Analyzing Conflicting Results in Rapid Point-of-Care COVID-19 Testing

Title:

Analyzing Conflicting Results in Rapid Point-of-Care COVID-19 Testing

Running Title:

Rapid Point-of-Care COVID-19 Testing

Authors:

Michael Mina MD, PhD, Center for Communicable Disease Dynamics, Department of Epidemiology and Department of Immunology and Infectious Diseases, Harvard T. H. School of Public Health. Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA

Steve Miller MD, PhD, Director, UCSF Clinical Microbiology Laboratory, San Francisco, California, USA

Michael Quigley MD, PhD – Medical Director, Scripps Health Sorrento Mesa Core Laboratory, San Diego, California, USA

Tyler Prentiss, BA, Global Health Initiative, Henry Ford Health System, Detroit, Michigan, USA

John Zervos, PD, Global Health Initiative, Henry Ford Health System, Detroit, Michigan, USA

John E. McKinnon, MD, MSc, Infectious Diseases & Translational and Clinical Research Center, Henry Ford Health System, Detroit, Michigan, USA

Stewart Comer MD, MA, FCAP, – Laboratory Director, Santa Barbara Cottage Hospital, Santa Barbara, California, USA

Corresponding Author:

Michael Mina, Harvard T.H. Chan School of Public Health, Department of Epidemiology, Boston, MA. 518 698 2756. mmina@hsph.harvard.edu

Disclosures:

MJM has received consulting fees from Roche Diagnostics, Sanofi Pasteur, and Abbott. SM has received consulting fees from Roche Diagnostics and Abbott. MQ has received consulting fees from Abbott. TP, JZ and JEM received grant support from Abbott for the ID NOW validation study. SC reports no financial disclosures.

Acknowledgements:

Abbott provided funding for the development of this manuscript. The authors wish to acknowledge Christopher G. Parkin for his editorial assistance in developing the manuscript.

Abstract

Rapid point-of-care (POC) testing for the SARS-CoV-2 virus provides an ability to quickly test individuals, facilitating near real-time action if infection is determined, such as isolation and initiation of contact tracing. Rapid POC testing also facilitates more frequent testing in individuals, which is crucial for screening programs and for environments where as up-to-date results are essential, for example to facilitate safe entrance into nursing homes. The Abbott ID NOW is a rapid POC test that delivers an accurate result in 5-13 minutes and that has demonstrated a high level of performance in variety of patient populations. However, recent anomalous results by Basu et al. have raised important questions about the sensitivity of the Abbott test. Understanding the validity of these anomalous results cannot be overstated. The ID NOW is currently used widely as a mainstay of many COVID-19 testing programs: urgent-care clinics, hospitals, businesses and other institutions rely on the test to provide fast and accurate results every day. Here, we assess the anomalous findings and find significant limitations in the study design, comparisons used, and patients evaluated. We describe how these limitations can account for essentially all of the losses in sensitivity described by Basu et al. We also discuss interim results from an ongoing multi-site clinical study in urgent-care clinics that indicate that the ID NOW COVID-19 rapid test is highly sensitive (≥94.7%) and specific (≥98.6%).

Rapid POC testing in outpatient clinical settings and acute care facilities offers the ability to quickly diagnose and isolate infected patients and enhance safeguards against virus transmission. It is important that clinicians received accurate information about rates of sensitivity and specificity in order to feel confident using current POC rapid-testing systems.

Introduction

A key concern among healthcare providers and public health officials is the risk of COVID-19 virus spread by infected individuals including by individuals who remain asymptomatic while the infection resolves without medical assistance. Given that the SARS-CoV-2 virus can transmit prior to symptom onset and during the early course of the illness progression,(1-3) there is an urgent need for testing technologies that enable clinicians to quickly diagnose the disease and isolate patients for appropriate care. Because many people are now reluctant to be tested at local hospitals and are more likely to go to an urgent care or other out-patient testing facility, point-of-care (POC) testing provides an opportunity to quickly diagnose individuals who would not otherwise get tested.

Since April 2020, the U.S. Food and Drug Administration (FDA) has authorized more than 30 COVID-19 tests through its Emergency Use Authorization (EUA) pathway. A number of these new tests are intended for POC use, but only a few give high sensitivity molecular results in a timely enough fashion to directly aid medical decisions. To ensure accuracy of results, POC systems are required to undergo validation studies that use laboratory-based testing systems for comparison and confirmation of results.

