Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company’s public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Validity of at-home rapid antigen lateral flow assay and artificial intelligence read to detect SARS-CoV-2

Shannon Richardson, Michael A. Kohn, Jenna Bollyky, Julie Parsonnet

ARTICLE INFO

Article history:
Received 21 January 2022
Revised in revised form 28 June 2022
Accepted 2 July 2022
Available online 7 July 2022

Keywords:
COVID-19
SARS-CoV-2
Diagnostic accuracy
Validity
Rapid antigen test
Artificial intelligence

ABSTRACT

Background: The gold standard for COVID-19 diagnosis—reverse transcriptase polymerase chain reaction (RT-PCR)—is expensive and often slow to yield results whereas lateral flow tests can lack sensitivity.

Methods: We tested a rapid, lateral flow antigen (LFA) assay with artificial intelligence read (LFAIR) in subjects from COVID-19 treatment trials (N = 37; daily tests for 5 days) and from a population-based study (N = 88; single test). LFAIR was compared to RT-PCR from same-day samples.

Results: Using each participant’s first sample, LFAIR showed 86.2% sensitivity (95% CI 73.6%–98.8%) and 94.3% specificity (88.8%–99.7%) compared to RT-PCR. Adjusting for days since symptom onset and repeat testing, specificity was 97.8% (89.9%–99.5%) on the first symptomatic day and decreased with each additional day. Sensitivity improved with artificial intelligence (AI) read (86.2%) compared to the human eye (71.4%).

Conclusion: LFAIR showed improved accuracy compared to LFA alone, particularly early in infection.

© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

1. Introduction

The COVID-19 pandemic is a global public health crisis with over 500 million cases and 6 million deaths globally to date [1]. To contain infectious outbreaks like COVID-19, rapid and accurate diagnostic testing is critical. Reverse transcriptase polymerase chain reaction (RT-PCR) is the current reference standard [2]. Although it is the reference standard, RT-PCR testing falls short of being a perfect test for several reasons. Even where widely available, RT-PCR can take hours to complete and days to get results. In the United States (US), turnaround time for RT-PCR is 1 to 2 days during which time, infection is likely to be transmitted [2,3]. RT-PCR is also comparatively expensive, requiring trained laboratory technicians. The average cost of one RT-PCR test ranges from $100 to $300 in the US [4]. In low- and middle-income countries (LMICs), cost and personnel constraints can make RT-PCR screening for COVID prohibitive [5–7]. Thus, despite high analytic sensitivity, RT-PCR tests are ineffective for surveillance because of slow turnaround and limited availability [8].

Antigen tests have been touted as a better screening tool to detect infectious agents like SARS-CoV-2 [9]. Antigen tests are less expensive than RT-PCR; the average cost of antigen tests in the US is $25 per test kit with the global cost around $5 [6]. The World Health Organization (WHO) also predicts the price of these antigen tests will decrease [6]. Additionally, antigen tests do not require trained health care professionals and can provide results in 15 to 20 minutes. This rapid time to results facilitates quicker isolation of infected individuals. A recent study assessing various surveillance approaches through simulation models found that the speed of case identification was far more important than test sensitivity [10]. With this in mind, the CDC recently published interim guidelines for the use of antigen tests instead of PCR as an effective testing solution [11]. To date, 4 antigen tests have earned Emergency Use Approval (EUA) from the U.S. Federal and Drug Administration (FDA) for at-home use [9,12]. Although sensitivity and specificity of these tests in symptomatic adults has been published (range 64%–90% and 85.7%–99%, respectively), little is known about their performance throughout the course of illness in comparison to RT-PCR [13–15]. This comparison is of particular interest since RT-PCR often detects non-viable or non-transmissible virus late in the course of illness [16].

