(PIVD) For SARS-CoV-2 in Vulnerable and Underserved Populations in Chelsea, MA

GRANT #02282022 RUF
Executive Abstract
Project Period [April, 2022 -February 2023]

This Abstract provides an overview of the project's goals, methods, and results. It outlines the specific aims of the project, the data and background methods used, and the results and developments achieved.

The project aimed to create tools for quantifying and comparing the real-world performance of different COVID-19 antigen tests. This was achieved through a prospective clinical study conducted by IDx20, which evaluated the performance of several antigen tests using laboratory methods and real-world conditions. The study employed computer image processing and mathematical probabilistic modeling techniques to quantify test results.

The specific aims of the project were as follows:

1. Validate antigen tests in the laboratory using visual and quantitative methods under controlled conditions. This involved analyzing visual assessments and digital signal intensity to establish the limit of detection and assess test performance.
2. Measure clinical positive and negative percent agreement data compared to the gold standard detection method (qRT-PCR) for three antigen tests self-administered by participants. The study also evaluated positive and negative agreement for up to eight days in participants who contracted COVID-19 during the study.
3. Compare laboratory evaluation and real-world visual binary assessment performance statistics to establish a relationship between lab-measured performance and real-world performance. The goal was to develop rapid, laboratory-based methods for assessing test performance without the need for clinical or real-life testing.

4. Evaluate the usability of antigen tests for self-testing and assess the results of a weekly testing regimen in low socioeconomic status groups and racial or ethnic minority populations.
5. Develop a digital reporting system for collecting and reporting test results in real-time. The platform allows users to log their antigen test results, review testing history, and provides various reporting options.

The study collected health questionnaires and at-home antigen tests, with images of the tests uploaded to a digital database. The results and developments of the project included:

- Laboratory evaluations showed performance differences among different antigen test brands.
- The laboratory and real-world assessments provided equivalent results in evaluating test performance.
- The study successfully predicted the real-world performance of antigen tests based on input information and the relational model.
- Longitudinal data analysis showed a correlation between SARS CoV-2 genome copy number and test line pixel intensity.
- Analytical tools were developed for generating limits of detection and quantifying performance.
- Usability research conducted with participants identified areas for improvement in test design and instructions.
- The insights gained from user research can inform the development of more accessible and user-friendly antigen tests.

Overall, the project demonstrated new techniques for evaluating the performance of COVID-19 antigen tests, including computerized calculations and mathematical modeling. These techniques offer advantages in comparing data among different clinical trials and assessing the performance of antigen tests in real-world conditions. The project's findings have implications for future testing programs and can help mitigate the spread of COVID-19.

Data and background methods

The clinical study involved data from enrolled participants ranging from 12 to over 90 years old in real-world test settings. The data collected consisted of health questionnaires and at home antigen tests (AT) done each week. The images of the antigen tests were uploaded to a digital database in real time at the moment of testing. Of the total data, a subset of 1047 AT images were paired with information from qRT-PCR of the same participant and same sample date.

The immunological basis of a lateral flow test relies on the principles of antigen-antibody interactions. Antigens are substances, such as the nucleoprotein of SARS CoV-2. The lateral flow immunoassay consists of a nitrocellulose membrane that can absorb liquids

through its network of pores. The sample contains the antigens that react with the gold nanoparticles. These colloidal spheres are very small in the nanometer scale and form a colloidal liquid which is red-purple color and have an antibody component that specifically binds the virus protein. The complex formed of nanoparticle-antibody and antigen migrates into the white color strip of nitrocellulose by capillarity force. The test line or detector area is the region on the strip that contains another specific antibody and can capture the nanoparticle-colored complex, therefore forming a visible detection line in that area of the strip. The number of nanoparticles captured in that area results in a color intensity that correlates with the amount of target or virus load of the sample.

The virus protein contained in the epithelial cells and nasal secretions is captured by the nasal swab; if the sample is positive the test line appears visible to the eye. The technological novelty we demonstrate here is to be able to calculate the amount of virus by using the test line area color intensity measured via computer vision by capturing an image of the test. The other component of the test strip is a control line that captures all the remaining of the gold nanoparticles and it is used to validate the test and it shows when the sample has migrated properly into the test.

For the laboratory analysis of the AT performance, we utilized heat inactivated virus and nucleoprotein dilutions to perform the image analysis of the tests. We analyzed the test band of the cassette (i.e. the band of the nitrocellulose strip that detects the presence of the virus). We calculate the signal intensity as the difference of the brightness of the white background of the strip minus the brightness of the test band, and normalize this value (in the range 0-1) along the dilution series. The signal intensity increases with the amount of virus in the sample. To quantify this relationship, we use a modified Langmuir model with parameters calibrated to the normalized signal intensity (NSI) of the test laboratory data.

