AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Clinical Microbiology®

EDITORIAL

Can We Test Our Way Out of the COVID-19 Pandemic?

Matthew A. Pettengill,[a] **Alexander J. McAdam,**[b] Editor in Chief, *Journal of Clinical Microbiology*

[a]Department of Pathology, Anatomy, and Cell Biology, Thomas Jefferson University, Philadelphia, Pennsylvania, USA
[b]Department of Laboratory Medicine, Boston Children's Hospital, Boston, Massachusetts, USA

ABSTRACT Frequent, low-cost, universal testing for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) with quarantine of those with a positive result has been suggested as a strategy to address the coronavirus disease 2019 (COVID-19) pandemic in the United States. Specifically, home or community use of tests that use paper strip detection devices, which may have reduced sensitivity for SARS-CoV-2, has been advocated. There are several potential challenges or problems with this strategy, including the limited availability of such tests, consequences of incorrect test results, difficulties with adherence to testing, and the questionable accuracy of such tests for detection of infectious people. Because of these, we think it is premature to strongly advocate for such a testing strategy, as the adverse consequences may outweigh any benefits. High-quality outcome data demonstrating the efficacy of this testing strategy are needed before widespread implementation.

KEYWORDS COVID-19, SARS-CoV-2

Recently, the use of universal, frequent testing for coronavirus disease 2019 (COVID-19) to guide quarantine has been proposed and widely discussed as a method to dramatically reduce or eliminate community spread of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (1).The core of the proposal is the mass production and utilization of inexpensive, rapid paper strip-based tests that can be used frequently (e.g., daily) in the community, whether in the home, workplace, or school. The proposal is that if such results were available daily to everyone (or to most people) in the country, transmission could be reduced enough to suppress the pandemic. We have several concerns about this approach, both technical and practical.

The current tests for acute SARS-CoV-2 infection in the United States include mainly commercial or laboratory-developed nucleic acid amplification tests (NAATs, such as PCR) with excellent technical sensitivity but somewhat lower clinical sensitivity (2) and a smaller number of antigen-based tests with lower technical and clinical sensitivity. Since early July 2020, more than 600,000 SARS-CoV-2 tests have been performed per day in the United States, and currently the United States has the 2nd highest per-capita rate of testing in the world among the countries for which data are available (https://coronavirus.jhu.edu/testing/international-comparison, accessed 20 August 2020)—and the highest number of positive results and deaths. Many countries have dramatically reduced community spread of COVID-19 without the scale of per-capita testing currently available in the United States. For example, Canada has less than half of the U.S. per-capita testing capacity and dramatically lower per-capita rates of transmission (see the Johns Hopkins database cited above). Perhaps enhanced SARS-CoV-2 testing capacity in the absence of a robust well-funded public health infrastructure and contact tracing capacity is similar to rapid molecular diagnostics for patients with sepsis that are not paired with antimicrobial stewardship intervention—compelling on paper but clinically ineffective (3). Unfortunately, results for NAATs for SARS-CoV-2 in the United States currently have a long turnaround time (TAT) in many laboratories, with some taking several days. Such delayed results are not useful for guiding decisions on

Citation Pettengill MA, McAdam AJ. 2020. Can we test our way out of the COVID-19 pandemic? J Clin Microbiol 58:e02225-20. https://doi.org/10.1128/JCM.02225-20.

Editor Melissa B. Miller, UNC School of Medicine

Address correspondence to Matthew A. Pettengill, Matthew.Pettengill@jefferson.edu.

The views expressed in this article do not necessarily reflect the views of the journal or of ASM.

Accepted manuscript posted online 25 August 2020
Published 21 October 2020

quarantine to reduce transmission and thus decrease the public health impact of the testing capacity in the United States. Testing has been faster at hospital and public health labs, whereas testing at many commercial reference laboratories has been slower as they have been overwhelmed by the testing volume from clinical facilities or testing sites without internal testing capacity. Large-scale testing alone does not appear to have been the key to transmission control in some other countries.