As part of the process for bringing new tests into healthcare organizations, institutions validate performance by comparison with other testing methods. These validations are not typically true clinical studies, per se, but instead provide a way for clinical laboratories and, in the case of POC tests, other institutions to define acceptance of new testing modalities. Recent results from a validation study, authored by Basu et al., have called into question the performance of the Abbott ID- NOW test platform. (4) The pre-print upload led the FDA to

inform the public about possible accuracy concerns of the ID-NOW assay. Given the wide use of the ID-NOW system for critical POC detection of SARS-CoV-2 infections, there is an immediate need to evaluate the reported findings of that validation study.

Here, we assess the validation approach and findings in that study. We identify numerous aspects of the study design and specimens used that are likely to have driven the anomalously low sensitivity measured for the ID-NOW assay. We also describe, in comparison, new data from interim findings in ongoing Abbott ID NOW™ studies and how the Basu et al. study is anomalous to other studies.

Although the ID NOW test has consistently demonstrated a high level of performance against a gold standard (nasopharyngeal (NP) swab, ranging from 83.3% - 95.0% positive agreement (sensitivity) and 96.5% - 100.0% negative agreement (specificity), in a variety of patient populations, results from a recent study by Basu et al. have raised important questions about the accuracy of the Abbott test.(4) In the study, investigators compared results from the ID NOW and Cepheid GeneXpert® Xpress POC tests, using the Cepheid GeneXpert® Dx laboratory instrument as their comparator reference method. Two key comparisons were made. First a comparison of samples collected using nasopharyngeal (NP) swabs in VTM were compared on both instruments. Second, a comparison of samples collected using anterior nare (AN) nasal swabs (introduced as 'dry' swabs directly into the ID-NOW assay) were tested on the ID-NOW instrument and were compared to NP swabs transported in VTM and tested on the GeneXpert® Xpress test. Investigators reported that the ID NOW test missed a third of the samples (5 of 15) detected positive by Xpert Xpress when using NP swabs in VTM and over 48% when using dry nasal swabs.

Here we describe three problems that can explain essentially all of the loss in sensitivity measured by Basu et al to explain this important anomalous finding. 1) In a number of key comparisons, specimens were diluted into VTM rather than direct inoculation of the swab into the ID-NOW test reagent, suggested by the manufacturer. This likely reduced the inoculated virus by up to an order of magnitude compared to manufacturers suggested protocol of direct inoculation of the swab into the assay. 2) In all populations tested, the specimens used in the study represent an extra-ordinarily skewed distribution of samples with very low RNA concentrations. All (100%) specimens missed by the ID-NOW in the study represent such low RNA concentrations that they likely reflect non-culturable or non-viable virus RNA remaining after infectious virions have been cleared. 3) When direct inoculation of the swab into the ID NOW assay reagent was performed, a useful assessment of the assay was obscured by the choice to compare anterior nasal swabs on the ID-NOW to nasopharyngeal swabs on the Xpert Xpress. The discrepancy between anatomic sites sampled (AN versus NP) causes any assessment to measure differences in the anatomic niche of the virus rather than provide a useful comparison between assays and ensuring that the technique used for collection is appropriate.

In sum, we find multiple methodological problems with the assessments put forth in the Basu et al. validation study. Collectively, these methodological flaws can explain all of the loss in sensitivity determined by Basu et al. and help to explain how this particular study provided such anomalous results. The importance of in-depth scrutiny of this particular study by Basu et al. cannot be overstated. The consequences of confusion over the sensitivity of the ID NOW,

including the FDA's letter informing the public of possible sensitivity issues of the assay, have undoubtedly caused enormous concern among users of ID NOW instruments.

Pre-dilution into VTM

In three of four assessments made by Basu et al. specimens were diluted in VTM. The ID-NOW instrument is intended for direct inoculation of the swab into the assay reagent, thereby increasing sensitivity by removing the initial dilution step. Other assays like the Xpert Xpress and the CDC assay are unable to accept a direct inoculation with the swab. Assuming that 0.5 mL of VTM was added to the ID-NOW in the analysis, the dilution process likely reduced the amount of tested virus that entered into the ID-NOW assay by an order of magnitude compared to the manufacturers recommended direct inoculation. Thus, measurement of assay sensitivity using specimens that are pre-diluted in VTM do not reflect the true sensitivity of the ID-NOW assay.