In the analysis of the naked-eye (i.e., binary) data obtained from the clinical study, we quantify the *probability of positive agreement as a function* of the qRT-PCR cycles (Ct). – the probability positive agreement (PoPA)-is known to be strongly dependent on the viral load of the sample. We model the PoPA function with a logistic function and estimate its parameters by regression analysis to fit the real-world binary data. In contrast to reporting the overall positive percent agreement, also known as PPA, the PoPA function is presented in our project as the alternative to describe the test performance in real-world conditions. We exemplify with antigen tests that the PPA value is strongly dependent on the circumstantial distribution of viral load in the set of participants, whereas the PoPA function is independent of this distribution – providing a robust standard for performance comparison.

The relational model formulated combined the laboratory evaluation and additional information to predict the real-world performance of the ATs. It is developed on a Bayesian network framework, which relates the performance of the tests across the different variable domains involved: NSI, target protein concentration, inactivated virus concentration and PCR cycles. The network represents the dependent components of the model and their calibrated relationships. In the Bayesian framework the model

parameters and test outcomes are treated as random variables, described by probability density functions (PDFs). This allows our model to account for full uncertainty calculations.

Results and developments

We performed laboratory evaluations based on the signal intensity analysis on eight test brands for Covid-19. Performance results were compared across the AT brands, using the estimated NSI curves against protein concentration, inactivated virus concentration and limits of detection. The AT brands showed important performance differences in these analyses as expected. We also performed in the laboratory naked-eye assessments and corresponding PoPA function calculations on the dilution series and compared the evaluation with the NSI evaluation. Both types of evaluation provided equivalent results in performance across the brands.

Following the study protocol, we provided real-world data to evaluate the performance of a final number of three test brands, by means of the basic statistics (e.g. PPA and NPA) and the PoPA function. These data also allowed us to define, in the domain or variable of normalized signal intensity (NSI), the visual acuity in assessment of the AT by the participants, by means of the limit of detection PDF of the participants' population.

The laboratory NSI calibrated models for the ATs were further used in mathematical modeling to simulate the PoPA function and PPA in real-world conditions for the three brands involved in the clinical study. The types of information represented in the network are: (1) a PDF as function of NSI describing the visual acuity of the participants in assessing the AT positiveness, (2) the laboratory model of NSI as function of target protein concentration, (3) the laboratory model of NSI as function of inactivated virus concentration, and (4) the calibrated model relating viral concentration and PCR cycles. We showed successful prediction of the real-world performance of the antigen tests based on the input information and the relational model, by comparing the expected and observed PPA for the three test brands analyzed.

Using serial day collection during the acute phase of the infection, our study showed a high correlation between SARS CoV-2 genome copy number by beans of "delta" Ct and the test line pixel intensity (red- purple color of the test line if we visualize it by naked eye). The longitudinal data analysis would inform the user or health provider about the viral load as well as the duration of the infection during the acute phase, and may contribute to the diagnosis and prognosis via a simple testing implementation tool.

The analytical tools presented here support the potential use for quantitative endpoints beyond the current binary (yes or no) outcome. The calculator tools for generating the limits of detection (LoD) are available to the Reagan Udall Foundation and the FDA via the web links provided in this report.

Intensity analysis web-tool: This tool is designed to be used by analysts from test manufacturers' laboratories or regulatory agencies. It quantifies the performance of the ATs in laboratory conditions. The input data is a two-column list with virus or target protein at various concentrations and the corresponding AT signal intensity of the test band. The devices should be photographed and images processed to generate the normalized signal intensity (NSI) file. The web-tool adjusts the NSI model (we use a Langmuir-Freundlich function) to the experimental data and estimates the limit of detection.

http://35.174.80.89:3838/idx20_pixel_intensity_analysis/

Probability of positive agreement web-tool: This tool is designed for analysts from research laboratories, test manufacturers or regulatory agencies. It quantifies the performance of the ATs based on naked-eye assessments. It estimates the PoPA as function of various eligible variables: qRT-PCR Cts, viral concentrations from a tissue culture infection concentration (plaque forming units or tissue culture infectious dose TCID50 titers), protein concentration or NSIs. The input data is a table with the naked-eye assessment (0 or 1) and the corresponding related variable. These data are generated either in the laboratory or in real-world conditions. The PoPA function is estimated by a logistic regression of the naked-eye assessment data.

http://35.174.80.89:3838/idx20_probability_positive_agreement/

We proved the usability of the ATs in self-testing and self-reporting using an informatic platform by participants of the clinical study, the setting included elderly vulnerable population of low socioeconomic status. In addition, we developed a platform for digital data collection in real-time and improved the user frontends.