Two biotechnology companies, Cellex and Exa Health, have recently partnered to create a twenty-minute, lateral flow antigen test that is read by an artificial intelligence algorithm accessed through a smartphone application (Lateral Flow with Al Read; LFAIR). This rapid test can provide results at home, but also, if desired, simultaneously transmit results to health care providers and public health authorities. Therefore, the LFAIR has the potential to make...
SARS-CoV-2 testing available to anyone with a smartphone and a cellular connection, whether in the US or in resource limited settings.

We assessed the diagnostic accuracy of the Cellex and Exa Health, LFAIR to detect SARS-CoV-2 compared to the reference standard, RT-PCR. We then conducted analyses to assess how test results relate to days since symptom onset, test the sensitivity of AI compared to users’ interpretations, and SARS-CoV-2 viral load as approximated by RT-PCR cycle threshold (CT) values.

2. Material and methods

2.1. Participants

Participants, 125 in total, were recruited from 2 study populations that were obtaining nasal swabs for RT-PCR. These participants included:

1) Thirty-seven from COVID treatment trials [NCT04524663 and NCT04346628] that included outpatients with mild to moderate COVID-19 symptoms who had tested positive for SARS-CoV-2 via RT-PCR within 72 hours of enrollment. These subjects underwent anterior nares swab by study personnel on days 1, 5, and 10 of study participation and swabbed themselves on days 2 to 4 and 6 to 9. Accumulated samples were refrigerated in viral transport medium before being sent for RT-PCR on days 1, 5 and 10. Subjects in this group were provided with 5 kits and asked to conduct them on 5 days on which nasal swabs were obtained for RT-PCR.

2) 88 from a longitudinal surveillance study of COVID-19 incidence (TrackCOVID study). These subjects were not known to be infected with SARS-CoV-2. The subjects were provided with a single LFAIR kit to be performed at home on the same day that study personnel swabbed the subjects’ anterior nares and sent the swabs for RT-PCR.

Inclusion criteria included fluency in English, age over 18 years, and reported comfort using iPhone applications. Exclusion criteria included receiving treatment with convalescent plasma or other antibody therapy related to SARS-CoV-2 infections, participating in a SARS-CoV-2 vaccine study, and testing positive for SARS-CoV-2 more than 2 weeks prior to enrollment without an interim negative RT-PCR test. Eligible participants were recruited from December 2020 to March 2021.

All subjects provided informed consent and the study was approved by Stanford’s Institutional Review Board (protocol #58444).

2.2. Data collection

Participants completed a survey with demographic information, symptom status, and if symptomatic, date of symptom onset and of prior positive RT-PCR. If a participant did not own an iPhone, one was loaned to them.

Each home test kit had a unique quick response (QR) code that was recorded before participants took the kit home. To simulate individuals purchasing kits at a store, researchers did not provide verbal instructions for the rapid antigen test. A brief instruction sheet directed subjects to scan the QR code to download the Exa Health app, which was only compatible with iPhones at the time of the study. The app then provided detailed video instructions to complete the test. A timer within the application indicated when 15 minutes had transpired and an image of the LFAIR should be uploaded to Exa Health with the smartphone camera. Only the image, no PHI or clinical information, was transferred. Participants were not provided with their results.

Results for RT-PCR assays, including CT values, were provided to this study by the PIs of the parent trials using study IDs connected to the QR codes. Subjects in NCT04346628 had their RT-PCR assays performed by Quest which amplified 2 sequences of the nucleocapsid gene (N1 and N3). All other RT-PCR assays were performed at the Stanford clinical lab using an RT-PCR of the E gene. Deidentified antigen test results, including AI binary results, the intensity score for the LFA antigen test line, and user read, were received from Exa Health with their attached QR codes.

2.3. LFAIR assay

At-home antigen test kits include a SARS-CoV-2 antigen test device, a buffer solution, a sterile swab, an extraction tube, and an instruction sheet. The plastic antigen test device includes the LFA test strip, consisting of a sample pad, reagent pad, reaction membrane, and absorbing pad. The reagent pad is made of colloidal-gold conjugated with monoclonal antibodies that bind to SARS-CoV-2 nucleocapsid protein (N protein); the N protein is detectable in upper respiratory specimens during the acute phase of SARS-CoV-2 infection [9].