Will rapid, inexpensive strip-based tests be available in the United States? Paper strip-based tests which are simple, inexpensive, and appropriate for community use to detect SARS-CoV-2 are not yet available. There are currently three antigen tests with emergency use authorization in the United States (https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices/vitro-diagnostics-euas#individual-antigen). None is approved for home use, although all can be used in patient care settings with a Clinical Laboratory Improvement Amendment (CLIA) Certificate of Waiver. What about the tests that are in development? Tests in development at Sherlock Biosciences and Mammoth Biosciences that have been promoted as "paper strip and other simple, daily COVID-19 tests" (1) use preamplification by reverse transcriptase and loop-mediated isothermal amplification (LAMP, a NAAT) before detection by a lateral flow assay (4, 5). Each of these two tests takes about an hour to run and requires specific temperatures for reverse transcription and amplification and so would require a testing device or platform. If these tests could be modified for use in the community, we cannot see how they would be available for $1 to $5, less than the cost of most home pregnancy or ovulation tests. Low-cost (or no-cost) testing would be particularly important, as means of access to testing have been unequal, with less testing availability in low-income and minority neighborhoods (https://www.sciencemag.org/news/2020/07/huge-hole-covid-19-testing-data-makes-it-harder-study-racial-disparities, accessed 24 August 2020). Such a low cost for end users could only be realized with tremendous financial subsidies, for example, from the federal or state governments. Paper strip-based tests targeting viral antigens have been developed as well, but details on their performance, cost, and availability are lacking. It is critically important to understand the real-world performance parameters of such tests before being able to evaluate any potential role in diagnostics or screening to reduce transmission. It is also an open issue whether such tests can be developed and produced at the massive needed scale any more quickly than vaccines, which can be expected to dramatically reduce the spread of SARS-CoV-2.

What would be the impact of false-positive results? Studies in which models of test parameters have been used to evaluate effects on transmission of SARS-CoV-2 have paid little or no attention to the specificity of tests or the practical impacts of imperfect specificity. According to these models, both turnaround time and test frequency are more important than test sensitivity for preventing transmission (6, 7). These models rely on theoretical test performance parameters and assume ideal test utilization and human behavior. In one study, the impacts of limit of detection (LOD, directly related to clinical sensitivity), test frequency, and turnaround time were compared for their relative impacts on predicted transmission, with a viral load of 10^3 copies of RNA/ml used to approximate the molecular test LOD and 10^5 copies of RNA/ml used for low-sensitivity daily tests (LSDT) LOD (7). LOD was the least impactful of these parameters, and turnaround time and testing frequency were suggested to be the keys to surveillance leading to reduced transmission for COVID-19. Test specificity was explicitly excluded from consideration in that particular study, as it could not have had an impact given the study design; the automatons in the model accepted their fate as potential transmitters of disease and altruistically isolated themselves for the good of society. We agree that specificity is of reduced importance when the result for false positives is that people comply with quarantine.

In reality, if specificity is even modestly compromised, it will strike at the core of an important parameter for real-world impact of testing: the reliability of the results. Some of the LSDTs under development are antigen-based tests. The few available studies have found SARS-CoV-2 antigen tests to have high specificities but have generally

included relatively small numbers of samples, making it difficult to detect small flaws in the specificity of these tests (8). The real-world specificity of antigen tests may be significantly lower than that suggested by studies performed by trained laboratory staff, as demonstrated by a large cluster of false-positive antigen tests in Vermont (https://www.burlingtonfreepress.com/story/news/2020/07/23/covid-19-testing-how-versions-differ-speed-use-and-accuracy-coronavirus/5446176002/). We are unable to find specificity data for the assays that use LAMP and lateral-flow detection.

Antigen test performance has been evaluated in detail for another respiratory virus—influenza virus. A meta-analysis of 159 studies, including 26 different antigen-based influenza tests, found pooled sensitivity of 62.3% (which is why many labs have abandoned this type of test for influenza) and pooled specificity of 98.2% (9). A recent publication on a newer influenza virus antigen test with improved sensitivity found specificity in a range similar to that found in the meta-analysis cited above, 98.4% (10). Specificity in the range of 98% is typically acceptable (even very good) for use in patients with high pretest probability of having the infection, but when used on a massive scale among asymptomatic subjects there are consequences for even modest reductions in specificity. For example, if LSDT specificity were 98% and tests were used by each of the ~325 million people in the United States every day, there would be a staggering 6.5 million false-positive results each day. In low-prevalence areas of the country, a very large proportion of positive tests would represent false positives. If we consider following up all positives with a NAAT, in the example above the number of confirmatory tests required per day would be an order of magnitude higher than current national capacity (approximately 650,000 tests per day) (https://coronavirus.jhu.edu/testing/international-comparison, accessed 20 August 2020). How could we expect the general public to take the test results seriously and respond to them appropriately if the test is insensitive and many positive results are wrong?

