Diagnosing COVID-19 infection: the danger of over-reliance on positive test results

The evidence shows that false positive PCR results are common enough to impact clinical and policy decisions.

Andrew N. Cohen, Ph.D.¹*, Bruce Kessel, M.D.², Michael G. Milgroom, Ph.D.³

¹ Center for Research on Aquatic Bioinvasions, Richmond CA, USA. acohen@bioinvasions.com
² John A. Burns School of Medicine, University of Hawai‘i, Honolulu HI, USA. bkessel@queens.org
³ School of Integrative Plant Science, Plant Pathology and Plant-Microbe Biology Section, Cornell University, Ithaca NY, USA. mgm5@cornell.edu

* Corresponding author

Abstract: Unlike previous epidemics, in addressing COVID-19 nearly all international health organizations and national health ministries have treated a single positive result from a PCR-based test as confirmation of infection, even in asymptomatic persons without any history of exposure. This is based on a widespread belief that positive results in these tests are highly reliable. However, data on PCR-based tests for similar viruses show that PCR-based testing produces enough false positive results to make positive results highly unreliable over a broad range of real-world scenarios. This has clinical and case management implications, and affects an array of epidemiological statistics, including the asymptomatic ratio, prevalence, and hospitalization and death rates. Steps should be taken to raise awareness of false positives, reduce their frequency, and mitigate their effects. In the interim, positive results in asymptomatic individuals that haven’t been confirmed by a second test should be considered suspect.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Key messages

The high specificities (usually 100%) reported in PCR-based tests for SARS-CoV-2 infection do not represent the real-world use of these tests, where contamination and human error produce significant rates of false positives.

Widespread misunderstanding of these false positive rates affects an array of clinical, case management and health policy decisions. Similarly, health authorities’ guidance on interpreting test results is often wrong.

Steps should be taken immediately to reduce the frequency and impacts of false positive results.

Tests for current infection with SARS-CoV-2 are based on detection of the virus’ RNA. Most of these tests use the polymerase chain reaction (PCR) to amplify diagnostic sequences within the virus’ genome. According to leading health authorities, while negative results from these tests are frequently wrong, positive results are highly reliable.1-3 Accordingly, the World Health Organization (WHO) and most government health ministries accept a positive PCR result on a single sample as confirmation of SARS-CoV-2 infection, even in asymptomatic persons without any history of exposure.4-10

However, in practice, PCR-based testing produces a significant number of false positive results, making positive results highly unreliable over a broad range of real-world scenarios. Consequently, the frequent assertion that positive test results for SARS-CoV-2 are more reliable than negative results is wrong most of the time, and the widespread and official reliance on a single positive PCR result as a sufficient basis for diagnosis—which was not the rule in previous epidemics—has been a mistake. The general misunderstanding of the rate of false positives in SARS-CoV-2 testing impacts clinical and case management decisions, and through flawed interpretations of test statistics, has affected health policy decisions.

False positives

The accuracy of a diagnostic test is measured by sensitivity, which is the proportion of infected individuals that test positive, and specificity, the proportion of uninfected individuals that test negative. Although SARS-CoV-2 PCR assays are widely reported to have 100% specificity (Supplemental Material-Version 3: Table S1), this refers only to the tests’ lack of reaction with substances other than SARS-CoV-2 genetic material (analytical specificity), and not to the potential for incorrect results in the real-world implementation of testing (clinical specificity) where contamination or human error can generate false positives during sample collection, transport and analysis.

In previous epidemics, health authorities voiced concerns that false positive results from PCR-based tests could harm both the individuals tested and the ability of government agencies to assess the outbreak, and they adopted measures to limit the occurrence of false positives. For example, the World Health Organization and the U.S. Centers for Disease Control and Prevention limited PCR-based testing to individuals with a high probability of infection (those with symptoms and/or significant exposure) and usually required confirmation of positive results by a second, independent test (Box 1). These warnings and requirements are absent from the same organizations’ guidance on SARS-CoV-2 testing.
Box 1: Measures minimizing false positive results in PCR-based tests

Then (US CDC and WHO recommendations):

SARS-CoV-1
"To decrease the possibility of a false-positive result, testing should be limited to patients with a high index of suspicion for having SARS-CoV disease...In addition, any positive specimen should be retested in a reference laboratory to confirm that the specimen is positive. To be confident that a positive PCR specimen indicates that the patient is infected with SARS-CoV, a second specimen should also be confirmed positive."

