Executive Summary

Motivation. In an ideal world, each new COVID-19 antigen test under consideration for approval by the FDA would be tested on many patient groups to assure benefit for all people at the time the test becomes available. In reality, such evaluation is impossible. It is reasonable to expect a test to perform equally well across many patient groups, but this is a hypothesis that must be tested. We reasoned that we could test this hypothesis without multiple trials by computationally applying the analytical sensitivity of the test in question, in the form of its limit of detection (LOD; the concentration of virus that can be detected with 95% confidence) to real-world patient data.

Methods. The real-world data we used was composed of viral loads, patient data from our hospital's electronic health record, a SARS-CoV-2 contagiousness threshold, and a small all-comers trial of two specific antigen tests with viral loads on samples taken at the same time. (Contagiousness, also known as infectiousness, is the likelihood that a person will be able to spread the virus, which is related to the amount of virus in a person.) Our specific aims were to test for viral-load differences in different patient groups and to characterize the two antigen tests, using in vitro infectivity testing and whole-virus sequencing to ensure our findings generalize temporally (to recent phases of the pandemic) and geographically (from eastern Massachusetts to the country).

Results. We accomplished these aims by constructing a public web application (https://arnaoutlab.org/coviral/). In this application, the user selects patient groups of interest—a cohort defined by demographics, socioeconomic status, comorbidities, treatment history, vaccination status, medication use, and so on—and either a named antigen test (if one of the two we trialed) or the LOD (for any antigen test, including the ones we trialed). The application then displays (*i*) the distribution of viral loads for those groups and (*ii*) the antigen test's clinical sensitivity for detecting contagiousness. (The rationale for use of clinical sensitivity is these tests are used for screening, so as not to miss infections.) These results are presented graphically as well as with p-values for comparisons.

Viral load distributions (or histograms) from different patient groups were largely similar, as judged by high p-value. As a result, the sensitivity and specificity of the two antigen tests, which have similar LODs, were also very similar across patient groups. This observation supports the proposal that separate trials are by and large not necessary for evidence-based approval of antigen tests (although there may still be some patient groups that require trials), grounding this conclusion in real-world data from close to 50,000 COVID-19 positive test results. Cohorts for which sensitivity would have been lower or higher were indicated by testing, without need for trials.

Nevertheless, the web application revealed numerous interesting cases where distributions differed in ways that were not expected (e.g. by self-reported race/ethnicity) or did not differ where differences were expected (e.g. by smoking status or pulmonary disease). It also revealed numerous cases where distributions followed predicted patterns (e.g. that survivors and patients who died with COVID-19 as an incidental finding had lower viral loads than patients who died from COVID-19). These observations illustrated the value of the web application for

generating hypotheses and more generally its value as a public research tool.

The OTC antigen tests that have been widely available on the market since 2021 are considerably less sensitive than RT-qPCR for detecting SARS-CoV-2 infection.[6] However, because their LODs are generally above the contagiousness threshold, they are quite sensitive for detecting contagiousness.[6] Based on our clinical experience, we hypothesized that antigen tests would perform quite similarly on different patient groups and subgroups; this hypothesis was largely supported. The BinaxNow COVID-19 Ag Card and CareStart COVID-19 Antigen Home Test had similar sensitivities, generally in the 0.90-0.93 range.

Conclusions. All major goals of this project were met. We demonstrated proof of principle that the appropriate real-world data can be used for rapid and highly cost-effective "in silico" trials that obviate the need for slow/expensive traditional trials. We showed that group-specific COVID-19 antigen-test trials are largely unnecessary. And, via data download and p-values (to quantify similarities/differences), we illustrated the value of the web as a discovery tool for unlocking the value in clinical laboratory data, setting the stage for future expansion and further development of such tools. Limitations include incomplete data for presentation and vaccination status, and that the variation of viral load over the course of disease were not was not explicitly evaluated.