Skewed specimen distribution toward very low viral RNA loads.

In the first part of their validation study, Basu et al. obtained 25 NP specimens diluted in VTM that were positive on the Xpert Xpress assay and used these to measure assay sensitivity. In their primary analysis of 15 specimens they found that the ID-NOW failed to detect RNA in 5 (33%) of the 15 specimens. Importantly here, four of the missed specimens had Ct values >40 (44.3, 44.1, 43.1, 40.7), representing such low RNA concentrations that they would have been reported as negative by all other gold-standard quantitative PCR assays, including the CDC assay (5; 6) nearly all of which use a cutoff Ct value for positivity of less than 40. In addition, in

the three specimens with Ct values >43, the Xpert Xpress detected RNA with only one of the two primers (N2 positive, E negative), calling into question whether these may have been false positive results, which is not uncommon when attempting to interpret quantitative PCR beyond 40 cycles of amplification. Removal of only these four exceedingly low viral RNA concentration samples from analysis, in order to align with all other gold-standard qPCR assays, alone changes the sensitivity results dramatically, from a measured sensitivity of 66% in their primarily analysis to 93% - in line with other assessments of the ID-NOW. In addition, the fifth of five specimens that was missed in that analysis had a Ct value of 36 on the Xpert Xpress assay. Though not above 40, this too represents a *very* low viral RNA concentration, so low that the RNA detected likely does not reflect culturable virus and does not reflect the transmitting stage of the virus. (7) As well, and as discussed above, the pre-dilution of the specimen into VTM likely reduced the RNA molecules introduced into the assay by an order of magnitude compared to direct swab inoculation. While we cannot know for certain, the ID-NOW would likely have detected the specimen if it was a direct inoculation per manufacturer suggested protocol.

In addition to the 15 specimens just discussed, an additional ten NP specimens, also diluted in VTM, were evaluated by Basu et al. In this assessment, the ID-NOW correctly detected RNA in all specimens that had Ct values less than 40. In one sample with a Ct value of 38 that was evaluated three times, the ID-NOW identified the specimen as positive in 2 of 3 attempts. In addition, the ID-NOW detected a positive result in one sample with a Ct value of 40.2. These results represent that the ID-NOW assay is an assay with potential for very high molecular sensitivity, despite pre-dilution of all specimens in VTM.

Overall, if specimens with Ct values >40 were removed from the Basu analysis – to align with a comparison against essentially any other gold-standard qPCR assay – then of the 22 qualifying (i.e. Ct <40) measurements made, only two would have been missed (both also with very high Ct values: 38 and 36) resulting in a measured sensitivity against the comparison assay qPCR of 90.9%. Among samples with likely culturable virus (i.e. Ct <34) the sensitivity was 100%. This is a very large departure from the 66% reported by Basu et al. for the portion of their analysis that relies on NP swabs diluted in VTM. Had these specimens all been directly inoculated into the ID-NOW instrument, it is likely that the measured sensitivity in these 22 specimens would have approached 100%.

Comparison of AN swabs to NP swabs is a referendum on the anatomic niche, not the assay.

In the results discussed above, the comparisons made compared apples to apples (i.e. NP swab to NP swab) but as discussed, the true assay sensitivity was obscured by an enrichment of the sample pool with exceptionally low viral RNA load samples, causing any minor reductions in sensitivity, if they exist, to be inappropriately amplified. In addition, the NP swabs were diluted in VTM against the ID-NOW manufacturer's suggested protocol. In the final major assessment of the ID-NOW by Basu et al. the authors did not dilute the swabs in VTM. However, potentially worse, the authors compared specimens collected using AN nasal swabs on the Abbott ID-NOW to NP swabs on the Xpert Xpress. NP swabs are the gold-standard sample collection method for COVID-19 testing and offer the best efficiency in capturing respiratory viruses, including SARS-CoV-2 onto a swab for testing. (8) However, NP swabs are cumbersome to perform, requiring trained medical professionals and extensive resources, both

