.0%
6	Number of True Results in every Ten Tests Reviewed (Ideal is ten of ten)						
7	Positive Test	10.0	10.0	10.0	4.1	10.0	4.2
8	Negative Test	10.0	9.9	9.9	10.0	10.0	10.0
9	One in x Test results is/are True (Ideal is one in one)						
10	Positive Test	1.0	1.0	1.0	2.5	1.0	2.4
11	Negative Test	1.0	1.0	1.0	1.0	1.0	1.0
12	**Post-Test Projections: With 50% Pre-Test Probability**						
13	Post-test Probability of COVID-19 (Ideal is 100% with positive test; 0% with negative test)						
14	With Positive Test	100%	100%	100%	96%	100%	96%
15	With Negative Test	0.0%	22.1%	17.5%	2.2%	0.0%	0.5%
16	Number of True Results in every Ten Tests Reviewed (Ideal is ten of ten)						
17	With Positive Test	10.0	10.0	10.0	9.6	10.0	9.6
18	With Negative Test	10.0	7.8	8.3	8.7	10.0	9.9
19	One in 'x' Test results is/are True (Ideal is one in one)						
20	Positive Test	1.0	1.0	1.0	1.0	1.0	1.0
21	Negative Test	1.0	1.3	1.2	1.0	1.0	1.0

Table 5 presents confidence intervals for PPA, NPA and likelihood ratios. Numbers in red in the table differ from others, are far from ideal, and/or do little to assist the clinician with diagnosis.

	Table 5. Confidence Limits for Drivers of Post-Test Interpretation						
		Mfg-1	**Lab A**	**Lab B**	**Mfg-2**	**Lab C**	**Lab D**
1	PPA Low Confidence Interval	88.0%	56.6%	61.2%	87.8%	54.6%	97.3%
2	PPA High Confidence Interval	100.0%	83.8%	90.9%	100.0%	100.0%	100.0%
3	NPA Low Confidence Interval	88.0%	78.5%	99.3%	84.5%	41.9%	92.5%
4	NPA High Confidence Interval	100.0%	100.0%	100.0%	100.0%	100.0%	97.9%
5	LR+ From Low PPA NPA	7.33	2.63	82.60	7.33	0.94	12.91
6	LR+ From High PPA NPA	999.9	999.9	999.9	999.9	999.9	999.9
7	LR- From Low PPA, NPA	0.14	0.55	0.39	0.14	1.08	0.39
8	LR- From High PPA NPA	0.00	0.16	0.09	0.00	0.00	0.09

Table 6 shows the range of possibilities with low and high confidence intervals.

Table 6. Range of Possible Interpretation of Test Results with Confidence Intervals

1	**Post-Test Projections: With 3% Pre-Test Probability**						
2	Post-test Probability of Covid-19 (Ideal is 100% with positive test; 0% with negative test)						
3	**Range of Possible Values**	**Mfg-1**	**Lab A**	**Lab B**	**Mfg-2**	**Lab C**	**Lab D**
4	With Positive Test- Low	18.5%	7.5%	71.9%	14.9%	2.8%	28.5%
5	- High	100%	100%	100%	100%	100%	59%
6	With Negative Test - Low	0.00%	0.50%	0.30%	0.00%	0.00%	0.00%
7	- High	0.40%	1.70%	1.20%	0.40%	3.20%	0.10%
8	Number of True in Results in every Ten Tests Reviewed (Ideal is ten)						
9	Positive Test- Low	1.8	0.8	7.2	1.5	0.3	2.9
10	- High	10	10	10	10	10	5.9
11	Negative Test - Low	10	9.8	9.9	10	9.7	10
12	- High	10	10	10	10	10	10
13	One in every 'x' Test results is/are True (Ideal is one)						
14	Positive Test- Low	1.0	1.0	1.0	1.0	1.0	1.7
15	- High	5.4	13.3	1.4	6.7	35.4	3.5
16	Negative Test - Low	1.0	1.0	1.0	1.0	1.0	1.0
17	- High	1.0	1.0	1.0	1.0	1.0	1.0
18	**Post-Test Projections: With 50% Pre-Test Probability**						
19	Post-test Probability of Covid-19 (Ideal is 100% with positive test; 0% with negative test)						
20	With Positive Test- Low	88.0%	72.4%	98.8%	85.0%	48.5%	92.8%
21	- High	100%	100%	100%	100%	100%	98%
22	With Negative Test - Low	0.0%	13.90%	8.3%	0.0%	0.0%	0.0%
23	- High	12.0%	35.6%	28.1%	12.6%	52.0%	2.9%
24	Number of True Results in every Ten Tests Reviewed (Ideal is ten)						
25	Positive Test- Low	8.8	7.2	9.9	8.5	4.8	9.3
26	- High	10.0	10.0	10.0	10.0	10.0	9.8
27	Negative Test - Low	8.8	6.4	7.2	8.7	4.8	9.7
28	- High	10	8.6	9.2	9.8	10	10
29	One in every 'x' Test results is/are True (Ideal is one)						
30	Positive Test- Low	1.0	1.0	1.0	1.0	1.0	1.0
31	- High	1.1	1.4	1.0	1.2	2.1	1.1
32	Negative Test - Low	1.0	1.2	1.1	1.0	1.0	1.0
33	- High	1.1	1.6	1.4	1.1	2.1	1.0