"[R]equirements for the laboratory diagnosis of SARS...almost always involves two or more different tests or the same assay on two or more occasions during the course of the illness or from different clinical sites...A single test result is insufficient for the definitive diagnosis of SARS-CoV infection."

MERS-CoV
"Confirmation of a "presumptive positive" result by CDC is required." Requirements for testing include both specific clinical features and epidemiologic risk.

Testing should be limited to persons with specified symptoms and, in most cases, elevated risk of exposure.

Ebola Virus
"CDC recommends that Ebola testing be conducted only for persons who...[have] both consistent signs or symptoms and risk factors...Any presumptive positive Ebola test result must be confirmed at the CDC...CDC considers a single diagnostic test...insufficient for public health decision-making."

Case confirmation requires specific clinical signs in addition to a single positive PCR test.

Zika Virus
Testing is recommended only for pregnant women with symptoms and recent exposure, or asymptomatic pregnant women with ongoing exposure. "[B]ecause of the potential for false-positive...results, updated recommendations include [PCR] testing of both serum and urine and concurrent Zika virus IgM antibody testing to confirm the diagnosis...with more than one test."

Testing is recommended only for symptomatic patients.

Now:

SARS-CoV-2
Except for validation of a laboratory's first few results, we found no requirement or recommendation for a second confirmatory test—either on a separate sample or a second aliquot from the initial sample—in guidance documents from the World Health Organization, the US Centers for Disease Control and Prevention, the European Centre for Disease Prevention and Control, Public Health England, the Public Health Agency of Canada, the Pan American Health Organization, or South Korea's Centers for Disease Control and Prevention; instead these entities require only a single positive PCR result to confirm infection in either symptomatic or asymptomatic persons. The Chinese Centers for Disease Control and Prevention requires clinical manifestations and in most cases epidemiological exposure in addition to a single positive PCR result to confirm a case. Since May 27 the Norwegian Institute of Public Health has recommended confirmatory tests of positive results in persons who are both asymptomatic and without a history of exposure.

In most regions testing was initially restricted to persons with clinical signs and symptoms and elevated epidemiological risk, but as tests became more available many authorities called for and implemented broader use of PCR-based tests, including testing of individuals with no symptoms or known exposure risk.
**Testing the tests**

External quality assessments (EQAs) test the implementation of medical diagnostic assays by providing participating laboratories with blind panels of positive and negative samples. The laboratories assay these samples using their normal procedures and report the results to the EQA manager, who compiles and analyzes the results. Since relevant data from EQAs of SARS-CoV-2 assays are not yet available, we reviewed the range of false positive rates (FPRs) in 43 EQAs of PCR assays of RNA viruses conducted in 2004-2019 (see Supplemental Material-Version 3: Methods). Each EQA enrolled between three and 174 participating laboratories, which together provided results for 4,113 blind panels containing 10,538 negative samples, of which 336 (3.2%) were reported as positive (Table 1). We considered two data sets comprising all 43 EQAs (full data set), and the 35 EQAs that analyzed at least 100 negative samples (subset). FPRs in each EQA ranged from 0 to 16.7% for the full data set, and 0 to 8.1% for the subset. The median and the interquartile range were lower for the full data set (median=2.3%, interquartile range=0.8-4.0%) than for the subset (median=2.5%, interquartile range=1.2-4.0%) (Supplemental Material-Version 3: Figure S2).