in person-time and personal protective equipment (PPE). On the other hand, an AN nasal swab, while having slightly decreased efficiency to capture viral particles that usually reside in the greatest concentration in the nasopharynx, can be more easily administered, including self-administered, with the major benefits of improving throughput of testing, reducing time, and conserving PPE. Because of these benefits, which are particularly important for the type of POC testing that the ID-NOW is intended, the ID-NOW SARS-CoV-2 testing kits are conveniently packaged with AN nasal swabs. However, the ID-NOW and the Xpert Xpress assays are each approved for NP, AN as well as oropharyngeal swabs. Therefore, because of the expected biological loss in picking up virus when swabbing the anterior nare versus the nasopharynx, to determine the relative sensitivity of the ID-NOW molecular assay, the same swab type should be used on the referent assay. Instead, in their primary comparison, Basu et al. compared specimens collected using NP swabs on the Xpert Xpress to those collected using AN swabs on the ID-NOW. Careful evaluation of their results shows that their primary findings are more a referendum on the anatomic niche of the virus, the swab type and the skew of the data towards exceptionally low viral load specimens, rather than an informative comparison of the molecular assays.

In that assessment, which measured a 51.6% sensitivity of the ID-NOW to detect SARS-CoV-2, the specimens were again heavily skewed towards very high Ct value (low viral RNA) specimens that do not reflect the transmitting stage of infection. Of 31 positives, 25 of 31 (81%) were at NP Ct values above 33.7 and, surprisingly, 14 of 31 (45%) were at NP Ct values >38, which is potentially a false positive and is essentially at or above the limit of detection for nearly all gold-standard qPCR assays. Thus, potentially only 20% or fewer of the tested

specimens in this sample set may have included specimens from individuals with transmissible virus.

As in the first series of analyses addressed above, this heavy skew towards high Ct value specimens runs the serious risk of greatly exaggerating subtle reductions in assay sensitivity near the limit of detection with the effect, here, of massively overestimating loss in sensitivity. Moreover, the specimens for the ID-NOW in this comparison were collected as AN swabs. Whether an AN swab is biologically capable of even retrieving viral RNA from an individual with an NP Ct value of 38 or above is questionable. For the 14 of 31 specimens with paired NP swab Ct values >38 and at least a portion of the 11 of 31 specimens with paired NP swab Ct values between 33.7 and 38, the inability to detect RNA from an AN swab is likely driven by the anatomical sampling location, without offering additional information on the ID-NOW assay itself. Importantly, of the 19% of specimens with NP Ct values below 33.5, indicating potentially culturable virus, the ID-NOW had 100% concordance. Combined, the heavy skew towards high Ct specimens measured in NP swabs, plus the use of AN swabs can likely explain essentially all of the loss in sensitivity, without invoking any significant loss in sensitivity of the molecular ID-NOW assay itself.

The outcome of our assessment of the Basu et al. validation places the results squarely in line with other evaluations of the ID NOW assay, which have demonstrated high levels of performance against the gold standard (nasopharyngeal (NP) swab, ranging from 83.3% - 95.0% positive agreement (sensitivity) and 96.5% - 100.0% negative agreement (specificity).

Current Findings from ID NOW Test Performance Evaluations

Urgent Care Clinic Study

Analysis of interim results from an ongoing multi-site clinical study in five urgent care clinics (in New Jersey, Tennessee, Louisiana, Texas and South Carolina) that has enrolled 256 subjects to date indicate that the ID NOW COVID-19 rapid test is showing strong agreement to laboratory-based PCR instruments. In that study, the ID NOW has thus far identified 29 of 29 (100% positive agreement) individuals when compared to a commonly-used laboratory-based molecular PCR assay (the Roche cobas® SARS-CoV-2 assay). Compared to the Roche assay, the ID NOW has also identified 226 of 227 individuals as negative – with the true status of the one individual identified as positive on the ID NOW but negative on the Roche currently under investigation. In parallel, ID NOW has demonstrated 94.7% positive agreement and 98.6% negative agreement compared to the Centers for Disease Control (CDC) 2019-Novel Coronavirus (COVID-19) Real-Time RT-PCR Diagnostic Panel. In comparison, Roche has demonstrated 95.0% positive agreement and 98.7% negative agreement when compared to the CDC assay. This multi-site study is one of the first studies conducted on the ID NOW COVID-19 test in a real-world setting as it is intended to be used.

The Everett Clinic Study

The Everett Clinic Study enrolled symptomatic patients presenting within 7 days of symptom onset. Results from this 955-subject study (763 symptomatic; 192 asymptomatic) conducted at the Everett Clinic have to date shown 91.3% positive agreement (sensitivity) and 100% negative agreement (specificity greater than 99.5% at the lower confidence limit) for ID NOW compared to lab-based PCR assays.

