

# The New York Times

## ***Coronavirus Antibody Tests: Can You Trust the Results?***

A team of scientists worked around the clock to evaluate 14 antibody tests. A few worked as advertised. Most did not.



*From left, Dr. Caryn Bern, Dr. Alexander Marson, Patrick Hsu and Dr. Jeffrey Whitman, who have helped lead an effort to assess antibody tests for the coronavirus. Credit...Anastasiia Sapon for The New York Times*

**By Apoorva Mandavilli**

- Published April 24, 2020 - Updated April 25, 2020, 8:56 a.m. ET

# The New York Times

The researchers worked around the clock, in shifts of three to five hours, hoping to stave off weariness and keep their minds sharp for the delicate task.

They set up lines of laboratory volunteers: medical residents, postdoctoral students, even experienced veterans of science, each handling a specific task. They checked and rechecked their data, as if the world were depending on it. Because in some ways, it is.

For the past few weeks, more than 50 scientists have been working diligently to do something that the Food and Drug Administration mostly has not: Verifying that 14 coronavirus antibody tests now on the market actually deliver accurate results.

These tests are crucial to reopening the economy, but public health experts have raised urgent concerns about their quality. The new research, completed just days ago and posted online Friday, confirmed some of those fears: Of the 14 tests, only three [delivered consistently reliable results](#). Even the best had some flaws.

The research has not been peer-reviewed and is subject to revision. But the results are already raising difficult questions about the course of the epidemic.

Surveys of residents in the Bay Area, Los Angeles and New York this week found that substantial percentages tested positive for antibodies to SARS-CoV-2, the official name of the new coronavirus. In New York City, the figure was said to be as high as 21 percent. Elsewhere, it was closer to 3 percent.

The idea that many residents in some parts of the country have already been exposed to the virus has wide implications. At the least, the finding could greatly complicate plans to reopen the economy.

Already Americans are scrambling to take antibody tests to see if they might escape lockdowns. Public health experts are wondering if those with positive results might be allowed to return to work.

But these tactics mean nothing if the test results can't be trusted.

In the new research, researchers found that only one of the tests never delivered a so-called false positive — that is, it never mistakenly signaled antibodies in people who did not have them.

Two other tests did not deliver false-positive results 99 percent of the time.

But the converse was not true. Even these three tests detected antibodies in infected people only 90 percent of the time, at best.

The false-positive metric is particularly important. The result may lead people to believe themselves immune to the virus when they are not, and to put themselves in danger by abandoning social distancing and other protective measures.

It is also the result on which scientists are most divided.

# The New York Times

“There are multiple tests that look reasonable and promising,” said Dr. Alexander Marson, an immunologist at the University of California, San Francisco, and one of the project’s leaders. “That’s some reason for optimism.”

Dr. Marson is also an investigator in the Chan Zuckerberg Biohub, which partly funded the study.

Other scientists were less sanguine than Dr. Marson. Four of the tests produced false-positive rates ranging from 11 percent to 16 percent; many of the rest hovered around 5 percent.

“Those numbers are just unacceptable,” said Scott Hensley, a microbiologist at the University of Pennsylvania. “The tone of the paper is, ‘Look how good the tests are.’ But I look at these data, and I don’t really see that.”

The proportion of people in the United States who have been exposed to the coronavirus is likely to be 5 percent or less, Dr. Hensley said. “If your kit has a 3 percent false-positive, how do you interpret that? It’s basically impossible,” he said. “If your kit has 14 percent false positive, it’s useless.”

Dr. Hensley said the study nonetheless was well designed and the results pressing, given the sudden proliferation of antibody tests on the market and the push to use them to lift lockdowns.

“I think this is exactly the kind of study that we need right now,” he said.

Dr. Marson and his colleagues said they were drawn to the study for that very reason.

As universities in the Bay Area shut down all research not related to the coronavirus, some researchers began focusing on ways to improve diagnostic tests for SARS-CoV-2.

Dr. Marson and his collaborator, Patrick Hsu, a bioengineer at the University of California, Berkeley, anticipated that antibody tests would also face questions about quality.

In mid-March, Dr. Hsu heard that a friend, a venture capitalist who owns a network of 1,000 community clinics in the New York area, had ordered thousands of rapid antibody tests. Investors and entrepreneurs seemed to be distributing them around San Francisco, too.

“I realized, ‘Gosh this is really the Wild West,’” Dr. Hsu said. “We needed to figure out which of these would really work.”

The duo recruited Dr. Jeffrey Whitman and Dr. Caryn Bern, who last year published an [analysis of antibody tests](#) for Chagas disease. Other graduate students and postdoctoral fellows volunteered to help perform the evaluations.

The team began with a modified version of the method Dr. Whitman had devised to validate Chagas tests. The researchers created a biosafety-certified space, obtained the needed approvals and procured hundreds of blood samples from two Bay Area hospitals.