After swabbing their noses, subjects place the nasal swab into the buffer solution. After 1 minute, 3 drops of the buffer are added to the sample well on the test device. When SARS-CoV-2 antigen is present in the sample, the antibody conjugates and the N protein will form an antibody-antigen complex, which is captured by the monoclonal antibodies on the test line (T) region. Presence of the T line indicates a positive test result. All valid test strips must also have a control line (C) to indicate that the sample had sufficient volume of buffer and that membrane wicking was successful (Fig. 1) [17].

The artificial intelligence algorithm (Exa Health, Inc., Los Altos, CA) determines if a test is positive, negative, or invalid by determining the intensity of both the C and T lines. Test results are stored as either positive, negative, or invalid. The algorithm also stores the test line intensity value, which reflects the amount of SARS-CoV-2 antigen bound to the antibody in the LFA device. This test line intensity value ranges from 0 (no line visible) to 1.0 (most intense line). We also collected the test users’ interpretations of their results. Participants were prompted through the smartphone application to indicate if they could see a line next to the control (C) and test (T) markers.

2.4. Statistical analysis

We report distribution and mean (standard deviation) of participant demographic and symptom characteristics, infection status determined by RT-PCR results, RT-PCR CT values, and antigen test line intensity.

RT-PCR results were used as the reference test to determine the sensitivity and specificity with 95% confidence intervals for each participant’s first antigen test regardless of symptom status.

Since SARS-CoV-2 positive participants underwent multiple reference and index tests, ranging from 1 to 5 pairs of tests, we also

---

**Fig. 1.** Example antigen test device results after sample has developed for 15 minutes. C indicates control line and T indicates test line. Invalid test strips do not have a C line, meaning there was insufficient sample volume or unsuccessful membrane wicking [18].
estimated sensitivity using a repeated-measures analysis including each index-reference pair completed on the same day. To understand how the LFAIR sensitivity varies over time, we stratified test sensitivity over days since symptom onset based on subject self-report. We also used the generalized estimating equation (GEE) approach to estimate sensitivity with robust standard errors for correlated observations [18-20]. We implemented a GEE model with a binary distribution and logit link function.

In subjects from the treatment cohort, we compared the sensitivity of LFAIR with that of the test user’s reading of the C and T lines using McNemar’s Exact test. We also used a linear regression to investigate the relationship between antigen test line intensity and RT-PCR CT values. Finally, we determined the LFAIR diagnostic accuracy compared to RT-PCR results stratified by participant age and education level. We created dichotomous variables for participant age (<40 and ≥40 years) and educational attainment (less than a 4 years college degree or college graduate). Data was analyzed using statistical programs R (Version 1.2.5019) and Stata (Version 16.1, StataCorp, College Station, TX).

3. Results

Of 37 participants recruited from the COVID treatment trials, 33 (89%) completed at least 1 antigen test successfully with a total of 119 assays performed (mean 3.6 per person). Paired RT-PCR results were available for 110 of these 119 LFAIR results (mean 3.3 per person). Of the 88 individuals from the longitudinal TrackCOVID study, 66 (75%) completed their LFAIR successfully and all 66 had paired PCR results.

The mean age of participants was 47.7 years (SD = 14.8) and 55.1% were female (Table 1). Of the 185 rapid LFA antigen tests completed, 1 yielded an invalid result and was removed from subsequent analysis.

Overall, 29 participants—28 from the COVID treatment trials—were positive by RT-PCR on at least 1 swab. Among the 70 SARS-CoV-2 negative participants, 65 (92.9%) were recruited from the TrackCOVID population-based study. Only 1 SARS-CoV-2 positive individual, recruited from the TrackCOVID trials, was asymptomatic.