Table 5, in rows 5 and 7, shows that Lab C has a low possible positive likelihood ratio of 0.94 and a high possible negative likelihood ratio of 1.08. When pre-test probability is 50%, the odds are 1:1 that the patient is infected. Multiply the pre-test odds times the likelihood ratio to calculate post-test odds – which

will be essentially unchanged in this case. The low range of post-test possibility with a positive test overlaps the high possibility with a negative test (Table 6 Rows 4 & 7, 20 & 24.) A positive or negative test result, as verified in Lab C, may not be able to differentiate infected from non-infested patients. This is not due to an inherent weakness in the test method, but to the low number of samples used by Lab C to verify method performance in their hands.

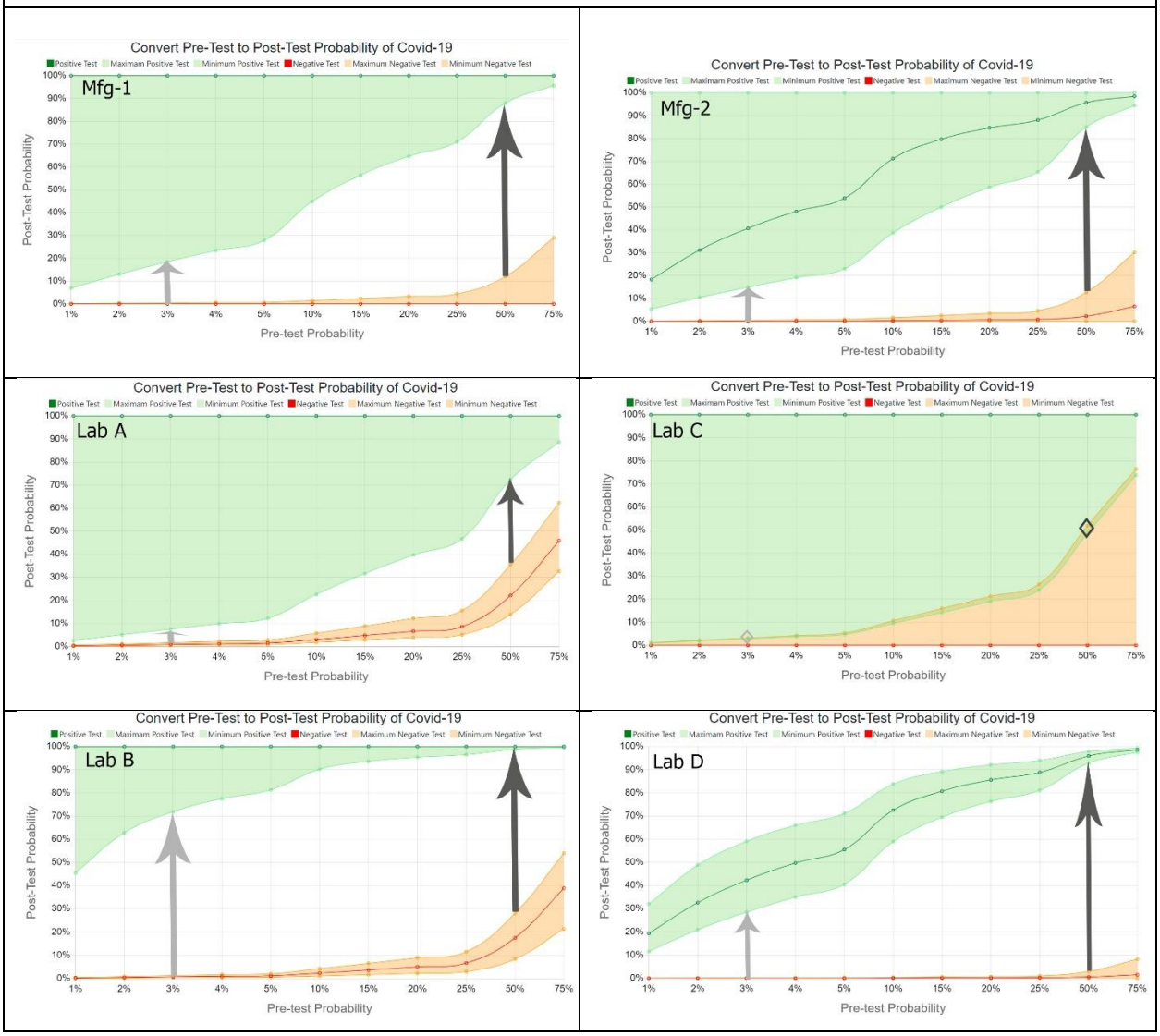
Where Table 4 presented the 'Probable' post-test interpretation of results, Table 6 shows the range of possibilities with low and high confidence intervals. Notice in row Table 4 - Row 4 that, where Mfg-1 and Lab A both reported 100% PPA, confidence intervals show that could actually be as low as 18% or 7.5%; in Lab C, the reported 100% may actually be as low as 2.8% due to the low number of samples tested. Row 9 shows that there may be less than two or three true results in every ten positives reviewed by clinicians.