They also purchased tests from Chinese manufacturers, clearing customs regulations and sometimes accepting Uber deliveries in the middle of the night.

# The New York Times

In all, the investigators analyzed 10 rapid tests that deliver a yes-no signal for antibodies, and two tests using a lab technique known as Elisa that indicate the amount of antibodies present and are generally considered to be more reliable.

Suited up in protective gear, the team worked in shifts of three to five hours in a sort of socially distanced factory line.

One researcher spotted the test with a blood sample, and another added the necessary chemical solutions; then two independent readers looked at the test, and a last person recorded the results. Still other team members analyzed the results, sometimes working through the night.

In the early hours of recent mornings, they handed the baton to Dr. Tyler Miller and his colleagues at Massachusetts General Hospital, who were conducting a slightly different analysis of three tests, including one evaluated in San Francisco.

The Bay Area team finished evaluating 12 tests in record time, less than a month. By comparison, the Chagas project required a team of three people working for more than a year just to compare four tests.

Having a study design already in hand helped speed the work, but there was one key difference. Decades of data have shown that Chagas disease elicits lifelong immunity. For this study, the team had no idea how quickly SARS-CoV-2 antibodies might turn up in the blood, or at what levels.

New tests also usually compete with an established gold standard. Lacking such a standard, the team instead compared all the tests in a “head-to-head bake-off to see who’s the winner,” Dr. Hsu said.

Each test was evaluated with the same set of blood samples: from 80 people known to be infected with the coronavirus, at different points after infection; 108 samples donated before the pandemic; and 52 samples from people who were positive for other viral infections but had tested negative for SARS-CoV-2.

Tests made by Sure Biotech and Wondfo Biotech, along with an in-house Elisa test, produced the fewest false positives.

A test made by Bioperfectus detected antibodies in 100 percent of the infected samples, but only after three weeks of infection. None of the tests did better than 80 percent until that time period, which was longer than expected, Dr. Hsu said.

The lesson is that the tests are less likely to produce false negatives the longer ago the initial infection occurred, he said.

The tests were particularly variable when looking for a transient antibody that comes up soon after infection, called IgM, and more consistent in identifying a subsequent antibody, called IgG, that may signal longer-term immunity.

“You can see that antibody levels rise at different points for every patient,” Dr. Hsu said. The tests performed best when the researchers assessed both types of antibodies together. None of the tests

# The New York Times

could say whether the presence of these antibodies means a person is protected from reinfection, however.

The results overall are promising, Dr. Marson added. “There are multiple tests that have specificities greater than 95 percent.”

Rapid antibody tests are generally used to get a simple yes-no result, but the team assigned the positive results — which appear as bands on a test strip — a score from zero to six. They trained readers to interpret those results, and found their decisions often agreed and were supported by the more quantitative Elisa tests.

“If you train the readers well, they can start to be reliable,” Dr. Marson said of rapid tests. “That is critical to understand if these tests could ever be deployed.”

The team at Mass General set a higher bar for specificity; they considered a score of one for the intensity of a band to be a negative result, rather than a score of zero.

Perhaps because they eliminated the fainter bands — the ones most likely to be erroneous — their estimate of specificity for BioMedomics, the one test that was evaluated by both teams, was more than 99 percent, compared with the San Francisco team’s estimate of 87 percent.

Other experts were skeptical of the scoring approach, however. “That’s not really a method that would give you a real quantitation,” said Florian Krammer of the Icahn School of Medicine at Mount Sinai in New York.

Dr. Krammer has developed a two-step Elisa test that he said has 100 percent specificity and delivers a measure of the quantity of IgM and IgG antibodies a person has. Scoring a rapid test’s bands might offer some data for a scientific study, he said, “but I would not make any decisions based on that.”

Dr. Krammer said false positives are less of an issue for assessing how widely the virus has spread in the population. If a test has a known false-positive rate, scientists can factor that into their calculations, he said.

But false positives become dangerous when making policy and personal decisions about who can go back to work. “You don’t want anybody back to work who has a false positive — that’s the last thing you want to do,” Dr. Krammer said.

Scanwell Health, a Los Angeles-based start-up, has ordered millions of test kits from Innovita, a Chinese manufacturer, and has applied to the Food and Drug Administration to market the tests for at-home use.

In the new study, the Innovita test detected antibodies in 83 percent of infected people and yielded a false-positive rate of 4 percent.

# The New York Times

Dr. Jack Jeng, chief medical officer of Scanwell Health, said the study looked at an earlier version of Innovita's test and not the "newer, improved version" his company had ordered. "It will be interesting to see how it performs," he said.

Dr. Marson and his colleagues have acquired tests from nearly 100 manufacturers, and plan to continue comparing them. The scientists also hope to expand their sample set to include people who were mildly ill or did not feel ill at all, and to stratify their data by age and the presence of chronic conditions.

"This is just the beginning," Dr. Marson said. "Our goal would be to keep going till we feel there's adequate supply in the market."