The mean (SD) CT values for the N1 gene, N3 gene, E gene, and RNaseP control sequence were 23 (4.9), 24 (4.9), 24.7 (6.8), and 21.3 (3.2), respectively. Stanford University and Quest report any CT values of 40 or less as suspicious for SARS-CoV-2.

4. Discussion

In our study cohort we found the LFAIR test had a sensitivity of 86.2% and specificity of 94.3% for the first LFAIR test completed, ignoring days since first symptom onset. These values are similar to the reported sensitivities and specificities of extant EUA-approved tests that range from 64% to 95% and 96% to 100%, respectively.

Because we only had sufficient samples with E gene CT values (58, compared to 14 for the N1 and N3 genes), we assessed the relationship of the E gene CT value with the LFAIR test line intensity. We detected a linear association between the E gene CT value and LFAIR test line intensity ($r = -0.53, P < 0.001$) (Fig. 3). A decrease in LFAIR test line was associated with an increase in CT value, a proxy for viral load.

We did not observe any significant differences in sensitivity and specificity of the LFAIR by age group or level of educational attainment (Supplementary Appendix, Tables S4 and S5).

Table 1

| Characteristics, n (%) | Participants (N = 98) |
|------------------------|-----------------------|
| Age                    |                       |
| 18-24                  | 1 (1.0%)              |
| 25-44                  | 42 (42.9%)            |
| 45-64                  | 41 (41.8%)            |
| 65+                    | 14 (14.3%)            |
| Race                   |                       |
| Asian                  | 19 (19.4%)            |
| Black                  | 4 (4.1%)              |
| White                  | 58 (59.2%)            |
| Other                  | 10 (10.2%)            |
| Ethnicity              |                       |
| Hispanic/Latino        | 17 (17.3%)            |
| Not Hispanic/Latino    | 81 (82.7%)            |
| Sex                    |                       |
| Female                 | 54 (55.1%)            |
| Male                   | 44 (44.9%)            |
| Education              |                       |
| ≥High school degree    | 4 (4.1%)              |
| Associates degree      | 10 (10.2%)            |
| Four-year college degree| 41 (41.8%)            |
| Graduate degree        | 43 (43.5%)            |
binary answer. However, the variance in the intensity was considerable, and at least for now, LFAIR should be treated as a binary test. Although our study was not adequately powered to assess the relationship of symptom type with LFAIR sensitivity or intensity, our findings suggest additional research into the relationship between symptoms and LFA accuracy is warranted.

A key strength of our study is the real-life conditions of our testing. Participants completed all of their LFAIR tests at-home, which is the setting for which these tests were developed.

Although these conclusions point towards the LFAIR being a useful public health tool for COVID-19, we note that the generalizability of our findings is limited by spectrum bias, iPhone application use, and low racial and ethnic diversity of our sample. For spectrum bias, since we only had one asymptomatic SARS-CoV-2-positive individual, we could not estimate sensitivity in asymptomatic individuals. Moreover, most participants were tested 3 or more days after symptom onset, when viral load peaks. Therefore, it is unclear how sensitive the at-home LFA rapid antigen test is in the first 2 days after symptom onset. This spectrum bias may mean that LFAIR’s sensitivity will be lower in practice [21].

Additionally, the diversity of our cohort does not adequately reflect the general public, limiting the generalizability of our findings. The similar LFAIR test diagnostic accuracy by age and education, however, suggest that the test can be used by a wide variety of individuals. That most positive patients had already received a confirmed diagnosis of SARS-CoV-2 also introduces the possibility of bias and subsequent user interpretation bias could have underestimated the discrepancy between the LFAIR and human eye sensitivity.

We should note how the persistence of RT-PCR positivity after the infectious period might affect our results. Since RT-PCR tests can be positive even when an individual is no longer contagious, some LFAIR false negatives that occurred in patients with non-viable or non-transmissible virus might actually be considered true negatives. We tried to assess the importance of this consideration by looking at the test sensitivity stratified by days since symptom onset.