Will people agree to frequent testing? Another important consideration is whether a significant proportion of the population would be willing to use LSDTs or to require their employees, clients, or students to use them. A meta-analysis of more than 100 studies found that physical distancing and wearing masks were associated with lower viral transmission (11). Both interventions are endorsed by the CDC and WHO, and yet their utilization has been politicized, and there is marked resistance to government mandates to wear masks in public spaces in many parts of the country. In locations where mask requirements exist, enforcement and compliance are variable. Similarly, regulations on large gatherings are inconsistently applied or nonexistent. It is difficult to envision LSDTs gaining significant penetration into the areas where transmission interventions are needed most, even if the results were reliable. If the sensitivity were low, as is expected for antigen tests, and it were known that many positives are false positive, the results would not be generally believable and LSDTs would be unlikely to gain widespread use. For example, if a test were 80% sensitive and 98% specific, the positive predictive value (PPV) would be 67.7% and negative predictive value would be 98.9% at a prevalence of 5%. If, however, the test described above were used at a massive scale in areas where prevalence is lower than 5%, which has likely been the case in most of the United States during the majority of time so far during the pandemic, the PPV would be reduced considerably: at 1% prevalence, the PPV would be 28.8%, and at 0.1% prevalence, the PPV would be 3.8%. It is also possible, and perhaps likely, that people would use negative LSDT results (possibly falsely negative) to justify abandoning proven interventions, such as wearing a mask and social distancing.

We are additionally concerned that mass utilization of LSDT has the potential to negatively impact the public perception and trust of medical tests in general. There are already many in the U.S. population that distrust some medical interventions (vaccines, for example) or existing tests (Lyme disease serology, for example). It is easy to envision a testing scenario with widespread daily use by untrained users in which a substantial proportion of LSDT results for subjects who actually have the virus may be wrong and

the majority of positive results will be wrong. If such a test is widely endorsed by the medical and laboratory communities but is ultimately considered by the public to be unreliable, it could lead to a dangerous erosion of the public trust in diagnostic tests in general. The ongoing saga of hydroxychloroquine as a potential therapy for COVID-19 may serve as a cautionary tale. There were promising *in vitro* data suggesting an antiviral effect of the drug (12), and small studies in China and an uncontrolled, nonrandomized trial in France were used by prominent physicians and scientists (and politicians) to suggest that hydroxychloroquine could help make the virus "disappear" (13). It did not (14).

Finally, if people do perform self-testing, some might decide not to report the results to their physician or to public health authorities. This would compromise the quality of the data needed to track the pandemic. Furthermore, asymptomatic people who have a positive result might not self-quarantine, perhaps because they do not believe the test results or perhaps because they need to go to work, attend in-person classes, or provide care to others.

Can insensitive tests accurately detect infectious individuals? One argument in favor of frequent use of rapid tests is that the insensitivity of some of these tests, primarily antigen tests, is acceptable because the tests still detect people who have high levels of virus and are therefore infectious (7). This raises two issues. First, do tests that could be used as LSDT specifically detect specimens with high viral loads? The data addressing this are limited, but one antigen test was recently found to detect only 82% of nasopharyngeal samples with threshold cycle (C_T) values below 25 in a group of well-validated NAATs for SARS-CoV-2 (8). These low C_T values indicate that high levels of viral RNA were present, and so this antigen test does not reliably detect samples that are likely to contain high levels of SARS-CoV-2. Second, suppose that we hypothesize that less-sensitive tests would reliably detect high levels of virus: does that mean that they would detect infectious individuals? Perhaps, but this is speculation. It is clear that samples with lower C_T values are more likely to contain SARS-CoV-2 that can be detected in viral culture (15, 16). But there are no data linking the C_T value or viral quantity to transmissibility, and it is possible that viral culture may be an insensitive indicator of the potential of an infected person to transmit SARS-CoV-2. Furthermore, there is no clear separation between samples that are infectious or not for viral culture using the C_T value, as there is significant overlap in the C_T values of samples that are positive and negative in viral culture (16). We do not dispute that people who have high levels of SARS-CoV-2 in their respiratory secretions are more likely to be infectious than those with low levels of virus, but we do question whether insensitive LSDTs can reliably detect whether someone is likely to be infectious.

