

Commentary

# Dr. Jay Bhattacharya Reveals Stanford University's Attempts To Derail COVID Studies

Paul D. Thacker, Jay Bhattacharya<sup>1a</sup>

<sup>1</sup> Stanford University

Keywords: Covid, Research, Media

<https://doi.org/10.35122/001c.88046>

In a lengthy first-person account, Professor Bhattacharya defends himself, colleagues, and family from attacks, and details apparent commercial interests driving Stanford science and media hit pieces.

Shortly after COVID reached US shores, the World Health Organization estimated that the virus would kill 11.2 million Americans—an alarming and frightening number that panicked public health officials. Today, we know that a much smaller number of Americans actually died, meaning the world's experts got it wrong.

But researchers at Stanford doubted those early numbers, and set out to run a study that would find how deadly the virus actually was, and how many people were being infected. Called the Santa Clara Study—for the California county where Stanford's professor ran their tests—the research was met with interference by Stanford administrators and a series of excoriating and sloppy articles in the now defunct newsite BuzzFeed.

Stanford Professor Jay Bhattacharya found himself at the center of this controversy and reveals, for the first time, what happened. Critics inside Stanford had commercial interests blinding them to the public health importance of the Santa Clara Study and a desire to protect the university from inconvenient research that might discover pandemic policies were misguided. While trying to shut down research by Bhattacharya and colleagues through Stanford's bureaucracy, internal Stanford critics appear to have also launched public attacks through strategic leaks to BuzzFeed reporter Stephanie M. Lee.

Following the money trail and naming names, Bhattacharya recounts getting his research published despite being under the microscope of Stanford administrators, and muses on the future of science and public health.

Bhattacharya's first-person account of the Santa Clara Study follows.

## PROFESSOR JAY BHATTACHARYA'S SANTA CLARA STUDY

### INTRODUCTION

The Santa Clara seroprevalence study was the first major paper to look for Covid antibodies in a large U.S. population



**Figure 1**

center. I am the study's senior author and a professor of 20+ years at Stanford's medical school.

When my colleagues and I released the study as a medRxiv preprint in mid-April 2020, it generated enormous public, media, and scientific interest. Every major newspaper, television station, and scientific news outlet covered our research. Within days of its release, I received hundreds of emails from academics worldwide and hundreds of questions and comments from reporters and laypeople. Social media began buzzing about the study, with over 20,000 accounts tweeting about it. It is the third most discussed medRxiv preprint in history and the 55th most discussed scientific paper ever. It is now published in the peer-reviewed journal, *International Journal of Epidemiology*.

Why all the attention? The study arrived shortly after much of the world had imposed an extraordinary lockdown – “two weeks to flatten the curve” – that shut down ordinary life for countless people. The Santa Clara study results had several enormous implications for the lockdown strategy:

1. The Covid virus had spread to at least 2.8% of the population after only a little more than a month of its arrival in the U.S. and despite the lockdown;
2. Most who had been infected and recovered had not come to the attention of public health authorities, and nearly 40% recalled no symptoms;
3. Infection fatality rate (IFR), or chance of infected people dying, was much lower than previously thought; and

a Jay Bhattacharya, MD, PhD

4. The pandemic had a long way to go before the end, and nearly everyone would become infected.

Many experts – especially those who had insisted a lockdown was necessary – did not want to believe these major implications.

My colleagues and I faced an unprecedented attack in the media and by some vocal scientists, who questioned the results and even our good faith conduct. This attack took place very publicly on Twitter and in some particular media outlets, like the now-defunct BuzzFeed News.

What remained hidden until now is that Stanford University administrative leadership intervened and interfered with the study in ways that violated accepted standards of academic freedom and harmed our ability to convey and defend our research findings. I believe their actions were partly motivated by a financial incentive to develop and sell antibody tests and partly to protect the Stanford University brand against media criticism. Violating academic freedom harmed public health policy, reverberating throughout the pandemic and continuing to this day.

This is my first-hand account of this untold story. I hope it will serve as a cautionary tale so that science, the media, and universities can function better in a future pandemic.

## A HYPOTHESIS

In early March 2020, the World Health Organization (WHO) announced a covid fatality rate estimate of 3.4% to the world.<sup>1</sup> With a population of around 332 million, the implication was that over 11 million people could have been expected to die, just in the United States.

For comparison, the typical seasonal flu has a case fatality rate that is more than ten times less—in the range of 0.1%, according to CDC data.<sup>2</sup> For the 2019-2020 flu season, the CDC estimates imply a case fatality rate of 0.07% and a fatality rate of 0.13% for the 2016-2017 season.

A crucial question is how many more people get infected but are not detected since they have no symptoms or are not tested.

I have been writing and publishing about infectious diseases for nearly my entire career. When I first realized in late 2019 that the covid outbreak might be a pandemic, my first thoughts went back to the 2009 H1N1 flu pandemic. When the H1N1 flu hit in 2009, I was interested in understanding how people reacted to the mortality threat of a new infectious disease. One of the curious things I noticed was a sharp difference between scientists' initial estimates of H1N1's lethality and what was later published in the scientific literature. The initial estimates were based on laboratory-confirmed cases of H1N1 and inferred a catastrophically high mortality rate—above 5%.