Fig. 2. At-home LFA rapid antigen test sensitivity by days since symptom onset. Sensitivities are illustrated with open circles with confidence intervals included. Data for this figure is in the Supplementary Appendix (Table S1). LFA = lateral flow antigen.

![Fig. 2](image)

Fig. 3. E gene CT value plotted by LFAIR test (T) line strength. E gene CT value associated with test line strength ($r = -0.53; P < 0.001$). LFAIR = Lateral Flow with AI Read.

![Fig. 3](image)
Results from this study indicate that this LFAIR test could be integral to improving SARS-CoV-2 screening and slowing the spread of and reducing the deaths caused by SARS-CoV-2. Moreover, this rapid testing has the potential to improve access to diagnostic testing because compared to the current reference standard, RT-PCR, it is less expensive and can produce test results in the absence of laboratories and medical personnel. This LFAIR test would be particularly useful in resource limited settings. For example, in developing countries with high smartphone use, the LFAIR test would enable diagnostic results to be sent from individuals’ phones directly to public health officials and physicians.

5. Conclusions

In summary, this lateral flow assay, when performed at home, has good sensitivity and specificity, particularly when read with artificial intelligence rather than with the human eye. The test appears most sensitive early in the course of illness when viral load is highest and transmission most likely. Because the artificial intelligence algorithm will improve as more testing is completed, and because results can be automatically transmitted to health department and health care providers, we expect this and similar assays will become useful tools for controlling SARS-CoV-2 in areas with high smartphone coverage.

Funding

This work was funded by Exa Health, Inc. (IC-2020 SPO 205666, Los Altos, CA). The company was not involved in the data analysis or in manuscript preparation or writing. The artificial intelligence algorithm that Exa Health produced for the smartphone application was set at the start of the study; none of the study’s data was used to develop the algorithm.

Authors statement

The authors contributed to the paper in the following manner: Shannon Richardson: Data curation, formal analysis, investigation, visualization, Roles/writing – original draft. Michael Kohn: Formal analysis, Methodology, Visualization, Writing—review and editing. Jenna Bollyky: Methodology, Project administration, Supervision, Writing – review & editing. Julie Parsonnet: Formal analysis, Funding Acquisition, Investigation, Methodology, Project administration, Writing – Reviewing and Editing

Authors’ contributions

Shannon Richardson: study design, data collection, data analysis and interpretation, drafted manuscript, and revised manuscript
Michael Kohn: data analysis and revised manuscript
Jenna Bollyky: data collection and revised manuscript
Julie Parsonnet: study design, data collection, and revised manuscript

Declaration of competing interest

The authors report no conflicts of interest relevant to this article

Acknowledgments

We thank Yvonne A Maldonado, MD and Upinder Singh, MD for their collaboration with the TrackCOVID study.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.diagmicrobio.2022.115763.