The EQAs did not report any relationship between false positives and the type of assay used. This is unsurprising since the likely sources of these false positives (contamination, human error) are more directly connected to laboratory practices and layouts than to which particular assay is used. Thus there is probably no systematic difference between the FPRs among the scores of different assays used to detect SARS-CoV-2 and the hundreds of different assays in the reviewed EQAs.

**False positives’ impact on the reliability of positive results**

We conservatively used the lower of the 25th percentile FPR values from the two EQA data sets to model the reliability of results. This FPR value (0.8%) is further conservative in that it doesn’t include false positives produced during sampling, and hasn’t been adjusted for any expected increase in error rate stemming from the rapid expansion of SARS-CoV-2 testing and the use of novel diagnostic assays. We used a false negative rate of 25%, based on published estimates that ranged from 0% to 52.2% (Supplemental Material-Version 3: Table S2). Sensitivity analyses that varied the false negative rate from 0% to 50% had little effect on the reliability of positive results (Supplemental Material-Version 3: Figures S3-S6).

Even a low FPR reduces the reliability of positive results when prevalence is low. This is apparent in model results based on test data for 82 countries and the 50 US states (Figures 1A, 2A), which show that in regions with low test positivity (toward the right side of the panels) much of the testing is targeted too broadly to be useful. In such cases, testing should be limited to persons with a higher probability of infection, or reliability should be improved by requiring the confirmation of positive results with a second, independent test.

Model results based on daily test data reveal the trajectories of test program reliability (Figures 1B-D, 2B-D). Some regions maintained broad testing even as test positivity declined to low levels and the reliability of positive results approached or reached zero. For example, the model results suggest that in South Korea after around April 7th nearly all positive results, especially for asymptomatic individuals, were probably false positives. Nonetheless, South Korea continued to conduct over 6,000 tests a day.

The reliability of positive results falls to near zero when test positivity approaches the FPR. However, even with positivities up to around four times the FPR, over 20% of positive results are likely to be false positives. Unless other respiratory diseases are pervasive in a test...
population, most of these false positive individuals would likely be asymptomatic, which could at least partially explain the reports of large numbers of asymptomatic carriers of SARS-CoV-2. Positive PCR results in individuals that don't develop symptoms should generally be considered doubtful unless confirmed by a second positive test.

**Interpreting individual test results**

Statements from health agencies and public health authorities often suggest that positive results from SARS-CoV-2 tests are more trustworthy than negative results. However, over a wide range of likely scenarios, the opposite is true: for example, in figures 1 and 2 wherever the blue columns showing the positive predictive value are lower than the orange columns showing the negative predictive value, positive results are more likely to be wrong than are negative results. The reason for this is that the FPR acts on samples from the uninfected fraction of the population, producing positive results for a percentage of the uninfected individuals, while the false negative rate acts on the infected fraction, producing negative results for a percentage of the infected individuals. When prevalence is low, the uninfected fraction is much greater than the infected fraction, so even a low FPR can have a larger effect than a high false negative rate.

**Sources of false positives**

Most false positives in PCR-based tests are probably due to contamination, derived from such sources as positive samples analyzed by the laboratory, positive controls, contaminated reagents, or infected workers. Massive amplification of nucleic acids makes PCR-based assays highly sensitive, but also highly vulnerable to minute levels of contamination which can produce false positives that are indistinguishable from true positives. False positives can also be produced by sample mix-ups or data handling errors.

In addition to EQAs, other types of studies have reported false positives in PCR assays for other coronaviruses. There is also laboratory evidence of false positives in SARS-CoV-2 assays: four studies conducting sensitivity or cross-reactivity assessments on SARS-CoV-2 PCR assays reported false positives when negative samples were tested, apparently due to contamination in those laboratories (Supplemental Material-Version 3: Tables S3, S4). Numerous cases of false positives in regular SARS-CoV-2 testing have been reported in the media (Supplemental Material-Version 3: Previous Reporting).