Later estimates were based on population samples of antibody levels—meaning a test confirmed people had been sick with the virus. This method found an infection fatality rate of around 0.01%—more than 100 times lower. This also meant that 99.99% of infected patients survived. The latter investigations are known as seroprevalence studies – “sero” derives from the Latin word for the watery compo-

nent of blood where the antibodies measured are found, and “prevalence” means frequency in a population.

The figure above –reproduced from a comprehensive study of mortality from the H1N1 pandemic published in 2014 – shows the disparity.<sup>3</sup> The blue squares depict mortality rates measured based on lab-confirmed active cases of H1N1 flu, while the red triangles are from seroprevalence studies published months after the pandemic began.

The y-axis depicts mortality rates on a log scale. Because this is a logarithmic scale, even slight differences show vast mortality rate changes. The seroprevalence studies showed mortality hundreds of times lower than the earlier case-based estimates.

The key reason for the difference between earlier and later studies is that earlier studies measured infection based on who came to the attention of public health authorities because they required medical care. But not everyone who was infected saw a doctor, meaning the actual number of infected was much higher. People with milder illnesses stayed home and were never counted. Undercounting the number of infected made the mortality rate appear higher than it actually was.

The seroprevalence studies avoid this problem by looking for evidence of infections based on antibodies in the blood. This procedure catches everyone, whether public health has identified them as sick or not. It also gives a more accurate and lower estimate of the fatality rate.

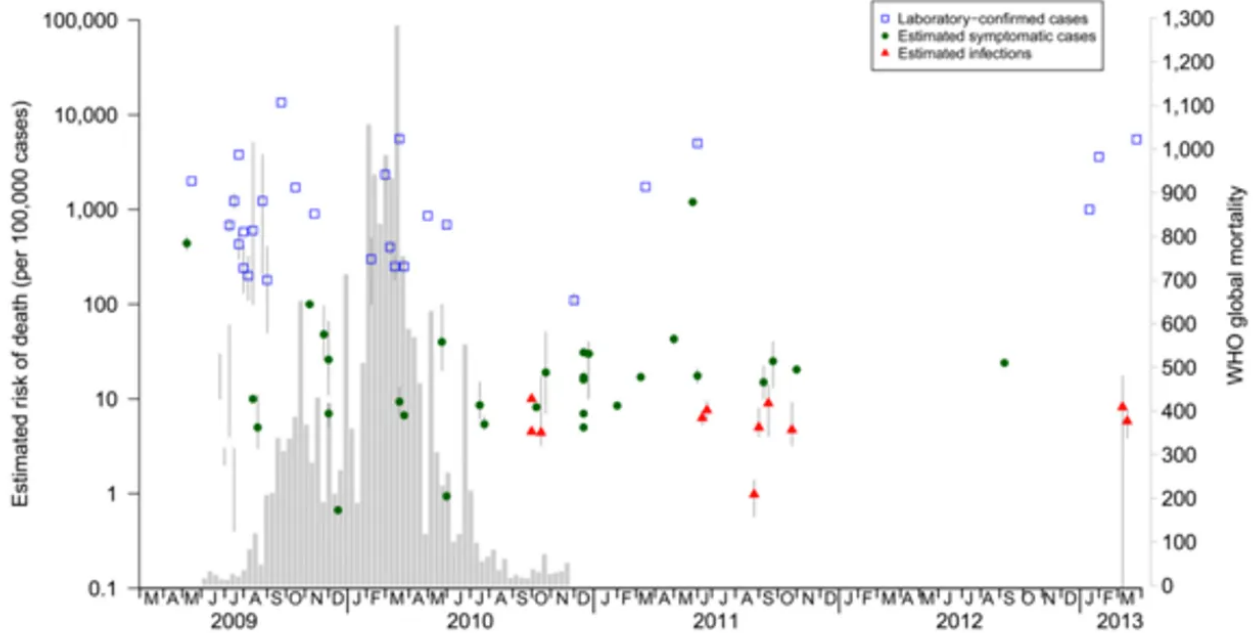
As it turns out, the SARS-CoV-2 virus that causes covid is similar in some ways to H1N1 in that many infected people have only a mild sickness or show no symptoms. And just like H1N1, the WHO and the media promulgated an initial fatality rate estimate based on a calculation that failed to account for people who got infected but showed no symptoms or never got sick enough to see their doctor. This bias skewed the fatality rate reports higher, making people think the virus to be more deadly than it actually is.

All this H1N1 background informed my thinking in February 2020. Turning this around, maybe this meant covid infections were much more common, and the virus had already spread more widely than public health officials realized.

If my scientific hypothesis was right, then covid might be simultaneously more infectious than previously understood, but also less deadly. Again, an accurate assessment of the spread and danger of the virus is vital for public health. And we had to get that right to know how to respond.

## AN OP-ED CALLING FOR A SEROPREVALENCE STUDY

In the middle of March 2020, governments worldwide instituted extraordinary lockdowns and stay-at-home orders to halt the spread of the virus. They reasoned that slowing infections would prevent the overcrowding of hospital systems already seen in Wuhan, China, and Lombardy, Italy. This policy responded to disease forecasting models like the Imperial College London's infamous Report 13, which predicted two million deaths in the U.S. occurring within



**Figure 2**

months.<sup>4</sup> A key parameter of that model was the IFR of the virus. And I knew this number could not be known unless there was a good seroprevalence study to estimate it.