Our user research for this project was divided into two phases with the first being foundational research and the second one being evaluative research. User observations revealed that 19.8% of participants had difficulty completing certain steps of the test procedure itself due to physical limitations. Most common limitations were in opening the swab package, removing the cap from the solution tube, placing the correct amount of sample droplets, checking the expiration data and hand washing. These insights provide valuable information for designing testing programs that address participants' motivations, pain points, and frustrations.

Conclusions

We demonstrated new techniques for obtaining in vitro antigen test's PPA calculations, or probability of positive agreement (PoPA), as a function of the test band pixel intensity and qPCR-PCR Ct cycles. Also, we unfolded a computational methodology, based on mathematical modeling, to infer the real-world positive agreement performance of ATs based on laboratory evaluation. Our techniques would have several advantages, including the computerized calculations, allowing comparisons of data among different clinical trials, with unequal number samples, from a range of virus concentrations and

from different brands of tests. The AT performance real-world prediction may be useful for assessment prior to long and costly common clinical validations. The computerized calculations based on image analysis have applications for comparative performances of the commercial manufactured lots of antigen tests and may prove to be useful for fast assessment of post-market monitoring of these types of tests.

The study was executed with the participation of senior housing, with over 65% Hispanic and Latino population facing economic challenges and vulnerability. A large portion of residents living in Chelsea, MA are of working class, low wage industries such as service, hospitality and construction who were severely affected during the COVID-19 pandemic, in this context the study provided extra layer of positive intervention for a vulnerable community.

We offer methodologies that potentially would improve the validation process for antigen tests in general.

Glossary

We describe below some of the technical terminology to facilitate the communication of our methods and results.

Test band: In a common antigen test device, refers to the band of the nitrocellulose strip that will be colored to signal the presence of the target pathogen in the sample. Commonly the location is marked with a “T” letter in one side of the device.

White background: In a common test device, refers to a zone of the nitrocellulose strip that will not be colored in the presence of the sample. We commonly select a zone between the test band and the control band.

Signal intensity: In the current context, refers to the average pixel brightness of the white background minus the average pixel brightness of the test band of the device strip. The brightness is measured in the digital image (photo) of the AT device fixed at the moment of the result reading by the user. The brightness is given in the range 0-255. In a RGB color photograph the brightness correspond to the average of the R, G and B colors.

Normalized signal intensity: Signal intensity after normalization in the range 0-1. For comparison, processing and analysis of a set of signal intensities, we normalize the signal intensity values based on the minimum and maximum values within the set.

Binary or *naked-eye assessment*: Refers to the positive or negative interpretation of the antigen test result produced by a human observing the device in plain sight after the manufacturer recommended time of reading the result after the activation with the sample.

Image processing: Refers to the process of extracting the signal intensity from the digital images of the test device. We employed in this project a software that allows our

laboratory analysts to upload the digital image and conduct a semi-automated selection of the test band and white background, and calculation of the signal intensity.

Known positive sample: We consider a sample to be known as positive in either of the following two contexts. In laboratory context, the known positive sample corresponds to samples containing the target recombinant protein, inactivated or active pathogen. In real-world context, the known positive sample corresponds to samples taken by a participant that has a qRT-PCR positive result with a sample taken the same day as the AT sample.

Probability of Positive Agreement (PoPA): Refers to the probability that a human observer of the AT interprets the result as positive in a device activated with a known positive sample.

Positive percent agreement (PPA): The positive percent agreement is the fraction of positive naked-eye assessments divided by the total number of tests expressed in percent, for a given set of ATs activated with known positive samples. In general, the PPA is different to the PoPA, due to the finite support of the sample set (number N of ATs in the set); the PoPA is the expected value of the PPA. For increasing N, the PPA converges to the PoPA.

Bayesian methods, modeling or inference: Refers to a branch of statistical science that represents knowledge as probability functions in parameter spaces. The parameters are considered random variables (i.e., have different outcomes when evaluated) that obey a probability model. The probability functions are updated according to the available information and data. In a common formulation, the method starts by incorporating prior information about the parameters and combines them with the likelihood function derived from the observed data. Through this combination, a posterior distribution is obtained, representing the updated knowledge about the parameters. This posterior distribution enables quantification of uncertainty and facilitates improved decision-making. By providing a framework to integrate prior information and iteratively update knowledge with new data, Bayesian modeling offers a powerful tool for robust statistical analysis.

Bayesian network: Generalizes Bayesian inference to complex problems with multiple types of datasets, information and model components. It describes the relationships across model components, and between model components and data components, using a graphical representation, given by direct acyclic graphs i.e. arrows that connect the model components according to causal dependencies. The state of the model parameters and the relationships are described by probability density functions. Bayesian networks are the framework of a large part of Artificial Intelligence tools in a wide spectrum of applications.