Detroit Method Verification Analysis

In a verification study performed by the Detroit Department of Health (DOH) and Michigan Department of Health and Human Services (MDHHS), the ID NOW showed 98% agreement with RT-PCR results from the 2019-nCoV Real-Time RT-PCR Diagnostic Panel assay.(9) Results from these demonstrated highest sensitivity in individuals tested during the early stages of symptom onset when viral loads were highest; whereas, even laboratory-based testing system can have false negatives toward the end of the infection cycle when the individual may no longer be infectious.(10)

Summary

Until proven therapies and effective vaccines become available, frequent use of POC rapid-testing technologies is emerging as one of the promising measures to prevent SARS-CoV-2 outbreaks. A significant advantage of POC instruments is the ability to test individuals more frequently and with greater safety.

For outpatient clinical settings and acute care facilities, POC testing that provides results in minutes rather than hours or days offers significant benefit through the ability to quickly diagnose and isolate infected patients and enhance safeguards against transmission. Moreover, use of these technologies will become increasingly important as communities gradually relax stay-at-home restrictions.

As new testing platforms and solutions become available, it will be important to confirm that these tests, when deployed against their intended use and setting, perform well and within

expectations. Validation studies, if performed well, support these uses. However, as discussed at length above, a validation done without proper comparisons and without appropriately representative populations can cause confusion to the medical community and can result in drastic but essential measures by regulatory bodies until any confusion is resolved.

Researchers should be cognizant that their findings – even when presented as preliminary – stand to be misinterpreted by non-scientific audiences to the potential detriment of public health. This is especially true in an era of citizen journalism and media who suddenly find themselves covering science, medicine, and public health without appropriate subject-matter expertise. Clinicians should feel confident in using current POC rapid-testing systems as intended given their known rates of sensitivity and specificity, capacity to reduce the spread of infection in society by identifying and isolating virus carriers, and overall positive impact to improving public health through early detection and treatment.

References

1. He X, Lau EHY, Wu P, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nature Medicine 2020;26:672-675.
2. Zou L, Ruan F, Huang M, et al. SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. N Engl J Med 2020; 382:1177.
3. To KK, Tsang OT, Leung WS, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. Lancet Infect Dis 2020;20(5):565-574.
4. Basu A, Zinger T, Inglima K, et al. Performance of the rapid Nucleic Acid Amplification by Abbott ID NOW COVID-19 in nasopharyngeal swabs transported in viral media and dry nasal swabs, in a New York City academic institution. J Clin Microbiol 2020 May 29;JCM.01136-20. doi: 10.1128/JCM.01136-20. .
5. Centers for Disease Control and Prevention (CDC). CDC 2019-Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel: Catalog # 2019-nCoVEUA-01 1000 reactions. https://www.fda.gov/media/134922/download. Accessed June 2, 2020.
6. Wang W, Xu Y, Gao R, et al. Detection of SARS-CoV-2 in Different Types of Clinical Specimens. JAMA. 2020;323(18):1843-1844. https://jamanetwork.com/journals/jama/fullarticle/276299. Accessed June 2, 2020.
7. Bullard J, Dust K, Funk D, et al. Predicting infectious SARS-CoV-2 from diagnostic samples. Clin Infect Dis 2020 May 22;ciaa638. doi: 10.1093/cid/ciaa638. https://academic.oup.com/cid/advance-article/doi/10.1093/cid/ciaa638/5842165. Accessed June 2, 2020.
8. Péré H, Podglajen I, Wack M, et al. Nasal Swab Sampling for SARS-CoV-2: a Convenient Alternative in Times of Nasopharyngeal Swab Shortage. JCM 2020;58(6);e00721-20; DOI: 10.1128/JCM.00721-20. https://jcm.asm.org/content/58/6/e00721-20. Accessed June 2, 2020.
9. Detroit Health Department (DHD), Michgan Department of Health and Human Services. Method Verification Summary Report. May 1, 2020. https://detroitmi.gov/sites/detroitmi.localhost/files/2020-05/abbot_verification_summary_report.pdf. Accessed May 24, 2020.
10. Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction–Based SARS-CoV-2 Tests by Time Since ExposureAnn Intern Med. 2020 May 13:M20-1495. doi: 10.7326/M20-1495. Accessed May 26, 2020.