With my colleague at Stanford, Eran Bendavid, I wrote an essay for the Wall Street Journal on March 24th, 2020, in which we discussed the sparse data we could find on the disease's infectiousness and its fatality rate. Bendavid is an infectious disease doctor, an expert in global health policy, and my former student. Extrapolating from information available out of Iceland, the National Basketball Association, and the doomed Diamond Princess cruise, we concluded that it was impossible to say whether covid had an infection fatality rate as high as 5% or as low as 0.01%.

Doing a very rough calculation, that meant that if every American got COVID, as few as 34,000 could die or as many as 17 million. If our hypothesis was correct, it was unlikely to be the latter. But neither we nor anyone else knew for certain, so we called for an immediate study of disease prevalence to clear things up.

The week before, another colleague at Stanford – the world-famous epidemiologist and scientist John Ioannidis – had published an article in Stat News making similar calculations regarding the uncertainty in prevalence and lethality of the virus.<sup>5</sup> Ioannidis suggested a wide plausible range of 0.05-1% of the IFR. Another colleague, the former vice dean of research at the Sol Price Public Policy School at the University of Southern California (and also a former student and good friend), Neeraj Sood, published another piece on March 15th in the Wall Street Journal, calling for random testing of populations with the same aims.

Together, we were calling for an immediate seroprevalence study to resolve the scientific uncertainty and to understand whether the lockdowns we were following had any chance of success. To this day, I am stunned by the failure

of the US Centers for Disease Control to run a nationwide seroprevalence study in March 2020. Not running a study was a catastrophic failure that hampered the ability of authorities to design an appropriate pandemic response.

## MAJOR LEAGUE BASEBALL AND TEST KITS FROM CHINA

The WSJ op-ed upended my life.

Immediately on its publication, I started receiving emails from people worldwide questioning my motives, including scientists, journalists, politicians, and even fellow professors at Stanford. Though the piece contained not a single word about the economic harms of lockdown, people wrote asking me, in effect, why I valued human life so little that I would trade economic well-being for money. Critics wrote that the op-ed committed the cardinal sin of appearing to support President Trump, though it did not mention him, and I do not recall seeing him ever proposing a seroprevalence survey. The op-ed was my public expression of a critical public health question about the infectiousness and the fatality rate of covid that needed immediate testing.

The op-ed also produced constructive responses from experts worldwide, most of whom I had never previously met, with some offering to help organize a seroprevalence study. Dr. Daniel Eichner, the head of the Sports Medicine Research and Testing Laboratory in Utah, had ordered thousands of lateral flow test kits from a company in China to check covid antibody levels for his clients in Major League Baseball. He offered to donate these kits so we could run the serosurvey. With this donation, we committed ourselves to running our study as rapidly as possible, because



**Figure 3**

the information was so important to inform public health policy.

Dr. Eichner had placed orders with two companies selling these kits – Biomedonics and Premier Biotech. Because the number of infected was likely low so early in the pandemic, we needed the kits to have a high accuracy and a very low false positive rate. Tests with a high error rate might falsely indicate that many people in the community had previously been infected and throw off our results.

In mid-March, the FDA approved the kits for research use in the U.S., but the agency had not evaluated them for accuracy. The information provided by both manufacturers suggested to me that these kits would work, but we would need to confirm the manufacturers' claims.

The antibody kits themselves look like pregnancy test kits. They require just a drop of blood collected from a finger prick and placed in a shallow well. A negative test will show a single line that indicates that the test was valid, but no antibodies were found. A positive test will show both the control stripe and one or two additional stripes. Positive lines indicate the presence of two different types of covid-specific antibodies. Technically, these are IgG & IgM stripes named after different antibodies that form after covid infection. The presence of either IgG or IgM line is evidence of prior covid infection.

Though Dr. Eicher had placed orders for these tests in early March, the kits were stuck awaiting shipment in a port in China, a victim of the Chinese lockdown. Biomedonics kept promising delivery but ultimately never came through. Luckily, the Premier Biotech test kits arrived just a week before we were set to collect study samples. We eventually used these test kits for three seroprevalence studies: the Santa Clara County study, a study led by Sood in L.A. County conducted a week later, and a nationwide study of

Major League Baseball employees conducted a week after that.

## ORGANIZING THE SANTA CLARA STUDY

Bendavid organized and directed the Santa Clara County Study. For data, we collected finger-prick blood samples from over 3,300 volunteers while strict stay-at-home orders were in place. Under normal circumstances, such a study would have taken a year or more to organize. Working night and day, Bendavid and the team accomplished this task within weeks.

The study team comprised hundreds of volunteers, including Stanford medical and doctoral students, to draw and analyze blood samples. Stanford undergraduates (including my daughter, then a Stanford sophomore) and others helped set up the three drive-through testing sites and directed traffic during blood collection days. Other volunteers like Jim Tedrow, owner and operator of the CRG Laboratories in Oklahoma City, flew out to Stanford to oversee blood sample testing for antibodies. Dr. Andrew Bogan, a molecular biologist and a successful Silicon Valley entrepreneur, helped run one of the drive-through sites and contributed to interpreting the study results. Even my wife, Dr. Catherine Su, volunteered, arranging to have the parking lot of my church serve as a drive-through testing site.