3. **The authors designed an online 'Clinician's Probability Calculator' to create a report that could accompany each laboratory's test results to show clinicians the post-test probability of Covid-19 based on their clinical opinion of each patient's pre-test probability.** Each laboratory's report would vary based on their chosen methodology and the data (PPA and PNA) generated by their on-site method validation studies (with sufficient samples.) The calculator overcomes the complexity of calculations that prohibit most laboratories from reporting post-test probability with confidence intervals. It is available at https://awesome-numbers.com/post-test-probability-calculator/. Users provide the number of known samples tested plus PPA and PNA determined; reports contain data as shown in Tables 3, 4, 5 and 6, with the figures with confidence intervals as in Figure 5.

   In the graphic, the x-axis is the pre-test probability, as estimated by the clinician or public health professional. The Y-axis is the post-test probability. The shaded green area shows confidence intervals that a positive COVID-19 test result indicates an infected person. The pale orange shaded area shows confidence intervals that negative COVID-19 indicates an infected person (false negatives.)

   Figure 5 shows Probability Calculator graphs from Manufacturer 1 and 2, plus Labs A and B who reported data from Manufacturer 1, and Labs C and D who reported data from Manufacturer 2. The arrows show the gap between the highest probability that a negative test represents an infected person and lowest probability with a positive test.

Figure 5. Probability Calculator results from Manufacturer 1 and 2, plus Labs A, B, C and D



Discussion:

Importance of PPA (sensitivity) and NPA (specificity):

PPA, Positive Percent Agreement (sensitivity), drives the rate of true positive and false negative test results. NPA, Negative Percent Agreement (specificity), drives the rate of true negative and false positive test results. PPA and NPA combine to drive the probability, number and cost of false-positive and -negative test results. (13) PPA and PNA are typically used by laboratory directors to compare inherent method quality and select test methods. They can also be used to calculate likelihood ratios that in turn

drive post-test probability of COVID-19 plus the graphs and other metrics displayed in Tables 3, 4, 5 and 6 and Figure 5.

Test methods are often verified by manufacturers under ideal conditions with hospital or contrived samples containing higher viral loads than those from asymptomatic individuals living in the community. As such, PPA and NPA in test laboratories might differ significantly from values reported by manufacturers. Notice in Figure-3 that none of the five laboratories reporting to FIND attained the 100% PPA claimed in the FDA IFU by Manufacturer 1[2]. In contrast, laboratories C and D reported higher PPA values than Manufacturer 2[3]. Thus, the PPA and NPA values reported by manufacturers cannot be assumed to accurately reflect performance in each laboratory.

PPA and NPA do not help clinicians decide if a specific positive or negative test result is true. PPA and NPA can be converted to likelihood ratios which can be used to convert clinical pre-test probability of disease for a specific patient to post-test probability.