Can we crush the curve without widespread testing? Inexpensive, reasonably achievable measures such as use of masks, hand hygiene, staying home when ill, and avoiding close contact are important in preventing transmission of SARS-CoV-2 (17–19) (https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html). In this context, the primary role of testing should be to monitor the effects of these well-established practices, in addition to its obvious role in patient care of symptomatic individuals and infection control prior to medical procedures or hospital admission. But using testing to prevent transmission of SARS-CoV-2 on a large scale is like using the weather report to prevent global warming.

While the idea of widespread, frequent use of LSDTs is appealing and while they may yet prove effective at reducing transmission in selected settings, there is considerable cause for caution in extrapolating predicted performance and utility from models or results of antigen-based testing for other respiratory viruses in the hands of testing professionals. We urge those promoting or considering LSDT strategies to demonstrate utility in a real-world trial and, at a minimum, to make lab-based performance characteristics publicly available before prompting the general public to petition the U.S. or state governments to approve and deploy LSDTs.

REFERENCES

1. Kotlikoff LJ, Mina M. 3 July 2020. A cheap, simple way to control the coronavirus. The New York Times, New York, NY. https://www.nytimes.com/2020/07/03/opinion/coronavirus-tests.html.
2. Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. 2020. Variation in false-negative rate of reverse transcriptase polymerase chain reaction-based SARS-CoV-2 tests by time since exposure. Ann Intern Med 173:262–267. https://doi.org/10.7326/M20-1495.
3. Timbrook TT, Morton JB, McConeghy KW, Caffrey AR, Mylonakis E, LaPlante KL. 2017. The effect of molecular rapid diagnostic testing on clinical outcomes in bloodstream infections: a systematic review and meta-analysis. Clin Infect Dis 64:15–23. https://doi.org/10.1093/cid/ciw649.
4. Broughton JP, Deng X, Yu G, Fasching CL, Servellita V, Singh J, Miao X, Streithorst JA, Granados A, Sotomayor-Gonzalez A, Zorn K, Gopez A, Hsu E, Gu W, Miller S, Pan CY, Guevara H, Wadford DA, Chen JS, Chiu CY. 2020. CRISPR-Cas12-based detection of SARS-CoV-2. Nat Biotechnol 38:870–874. https://doi.org/10.1038/s41587-020-0513-4.
5. Joung J, Ladha A, Saito M, Segel M, Bruneau R, Huang MW, Kim NG, Yu X, Li J, Walker BD, Greninger AL, Jerome KR, Gootenberg JS, Abudayyeh OO, Zhang F. 2020. Point-of-care testing for COVID-19 using SHERLOCK diagnostics. medRxiv https://doi.org/10.1101/2020.05.04.20091231.
6. Paltiel AD, Zheng A, Walensky RP. 2020. Assessment of SARS-CoV-2 screening strategies to permit the safe reopening of college campuses in the United States. JAMA Netw Open 3:e2016818. https://doi.org/10.1001/jamanetworkopen.2020.16818.
7. Larremore DB, Wilder B, Lester E, Shehata S, Burke JM, Hay JA, Tambe M, Mina MJ, Parker R. 2020. Test sensitivity is secondary to frequency and turnaround time for COVID-19 surveillance. medRxiv https://doi.org/10.1101/2020.06.22.20136309.
8. Lambert-Niclot S, Cuffel A, Le Pape S, Vauloup-Fellous C, Morand-Joubert L, Roque-Afonso AM, Le Goff J, Delaugerre C. 2020. Evaluation of a rapid diagnostic assay for detection of SARS-CoV-2 antigen in nasopharyngeal swabs. J Clin Microbiol 58:e00977-20. https://doi.org/10.1128/JCM.00977-20.