Bendavid contacted Dr. Sara Cody, the chief health officer of Santa Clara County, to solicit the help of Santa Clara County Public Health and ultimately secured its cooperation for our study on March 28th, 2020.

My experience in running studies before covid typically involved small research teams working for months or years on a project. This was the first time I had been part of an enormous volunteer effort, and the feeling of a community chipping in their talents to conduct vital scientific research was exhilarating. My pride in the Stanford community was bursting on the eve we began data collection.

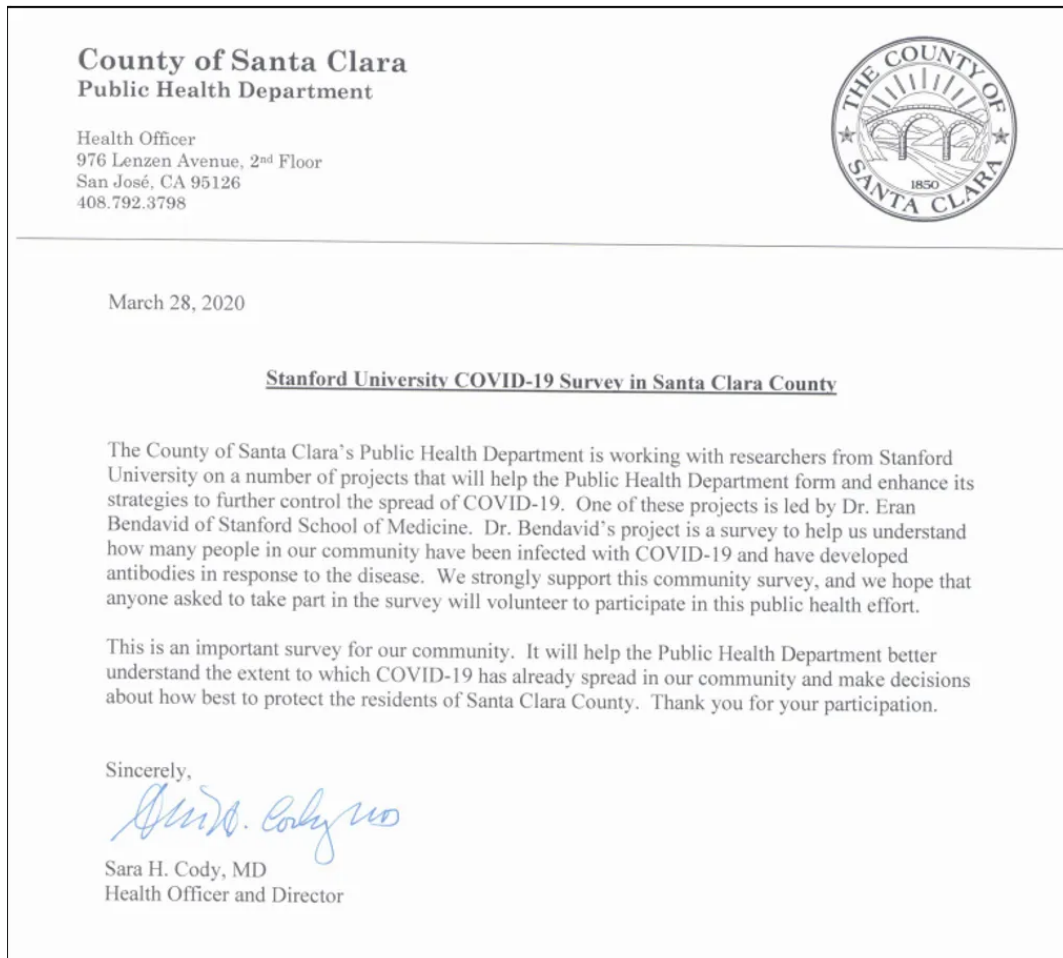
But that good feeling would not survive the month.

## FACEBOOK RECRUITING

An early outbreak in Seattle made it clear that the elderly in nursing homes were at grave risk from covid infection, but we did not want to inadvertently spread the virus by going to nursing homes and sampling this population. Instead, we decided to recruit a representative community-dwelling sample from across Santa Clara County. But identifying a representative sample of people was next to impossible as the county had an order to shelter-in-place.

We considered two different sampling schemes. The first involved sending recruiters to wait in front of grocery stores to find people willing to give a finger prick. Grocery stores were one of the only buildings open for public access during the shelter-in-place order. But we rejected this option. Vulnerable people were less likely to risk leaving their homes, which would bias our sample against them. Also, this process would put our recruiters in direct contact with study subjects, putting them at risk of infection and pos-





**Figure 4**

sibly turning our study into a focal point of community spread.

We settled on recruiting through Facebook. There was some precedent in the survey literature for this strategy, though it was not the norm. However, this scheme would limit face-to-face interactions and possible infections. Also, Facebook would give us demographic information about participants, including the zip code where they lived. We could then correct for imbalances in population representation through statistical reweighting.

Any study involving human subjects requires oversight by a review committee whose mandate is to protect the rights and interests of the study participants. The Santa Clara study was no exception. We worked closely with Stanford's human subjects review board and the Santa Clara County Department of Public Health to craft a recruiting message.