Relevance of likelihood ratios:

Likelihood ratios allow one to convert pre-test to post-test odds of infection. The mathematics of this process are complicated, but the logic is clear. (15) When pre-test probability is 50%, the odds are 1:1 that the patient is infected. One of every two people with 'these' clinical symptoms is expected to be positive before testing (50%). If the positive likelihood ratio is approximately 24, as in Lab D in Table 3, multiplying the pre-test odds by the positive likelihood ratio produces post-test odds of 24:1. Twenty-four of every 25 people with a positive test are actually infected; 24/25 = 96% post-test probability.

Importance of number of known samples tested:

The number of known positive and negative samples tested determines the confidence intervals around PPA and NPA (14). Figure 1 and Table 3 illustrate the dramatic difference in number of samples tested by individual sites reporting to FIND (1). Labs A, B and C each reported 100% NPA. Lab A made that assessment by testing 15 known negative samples, while Lab B tested 546 known negatives and Lab C tested only three. The lower limit of confidence for NPA in Lab A is 78.5% compared to 99.3% in Lab B and only 41.9% in Lab C (Table 5.) The low number of known samples tested in Lab C does not allow this lab to verify acceptable method performance.

Impact of confidence intervals:

Confidence intervals determine the range of possibilities for PPA and NPA, which drive likelihood ratios that drive post-test probability of COVID-19 with positive and negative test results. Post-test probability

drives the number of true and false positive and negative tests in every 10 positive or negative results seen, and how many positive or negative test results would be seen to find one true test result. Confidence intervals allow users to visualize the gap between the post-test probability that a positive, or negative, test indicates an infected person.

Value of graphs and metrics reported by the Probability Calculator:

Instead of either taking all positive or negative test results at face value or developing personal experience to 'guess' if results are true or false, clinicians can visualize a reliable scientific range of possibilities. Glancing at the six graphs in Figure 5 clarifies that when pre-test probability is only 3%, the probability of a person having COVID-19, even with a positive test, is less than 50% - except in Lab B where they tested enough samples to prove test reliability. These graphics and data eliminate the use of Fagan's Nomogram[7], which is typically used with likelihood ratios but are cumbersome for front-line use and does not include confidence intervals.

Laboratory directors and public health officials who are challenged to select and verify test methods can clearly see the ability of each test to project, or rule out, COVID-19 infection. Lab C shows no gap at all between the range of possibilities that a positive or negative test indicates an infected person. The test has not been verified to provide useful information by testing only five known-positive and three known-negative samples. Laboratory directors and clinicians can have little confidence in the values reported.

Clinicians can benefit from understanding the number of true and false positive results they can expect to see in every ten positive or negative results. Clinicians may see as few as one or two true positives in every ten positive test results according to the Manufacturer-1 and Lab A, while Lab B can be relied on to produce over seven of ten true positives (Table 6 Row 9.) Knowing the frequency of true test results reported could certainly change test selection and interpretation. According to data from Mfg-1, clinicians may see only one true positive result in every 5.4 tests reviewed; for Lab A, that could be one in 13.4 results (Table 6, Row 15.) Lab B, who tested more known negative samples and has less variation due to confidence intervals, clinicians will see one true positive in every 1.4 tests. In contrast, Lab C may produce only one true positive in every 35 positive results. This information, however, is not available by only examining the reported PPA and NPA values.

In the USA, Clinical Laboratory Improvement Amendments (CLIA) mandates that laboratory director responsibilities include "ensuring that your laboratory develops and uses a quality system approach to laboratory testing that provides accurate and reliable patient test results. [16]" Accuracy is the number of true results as a portion of all test results created[17]. The authors were shocked to discover that most

laboratories are not required to verify that they can, at least, reproduce PPA and NPA claims from manufacturers. A Linked-In survey confirmed that 83% of respondents agreed that "Testing labs should confirm that they can attain or exceed manufacturer's claims for PPA and NPA, sensitivity & specificity, for COVID-19." This does not preclude the laboratory director from performing this study as part of good lab practice. In fact, Stephanie L. Mitchell et al published an article in the Journal of Clinical Microbiology[18] outlining a process to verify method PPA and NPA with ten positive and negative samples. In order to ensure method accuracy, we recommend that each testing laboratory confirm PPA and NPA with sufficient known samples to provide reliable post-test probability of disease. We concur that each report should be accompanied by a statement from the laboratory indicating that test performance has been verified. The Clinician's Probability Calculator fulfils this need.