**Impacts**

Much attention has been paid to false negative rates in SARS-CoV-2 PCR testing (Supplemental Material-Version 3: Table S2) and to FPRs in antibody testing, but we found virtually no discussion in the scientific or medical literature of FPRs in SARS-CoV-2 PCR testing (Supplemental Material-Version 3: Previous Reporting). Failing to anticipate and either correct or adjust for false positive results has clinical and case management consequences, including waste of personal protective equipment, waste of human resources in contact tracing, and potentially dangerous sequestering of uninfected individuals with infected individuals. A false positive test result could impede a correct diagnosis, delaying or depriving patients of appropriate treatment. False-positive patients introduce noise into clinical observations, which may hinder the development of improved medical care based on clinical experience. False-positive individuals or their close contacts could be subjected to medically inappropriate therapies, including treatment with prophylactic or antiviral medications and antibody therapy. Individuals that have falsely tested positive may be less likely to avoid future exposure to infected individuals, believing they have immunity, and for the same reason may not seek vaccination when it becomes available. Clinical trials could lose statistical power by unwittingly...
enrolling false-positive individuals, who would be exposed to potentially harmful side effects without any mitigating potential for benefit. False positives also distort the estimates of an array of epidemiological statistics that affect policy decisions, including the asymptomatic ratio, prevalence, and hospitalization and death rates.

Fixing the problem

The impact of false positives in SARS-CoV-2 testing could be mitigated by increasing the awareness of false positives; by improving estimates of false positive rates with appropriately-designed EQAs, or by assessing results retrospectively with serological tests; and by reducing the frequency of false positives by requiring two independent positive tests to diagnose an individual as infected.

Like all tests, PCR-based assays are subject to error that includes both false negative and false positive results. A successful testing program must understand the error rates and use tests appropriately. While SARS-CoV-2 testing to date has clearly missed the mark, we can course-correct: we can reassess plans for group-testing or mass-testing using realistic estimates of the potential range of false-positive rates; reconsider the conclusions of studies that implicitly assumed a zero false positive rate; and reduce misdiagnoses and statistical miscasts by confirming positive results with a second independent test, especially in asymptomatic persons and in areas where test positivity is low. In the interim, where mass-testing or group-testing has been conducted without regard to symptoms or exposure—notably in certain regions, congregate-living facilities, workplaces and sports leagues—positive results in healthy individuals that haven't been confirmed by a second test should be considered suspect.
References


Contributors and Sources: ANC is a marine biologist with a focus on biological invasions; this analysis grew out of previous work on false positives in PCR-based assays used in environmental monitoring. BK is a clinical Ob/Gyn academic whose research has ranged from menopause to ovarian cancer. MGM has used PCR-based analyses to investigate plant diseases. ANC designed and conducted the analyses with input from MGM; ANC and BK collaborated on the review of clinical implications; and all authors participated in writing and editing the report. ANC is the guarantor. We thank Elisa Liberti, Dominic Chow, Thomas Taylor, James Carlton, Richard Norgaard and Bruno Pernet for helpful reviews and other assistance. This paper is dedicated to the memory of Kirk Smith, who kindly reviewed the manuscript and offered insightful advice a few weeks before his untimely death. All data is available in the manuscript or in the Supplemental Material-Version 3. No funding was received for this work. The authors declare no conflicts of interest.
Fig 1 Reliability of SARS-CoV-2 test results in different countries. Positive predictive value (the probability that a positive result is true) and negative predictive value (the probability that a negative result is true) calculated with a false negative rate of 25% and a false positive rate of 0.8%. (A) Results for 82 countries based on cumulative test data through the most recent available date (between May 5 and May 24, 2020). Countries arranged left to right in order of decreasing test positivity. (B-D) Reliability trajectories based on the previous-7-day moving average, showing countries where the reliability of positive test results has declined significantly (United Kingdom), sharply (Italy), and precipitously (South Korea). Cumulative test data are from Our World in Data (https://github.com/owid/covid-19-data/tree/master/public/data/ accessed May 24, 2020). Daily test data are from Our World in Data, the Italian Ministry of Health (http://www.salute.gov.it/portale/nuovocoronavirus/archivioNotizieNuovoCoronavirus.jsp?lingua=italiano&menu=notizie&p =dalministero&area=nuovocoronavirus&notizie.page=0 accessed May 24, 2020) and the South Korean Centers for Disease Control and Prevention (https://www.cdc.go.kr/board/board.es?mid=&bid=0030 accessed May 24, 2020).
Fig 2 Reliability of SARS-CoV-2 test results in the United States. Positive predictive value (the probability that a positive result is true) and negative predictive value (the probability that a negative result is true) calculated with a false negative rate of 25% and a false positive rate of 0.8%. (A) Results for the 50 U.S. states based on cumulative test data through May 24, 2020. States arranged left to right in order of decreasing test positivity. (B-D) Reliability trajectories based on the previous-7-day moving average, showing states where the reliability of positive test results has declined significantly (New York), sharply (Oregon), and precipitously (Hawai‘i). Test data are from The COVID Tracking Project (https://covidtracking.com/about-data accessed May 24, 2020).
Table 1  False positive rates in external quality assessments of RNA virus assays