At the time, scientists were uncertain whether antibody presence indicated immune protection against reinfection. So, we crafted a script that made clear that our goal was to measure the extent of disease spread in the population, not to provide individual medical testing. Nevertheless, I received many emails from people – including professors on campus – who asked to be included in the study because they were curious whether they had antibodies. We denied

these requests for ethical reasons and to avoid bias in the sampling scheme.

I did not myself take an antibody test until months later because I felt it was wrong to turn away so many people and then test myself or someone in my family.

Interest in the study in the more affluent zip codes of Santa Clara went viral shortly after the official Facebook ads went live. During the initial hours, we also heard rumors that people were directing friends to the recruitment advertisement. To avoid biasing our samples, we did not include these potential subjects as study participants. I later learned that my wife, Dr. Su, sent an email through my son's middle school parent's listserv telling people about the study.

Unauthorized recruiting efforts were well-intentioned (they seemed aimed to ensure we had enough people sign up for the study) but also unhelpful. They skewed our study population toward oversampling the relatively affluent people who live in the county's northern half and under-sampling less affluent people who tend to live in the county's southern half. Since covid incidence in those days showed the same economic inequality as many other health conditions in America, this sampling biased our prevalence estimate downward relative to the truth. This problem is unavoidable, and we employed statistical reweighting techniques to help solve them.

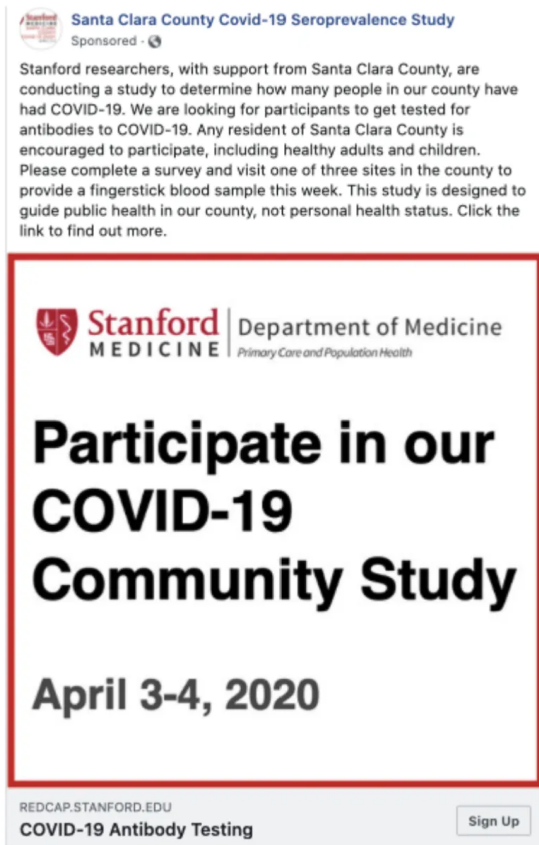


Figure 5

## HUMAN SUBJECTS REVIEW AND THE RETURN OF ANTIBODY RESULTS TO STUDY PARTICIPANTS

We knew that many were interested in joining our study because antibody tests were not available in March 2020. The US FDA had approved antibody test kits only for research use. So, the only way for a person to learn about their covid antibody status at the time was to participate in a study.

At the same time, it is unclear whether researchers should tell people who enroll in clinical studies about their own test results. Study participants may not understand the study result's nuances, such as the possibility of false positive or negative findings. Without understanding how to interpret results, participants may make decisions that harm themselves.

An example might be a study that tells a participant the result of a genetic test that doubles the lifetime risk of developing a rare cancer from 0.001% to 0.002%. Though this increase in risk is tiny, some study participants may choose to have the implicated organ removed. This will reduce the cancer risk by a mere 0.001% while leaving the patient to face the dangers of invasive surgery and living without that organ.

In our case, if the test produced an erroneous result, a patient could be harmed by coming away with a false perception about their antibody status and possible risk of infection.

### Script for Participants with a Positive Test

"A positive result could mean any of the following three things: (1) You got the virus that causes COVID-19, and your body developed antibodies that we are detecting with our test. Scientists do not yet know if this means you can get sick again, or whether you have immunity. (2) You may still have the coronavirus. Sometimes people that get infected develop antibodies even before the infection is over. (3) Although we double-checked your results, there remains a small chance that the tests were inaccurate, and you do not have antibodies to coronavirus."

Figure 6

### SCRIPT FOR PARTICIPANTS WITH A POSITIVE TEST

"A positive result could mean any of the following three things: (1) You got the virus that causes COVID-19, and your body developed antibodies that we are detecting with our test. Scientists do not yet know if this means you can get sick again, or whether you have immunity. (2) You may still have the coronavirus. Sometimes people that get infected develop antibodies even before the infection is over. (3) Although we double-checked your results, there remains a small chance that the tests were inaccurate, and you do not have antibodies to coronavirus."

The Stanford human subjects review board raised this very issue, but we insisted it would be unethical to withhold test results. People who took the antibody test would be intensely interested in their infection history, and we thought it paternalistic to deny them the result. We also accepted our responsibility to provide accurate guidance to those who tested positive based on what was known at the time.

In addition to the Stanford human subjects review board, we informed the Santa Clara County Department of Public Health that we would disclose antibody test results to study participants. They did not object and worked with us on a script to convey the findings. That script, reproduced below, emphasized the possibility of test errors and explained that scientists did not know with certainty at the time whether the presence of antibodies conferred any protection against subsequent covid infection.