Conclusion

Despite best efforts, false positive and false negative Covid-19 test results are unavoidable[4,5]. A positive or negative test result from one laboratory has a different probability for the presence of disease than the same result from another laboratory. Post-test probability, likelihood ratios and confidence intervals can help answer the question:" Does this person have Covid-19, or not?" by converting the physician or other healthcare professional's clinical estimate of pre-test probability to post-test probability. If the pre-test probability is below 5%, a positive test result may only raise that probability to less than 50%. A negative test with some methods in some labs may still convey a 20% post-test probability of disease. Ranges of probabilities differ depending on proven method PPA and NPA in each laboratory. The authors recommend that testing laboratories verify PPA and NPA of their Covid-19 method with sufficient sample numbers to verify acceptable performance and create a report with the Clinician's Probability Calculator (https://awesome-numbers.com/post-test-probability-calculator) to assist with interpretation of test results.

References:

1) FIND Foundation for Innovative New Diagnostic Test performance dashboard. Available from https://finddx.shinyapps.io/COVID19DxData/ Accessed Oct 17, 2020 (Data attached as supplementary material).

2) ID NOW COVID-19 For Use Under an Emergency Use Authorization – US only (EUA). Available from https://www.fda.gov/media/136525/download (Accessed Oct 10, 2020)

3) Instructions for Use For Use Under an Emergency Use Authorization (EUA) Only, Cepheid Xpert® Xpress SARS-CoV-2. Available from https://www.fda.gov/media/136314/download (Accessed Oct 10, 2020)

4) Surkova E, Nikolayevskyy V, Drobniewski F. False-positive COVID-19 results: hidden problems and costs [published online ahead of print, 2020 Sep 29]. *Lancet Respir Med*. 2020;8(12):1167-1168.

5) Mayers C, Baker K. Impact of false-positives and false-negatives in the UK's COVID-19 RT-PCR testing programme. June 3, 2020. Available from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/895843/S0519_Impact_of_false_positives_and_negatives.pdf (Accessed Aug 8, 2020).

6) Maxim LD, Niebo R, Utell MJ. Screening tests: a review with examples [published correction appears in Inhal Toxicol. 2019 Jun;31(7):298]. Inhal Toxicol. 2014;26(13):811-828.

7) Prinzi A.Why Pretest and Posttest Probability Matter in the Time of COVID-19 , American Society for Microbiology, June 2020. Accessed from https://asm.org/Articles/2020/June/Why-Pretest-and-Posttest-Probability-Matter-in-the

8) Fact Sheet For Healthcare Providers Id Now Covid-19. Abbott Diagnostics Scarborough, Inc. Updated: September 17, 2020. Available from https://www.fda.gov/media/136523/download

9) Fact Sheet For Healthcare Providers Cepheid A Xpert® Xpress SARS-CoV-2. August 8, 2020. Available from https://www.fda.gov/media/136313/download

10) Confidence Intervals. Available from http://www.stat.yale.edu/Courses/1997-98/101/confint.htm (Accessed Sept 15, 2020)

11) McGee S. Simplifying likelihood ratios. J Gen Intern Med. 2002;17(8):646-649.

12) Policy for Coronavirus Disease-2019 Tests During the Public Health Emergency (Revised). May 11, 2020. Available from https://www.fda.gov/media/135659/download (Accessed Sept 15, 2020)

13) Brooks Z, Das S. COVID-19 Testing: Impact of Prevalence, Sensitivity, and Specificity on Patient Risk and Cost, American Journal of Clinical Pathology. 2020; 154(5):575–584

14) Confidence Intervals for One-Sample Sensitivity and Specificity. Available from https://ncss-wpengine.netdna-ssl.com/wpcontent/themes/ncss/pdf/Procedures/PASS/Confidence_Intervals_for_One-Sample_Sensitivity_and_Specificity.pdf (Accessed Sept 15, 2020)

15) Likelihood ratios and probability of infection in a tested individual. Available from https://epitools.ausvet.com.au/probabilityofinfection (Accessed Sept 15, 2020)

16) Clinical Laboratory Improvement Amendments (CLIA) Laboratory Director Responsibilities Available from https://www.cms.gov/Regulations-and-Guidance/Legislation/CLIA/Downloads/brochure7.pdf

17) Baratloo A, Hosseini M, Negida A, El Ashal G. Part 1: Simple Definition and Calculation of Accuracy, Sensitivity and Specificity. *Emerg (Tehran)*. 2015;3(2):48-49.

18) Mitchell SL, St George K, Rhoads DD, et al. Understanding, Verifying, and Implementing Emergency Use Authorization Molecular Diagnostics for the Detection of SARS-CoV-2 RNA. J Clin Microbiol. 2020;58(8):e00796-20.