References

[1] Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 2020;20(5):533–4. doi: 10.1016/S1473-3099(20)30120-1.
[2] Weissleder R, Lee H, Ko J, Pittet MJ. COVID-19 diagnostics in context. Sci Transl Med 2020;12(546):1–7. doi: 10.1126/scitranslmed.abc1931.
[3] Sallam M. Covid-19 vaccine hesitancy worldwide: a concisive systematic review of vaccine acceptance rates. Vaccines 2021;9(2):1–15. doi: 10.3390/vaccines9020160.
[4] Krouse S, Abbott B. What kind of covid test should i get? Answers on cost, accuracy and more. Wall St J 2020;1–7. https://www.wsj.com/articles/covid-19-tests-answers-on-cost-accuracy-and-turnaround-time-11599143378. Accessed August 30, 2022.
[5] Bhoury P, Azar E, Hitti E. COVID-19 response in lebanon: current experience and challenges in a low-resource setting. JAMA J Am Med Assoc 2020;324(6):548–9. doi: 10.1001/jama.2020.12695.
[6] Mahase E. Covid-19: 120 million rapid tests pledged to low and middle income countries. BMJ 2020;371:1015. doi: 10.1136/bmj.m1957.
[7] Kobia F, Gitaka J. COVID-19: are Africa’s diagnostic challenges blunting response effectiveness? AAS Open Res 2020;3(4):1–11. doi: 10.12688/aasopenres.10061.1.
[8] Mina MJ, Parker R, Larremore DB. Rethinking Covid-19 test sensitivity — A strategy for containment. N Engl J Med 2020;383(22):e120. doi: 10.1056/nejmp2025631.
[9] Ravi N, Cortade DL, Ng E, Wang SX. Diagnostics for SARS-CoV-2 detection: a comprehensive review of the FDA-EUA COVID-19 testing landscape. Biosens Bioelectron 2020;165:112454. doi: 10.1016/j.bios.2020.112454.
[10] Larremore DB, Wilder B, Lester E, Shehata S, Burke JM, Hay JA, et al. Test sensitiv- ity is secondary to frequency and turnaround time for COVID-19 screening. Sci Adv 2021;7(1). doi: 10.1126/sciadv.abg5395.
[11] Centers for Disease Control and Prevention. Interim guidance for antigen testing for SARS-CoV-2. Centers Dis Control Prevent 2020. https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antigen-tests-guidelines.html. Accessed July 1, 2021.
[12] FDA. In vitro diagnostics EUAs - antigen diagnostic tests for SARS-CoV-2. FDA, 2021. p. 1-5. Available from: https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices/in-vitro-diagnostics-euas-antigen-diagnostic-tests-sars-cov-2. Accessed August 30, 2022.
[13] Prince-Guerra JL, Almendares O, Nolen LD, Gunn JK, Dale AP, Bueno SA, et al. Evaluation of Abbott BinaxNOW® rapid antigen test for SARS-CoV-2 infection at two community-based testing sites — Pima County, Arizona, November 3–17, 2020. MMWR Morb Mortal Wkly Rep 2021;70(3):106–5. doi: 10.15585/mmwr.mm7003e3.
[14] Pollock NR, Tran K, Jacobs JR, Cranston AE, Smith S, O’Connell MR, et al. Performance and operational evaluation of the access bio carestart rapid antigen test in a high-throughput drive-through community testing site in Massachusetts. Open Forum Infect Dis 2021;8(1):1–7. doi: 10.1093/ofid/ofab243.
[15] BD. BD Veritor SARS-CoV-2 System : easy-to-use, portable rapid point-of-care antigen diagnostic test in just 15 minutes. https://www.bd.com/en-za/about-bd/news-and-media/press-releases/jun-30-2021-bd-veritor-sars-cov-2-system. Published 2021.
[16] van Kampen JIA, van de Vijver DAMC, Kraaij PLA, Haagsman BM, Lamers MM, Okha N, et al. Duration and key determinants of infectious virus shedding in hospitalized patients with coronavirus disease-2019 (COVID-19). Nat Commun 2021;12(1):1. doi: 10.1038/s41467-020-20568-4.
[17] Spartan Medical. Cellex QSARS-CoV-2 Antigen Rapid Test 2022.
[18] Liang KY, Zeger SL. Longitudinal data analysis using GLM. Biometrika 1986;73(1):13–22. doi: 10.1093/biomet/73.1.13.
[19] Zeger L, Liang S. Longitudinal data analysis for discrete and continuous outcomes author(s): Scott L. Zeger and kung-yee liang published by: international biometric society stable. Biometrics 1986;42(1):121–30.
[20] Genders TSS, Sprock S, Sijen T, Sypherd SW, Lesaffre E, Hunink MGM. Methods for calculating sensitivity and specificity of clustered data: a tutorial. Radiol- ogy 2012;265(3):910–6. doi: 10.1148/rd.12120509.
[21] Kohn MA, Carpenter CR, Newman TB. Understanding the direction of bias in studies of diagnostic test accuracy. Acad Emerg Med 2013;20(11):1194–206. doi: 10.1111/acem.12255.