Armed with this script, the study's principal investigator, Bendavid, spoke with everyone who tested positive. To this day, I think this was the right decision. Later research found that the antibodies measured by the test kit were a good marker for long-lasting protection against reinfection and against severe disease upon reinfection. In retrospect, the script turned out to be overly cautious – emphasized that we did not yet know whether covid-recovery provided immunity.

<u>Sample</u>	<u>Gold standard N</u>	<u>Kit agreement N</u>	<u>Sample type</u>	<u>Notes/provenance</u>
<i>Specificity (negative agreement)</i>				
1	371	368	Serum	Manufacturer's data based on pre-COVID-19 era sera

Figure 7

Unfortunately, some Stanford leaders later used this choice against us.

### CHECKING THE ACCURACY OF THE PREMIER BIOTECH TEST KITS

As the study approached the launch date, we worried about the test kit's accuracy. For a study such as ours, conducted in a setting of likely low infection rates, a high false positive rate would doom the study, rendering it unable to distinguish between zero prevalence and something higher.

To measure the kit's false positive rate, scientists find blood samples that they know for sure do not contain covid antibodies. The manufacturer did this by testing their kits with blood samples from 371 people drawn and stored before the SARS-CoV-2 virus began infecting humans. Any positive result in this sample was obviously a false positive. The manufacturer's documentation reported that only two of these 371 samples were positive. This meant the test had a false positive rate of 0.5%. If correct, this rate was perfect for our purposes.

After completing our study and releasing our scientific paper, we discovered a separate manufacturer document indicating a third sample might be positive. This increased the false positive rate to 0.8%. In any case, a false positive rate of 0.8% was still low enough to run a meaningful study.

Ultimately, we would find many assessments of the false positive rate of our test kit from independent researchers who had conducted studies similar to the one the manufacturer had run. At the time, though, we had the one check, so we started compiling a list of studies measuring our antibody test kit's crucial false positive rate. In the first row, we put the manufacturer's own estimate.

### JETBLUE'S FOUNDER AND A STANFORD COLLEAGUE OFFER HELP

Because covid antibody tests were so new, we wanted independent confirmation of the kit's accuracy from a source other than the manufacturer. Bendavid approached two colleagues with laboratories at Stanford for help – Scott Boyd, a senior professor in the pathology department, and Taia Wang, an assistant professor in the microbiology department.

In late March 2020, both Boyd and Wang's laboratories were developing their own covid antibody detection methodologies using a technology called an ELISA (or enzyme-linked immunosorbent assay). ELISAs often provide

more accurate results than the type of test kit we used, but this technique required a venous blood draw, not a pinprick. That would require study participants to come to the medical center – impossible given the shelter-in-place order. In any case, this technology was unavailable to us.

Understandably, Boyd and Wang were preoccupied with developing their ELISA tests, and Boyd told Bendavid he was too busy to help. In the early days of the pandemic, there was a scientific race to be the first to make an extremely accurate test. The stakes were high – developing an accurate, FDA-approved antibody test for clinical settings might be worth millions for its developer and Stanford.

Our team's purpose was less commercial and more research-focused. We wanted to measure how widespread the disease already was in the population to inform public health policy, not sell antibody testing. The cheap and accurate test kit we found was a godsend for our project, but direct competition for other teams at Stanford.

Wang kindly agreed to test our kits for accuracy, and as the April 4th date for blood sample collection approached, we anxiously awaited her findings.

She reported the results on April 3rd, the day before we were set to go into the field. Her laboratory checked the kit against stored blood from New York, collected before the pandemic. Her check found a 0% false positive rate with all 30 samples collected before the pandemic. We entered a second row into our table of studies measuring the false positive rate.

Around the same time in late March, Mr. David Neeleman, the founder of the airline company JetBlue, contacted Ioannidis and my research team. Because of the shelter-in-place order, Mr. Neeleman's airline was effectively closed for business, and he was interested in better understanding how long this shutdown would be and what this would mean for his employees. He was also worried about his parents living in New York City because they were older and at high risk of a poor outcome if infected.

Mr. Neeleman wanted to fund some bigger studies, including one focused on New York, the epicenter of the pandemic in the U.S. at that time. However, we had already committed to the study in California and could not change course.

Conversations with Mr. Neeleman helped me understand how vital the Santa Clara results would be for millions of people. He mentioned, for example, that Elon Musk was interested in funding a nationwide study. After we completed the study, he ultimately donated \$5,000 as a gift to Stanford to help fund costs, but we never ran a nationally representative sampling.

<u>Sample</u>	<u>Gold standard N</u>	<u>Kit agreement N</u>	<u>Sample type</u>	<u>Notes/provenance</u>
<i>Specificity (negative agreement)</i>				
1	371	368	Serum	Manufacturer's data based on pre-COVID-19 era sera
2	30	30	Serum	Pre-COVID-19 era healthy adults from New York

Figure 8

## BLOCKBUSTER RESULTS

On April 4th and 5th, Bendavid oversaw the field collection of over 3,300 finger-prick blood samples from Santa Clara residents. Local news media gave us positive coverage, and despite inevitable hiccups, all in all, things went rather smoothly.<sup>6</sup> Jim Tedrow, an Oklahoma scientist and commercial laboratory owner, flew out to Stanford and oversaw our laboratory that measured the participants' antibodies.