9. Chartrand C, Leeflang MM, Minion J, Brewer T, Pai M. 2012. Accuracy of rapid influenza diagnostic tests: a meta-analysis. Ann Intern Med 156:500–511. https://doi.org/10.7326/0003-4819-156-7-201204030-00403.
10. Antoniol S, Fidouh N, Ghazali A, Ichou H, Bouzid D, Kenway P, Choquet C, Visseaux B, Casalino E, Emergency Department study group on respiratory viruses. 2018. Diagnostic performances of the Xpert® Flu PCR test and the OSOM® immunochromatographic rapid test for influenza A and B virus among adult patients in the Emergency Department. J Clin Virol 99–100:5–9. https://doi.org/10.1016/j.jcv.2017.12.005.
11. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schunemann HJ, COVID-19 Systematic Urgent Review Group Effort (SURGE) study authors. 2020. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. Lancet 395:1973–1987. https://doi.org/10.1016/S0140-6736(20)31142-9.
12. Andreani J, Le Bideau M, Duflot I, Jardot P, Rolland C, Boxberger M, Wurtz N, Rolain JM, Colson P, La Scola B, Raoult D. 2020. In vitro testing of combined hydroxychloroquine and azithromycin on SARS-CoV-2 shows synergistic effect. Microb Pathog 145:104228. https://doi.org/10.1016/j.micpath.2020.104228.
13. Gautret P, Lagier JC, Parola P, Hoang VT, Meddeb L, Mailhe M, Doudier B, Courjon J, Giordanengo V, Vieira VE, Tissot Dupont H, Honore S, Colson P, Chabriere E, La Scola B, Rolain JM, Brouqui P, Raoult D. 2020. Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. Int J Antimicrob Agents 56:105949. https://doi.org/10.1016/j.ijantimicag.2020.105949.
14. Boulware DR, Pullen MF, Bangdiwala AS, Pastick KA, Lofgren SM, Okafor EC, Skipper CP, Nascene AA, Nicol MR, Abassi M, Engen NW, Cheng MP, LaBar D, Lother SA, MacKenzie LJ, Drobot G, Marten N, Zarychanski R, Kelly LE, Schwartz IS, McDonald EG, Rajasingham R, Lee TC, Hullsiek KH. 2020. A randomized trial of hydroxychloroquine as postexposure prophylaxis for Covid-19. N Engl J Med 383:517–525. https://doi.org/10.1056/NEJMoa2016638.
15. La Scola B, Le Bideau M, Andreani J, Hoang VT, Grimaldier C, Colson P, Gautret P, Raoult D. 2020. Viral RNA load as determined by cell culture as a management tool for discharge of SARS-CoV-2 patients from infectious disease wards. Eur J Clin Microbiol Infect Dis 39:1059–1061. https://doi.org/10.1007/s10096-020-03913-9.
16. Huang CG, Lee KM, Hsiao MJ, Yang SL, Huang PN, Gong YN, Hsieh TH, Huang PW, Lin YJ, Liu YC, Tsao KC, Shih SR. 2020. Culture-based virus isolation to evaluate potential infectivity of clinical specimens tested for COVID-19. J Clin Microbiol 58:e01068-20. https://doi.org/10.1128/JCM.01068-20.
17. Lyu W, Wehby GL. 2020. Community use of face masks and COVID-19: evidence from a natural experiment of state mandates in the US. Health Aff (Millwood) 39:1419–1425. https://doi.org/10.1377/hlthaff.2020.00818.
18. Schwartz KL, Murti M, Finkelstein M, Leis JA, Fitzgerald-Husek A, Bourns L, Meghani H, Saunders A, Allen V, Yaffe B. 2020. Lack of COVID-19 transmission on an international flight. CMAJ 192:E410. https://doi.org/10.1503/cmaj.75015.
19. Hendrix MJ, Walde C, Findley K, Trotman R. 2020. Absence of apparent transmission of SARS-CoV-2 from two stylists after exposure at a hair salon with a universal face covering policy - Springfield, Missouri, May 2020. MMWR Morb Mortal Wkly Rep 69:930–932. https://doi.org/10.15585/mmwr.mm6928e2.