<table>
<thead>
<tr>
<th>Virus</th>
<th>Number of EQAs</th>
<th>Dates of EQAs</th>
<th>Laboratories per EQA</th>
<th>Negative samples per EQA</th>
<th>False positive rates$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARS-CoV</td>
<td>1</td>
<td>2004?</td>
<td>58</td>
<td>174</td>
<td>2.3-6.9%</td>
</tr>
<tr>
<td>MERS-CoV</td>
<td>3</td>
<td>2014-17</td>
<td>49-99</td>
<td>49-1134</td>
<td>&lt;0.6-1.0%</td>
</tr>
<tr>
<td>Influenza A viruses</td>
<td>17</td>
<td>2007-2019</td>
<td>64-174</td>
<td>114-332</td>
<td>&lt;0.6-7.0%</td>
</tr>
<tr>
<td>Hepatitis C virus</td>
<td>8</td>
<td>2005-07</td>
<td>5-104</td>
<td>21-728</td>
<td>2.1-7.0%</td>
</tr>
<tr>
<td>Hepatitis Delta virus</td>
<td>1</td>
<td>2015?</td>
<td>28</td>
<td>112</td>
<td>5.4%</td>
</tr>
<tr>
<td>Chikungunya virus</td>
<td>2</td>
<td>2007, 2014</td>
<td>31-56</td>
<td>108-297</td>
<td>1.9-8.1%</td>
</tr>
<tr>
<td>Chikungunya, Dengue</td>
<td>1</td>
<td>2015</td>
<td>20</td>
<td>40</td>
<td>2.5%</td>
</tr>
<tr>
<td>Dengue virus</td>
<td>1</td>
<td>2013</td>
<td>16</td>
<td>16</td>
<td>6.3%</td>
</tr>
<tr>
<td>Zika virus</td>
<td>1</td>
<td>2016</td>
<td>50</td>
<td>504</td>
<td>2.8%</td>
</tr>
<tr>
<td>Rift Valley Fever virus</td>
<td>1</td>
<td>2012</td>
<td>30</td>
<td>117</td>
<td>3.4%</td>
</tr>
<tr>
<td>Measles virus</td>
<td>1</td>
<td>2014</td>
<td>41</td>
<td>123</td>
<td>0.8%</td>
</tr>
<tr>
<td>Ebola virus</td>
<td>5</td>
<td>2014-16</td>
<td>3-82</td>
<td>3-317</td>
<td>0.3-16.7%</td>
</tr>
<tr>
<td>4 arboviruses</td>
<td>1</td>
<td>2017</td>
<td>51</td>
<td>204</td>
<td>4.9%</td>
</tr>
</tbody>
</table>

$^a$ "<" indicates a false positive rate below the detection limit; treated as zero in the analyses.

$^b$ See Supplemental Materials-Version 3: Table S5 for full data and references.