By the morning of the 6th, we knew we had bombshell results.

We found 50 people positive for covid antibodies out of 3,300 study participants. This doesn't sound like much, but once we corrected for our test kits' rate of false positives and false negatives and adjusted our sample to reflect Santa Clara County's demographics, we found that 2.8% of the county had already had covid.

This number was nearly 50 times more than county public health officials had estimated.

The implications for covid policy were threefold. First, the disease had already spread far more widely than people had realized. And this was despite three weeks of shelter-in-place orders, business closures, and school shutdowns. If lockdowns were supposed to eradicate the virus, our result made clear this was impossible. It also meant that the disease had arrived in California considerably before the official start date of the pandemic in late January 2020.

Second, the result implied that the infection fatality rate from covid infection was 0.2%, seventeen times lower than the case fatality rate announced earlier in the pandemic by the WHO. Of course, this did not consider nursing home residents, which we had excluded from the study to protect them from possible infection.

Finally, with only about 3% of people infected, there was still a long way to go before the pandemic would be over. The latter might happen only when enough people have acquired some form of immunity to the virus.

## THE STAR CHAMBER

Our field effort drew positive attention from the local papers, but the Stanford researchers developing their own commercial antibody tests were less happy with this attention.<sup>7</sup> California Governor Gavin Newsom, at the time, was focused on alleviating the lack of covid tests available to people early in the pandemic. He held a press conference coincident with our data collection in Santa Clara in which

he praised "Stanford" advances in antibody testing.<sup>8</sup> It was unclear whether he meant our study or Boyd's antibody test.

The Governor's remarks apparently drew the attention and ire of Boyd. He complained to Tom Montine, his department's chair, that our seroprevalence study would lead people in the community to confuse the test kit we used in the study with the ELISA antibody testing platform, called the Stanford ClinLab, that he was developing.

Montine and Boyd must have also contacted other leaders in Stanford Medical School because Bendavid and I started receiving increasingly alarming emails from Bob Harrington's (chair of Medicine) and Lloyd Minor's (Dean of the Medical School) offices expressing these same concerns. Bendavid responded with a detailed email to Montine explaining the evidence that our test kit was fit to run our study. In his email, Bendavid explained that the false positive rate for our test kit was low and that we had a distinct epidemiological and public health purpose for our study. We were not interested in creating a commercial test kit.

Bendavid's email to Stanford officials is below.

Ignoring Bendavid's explanation of our test kit's accuracy, Montine responded that disclosing the test kit results to study participants would "confuse" them. He wrote that we should not tell people their results "until [Boyd] determine[d] the concordance between [the Premier Biotech] kit and [Boyd's] Clinical Lab test."

In effect, the real concern of Montine and Boyd was that our epidemiological study posed a branding problem for Boyd's ELISA covid test that he was developing.

At the time of the email, Boyd was on the verge of releasing his test. Stung by criticism that his administration had not ramped up testing in California, Gov. Newsom bragged to the media about this new "Stanford" commercial antibody test.<sup>9-11</sup> Within a week, pharmaceutical industry trade journals were touting Boyd's new test, though it had not received FDA approval and could only be used on local patients.<sup>12</sup>

According to Boyd and Montine, the lateral flow test kit we used in our study – which they viewed without evidence as inferior – would confuse the public and potentially lead people not to trust Boyd's ELISA test. This opinion ignored the manufacturer's data and Wang's independent confirmation of our test kit's accuracy.

It is difficult to convey the extent of my shock at what happened next.

A bedrock idea of academia is that professors should be free to pursue their research without interference from



**From:** Eran Bendavid  
**Sent:** Monday, April 6, 2020 3:34 PM  
**To:** Thomas Montine  
**Cc:** Doug Owens ; Jayanta Bhattacharya ; Robert Harrington  
**Subject:** Validation data to date

Hi Thomas,

I wanted to present you with all the data we have to date on validating the test kit we used. The manufacturer's data is as follows:  
 371 pre-COVID samples were used as negative controls  
 Of those, 369 were kit negative on IgG and 368 on IgM. That's a specificity of 99.2-99.5%, with the lower bound on the lowest of those at 98.0%.

They used 75 positives that were clinically confirmed and known to have IgG as positive gold standard for IgG, and 85 that were known to be IgM positive as gold standard for IgM.  
 Of those, 75 were IgG positive and 78 were IgM positive. That's 91.8-100% sensitivity, with the lower bound on the lowest of those at 83.8%.

Once the kits arrived here, we engaged Taia Wang to validate the kits. (We also asked Scott Boyd and he declined on the grounds of being too busy.)  
 We gave her 100 kits to use for validation.  
 She ran 32 samples in her lab that were PCR+ and either IgM+ or IgG+ using her ELISA (32 IgM positive, 25 IgG positive).  
 Of those, 24 were IgM kit+ and 21 were IgG kit+. Overall, 25/32 positive using either band, 78% (60-91%).

She also ran 30 pre-COVID negative control and all 30 were negative.

We are considering several approaches to combining these data, and plan to present a range of possible sens and spec, and likely settle on a combination of the manufacturer's data with Taia's data, upweighting our local data because it's our and it's local. If we give Taia's data double the weight, we get a point estimate of specificity of 99.6% and sensitivity of 88.6%.

Since we are not providing clinical guidance for people, and the primary purpose of this study is for assessing the prevalence of covid-19 in our community, we can assess the uncertainty bounds around our prevalence point estimate and propagate the uncertainty around these sensitivity and specificity. We have worked out many of the approaches to doing that, and our sample size was big enough to give us enough power to make meaningful inferences using these data. Importantly, we are not going to provide any diagnostic results to individuals. The IRB approved us to call back those who tested positive, but the guidance we will give them makes it clear we do not know what the tests mean or what the implications are for individuals (we asked to report positives back to improve our survey response rate). The goal of this study is for population prevalence, not individuals guidance, and this was stated clearly in the consent.

I hope this is helpful in providing you a sense of the validation done to date and our plans for using this information.

We would also be very open to more validation of the existing kit. We have not many but some kits left over. We also saved all ~3200 samples from our study – in Taia's lab – but those are small-volume whole blood in EDTA samples that are now days-old, so I'm not sure how well they are holding up.

Best,  
 Eran

Figure 9

school administrators or other faculty members. Rather than dismiss Boyd and Montine's concerns as untoward interference in our research, the Stanford administrative hierarchy abetted that interference.

To ensure the protection of study participants, we had to gain approval from Stanford's human subjects review board before we were permitted to enroll anybody. We did this with a great sense of responsibility and worked closely with the board to protect our study participants and volunteers from harm. This approval is not just for the study itself but also for the specific study protocols and processes for conducting the research.

Surprisingly, Bendavid and I started receiving strident emails from medical school leaders ordering us to change our study protocol, even though they had been thoroughly vetted and approved by the Stanford human subject review committee and even though we had already collected our data. The human subjects review board is supposed to be independent of pressure and influence from medical school administrators.

I learned later that the medical school leadership had convened an ad hoc committee to supervise our research. The committee membership included many senior colleagues from Stanford, but none of us conducting the research.

I never learned the complete membership or who this group consulted, but I later learned that some members of that ad hoc committee had conflicts of interest – including some financial conflicts – that predisposed them against our study. I took to calling this group the “Star Chamber” because it did not follow any of Stanford's norms of research oversight, and it issued edicts that interfered with the conduct of our study. A star chamber is a secret court that the monarch used in medieval England to judge citizens accused of crimes against the crown, often including political opponents.

Among the people in the Stanford administrative hierarchy or faculty who I believe were members of or communicated with this informal but powerful group included Bonnie Maldando, a professor of pediatrics (and now interim chair of the Stanford Department of Medicine); Boyd, the professor of pathology we met earlier; Doug Owens, my immediate boss and chair of my department; and Robert Tibsherani, a professor of statistics. As chair of my division, Owens should have protested against the Star Chamber's intrusion on the academic freedom to pursue our research. Instead, he actively assisted them.

The Star Chamber dictated various orders, including demands to modify our study protocols that would have put us afoul of the human subjects committee, which approved our study. They also demanded to review and approve any

manuscripts we would write. Finally, they ordered me and Bendavid not to speak with the press about our research.

At first, we cooperated with the Star Chamber. This despite the fact that their orders directly violated our academic freedom and would cause us to fundamentally alter the protocols that the Stanford Human Subject Review Committee had already approved (more on that in just a bit). We obeyed until they warned us against publishing our research; at that point, we decided to ignore them.

## PUBLIC PEER REVIEW

We decided to release a preprint a week after completing data collection. We chose this process over a time-consuming peer-review journal. Researchers often post papers as preprints when they want immediate scientific feedback to improve the paper, release updates, and prepare for submission to a peer-reviewed journal. We also thought our results might help inform public health authorities to make better decisions about how to manage the pandemic.

Plus, other scientists were also interested in the accuracy of our antibody test kit and were testing it themselves. There was a race among companies to sell antibody test kits and check them for accuracy. After we released our preprint, scientists worldwide contacted us to offer the results of their own investigations.

Federal bureaucrats, reporters, and even members of the public read our preprint when we released it. Officials from the U.S. Food and Drug Administration (FDA) contacted us for help in evaluating the test kit we used. We sent them hundreds of our spare kits, and a few months later, the FDA granted the test kit emergency use authorization. In their testing, the FDA found the kit to be more accurate than we did. Among the 200 known-negative samples FDA tested, they did not find a single false positive. That same summer, Congress investigated the antibody testing industry and found many companies selling fraudulent test kits.

But the Premier test kit we used came through with flying colors.

We added these new studies testing the kit's accuracy to our list. With these results, the total number of samples we had to measure the false positive rate increased from just over 400 to over 3,400. The results confirmed the manufacturer's estimate of a ~99.5% specificity rate (0.5% false positive rate). We incorporated these new findings into an update of our paper, which we released two weeks after our first version.

On the release of the first version of our preprint, a debate ensued online about whether our result was credible – appropriate given how surprising our finding was. According to one statistic, our preprint generated more interest online than all but two other scientific preprints released that year.<sup>13</sup> Typically, a paper will receive maybe two or three peer reviews, and scientists may have months to respond. We received what felt like thousands of peer reviews within a day of releasing the paper. My email inbox filled beyond capacity.

One internet entrepreneur with a large online following published a blog post suggesting a remote possibility that

all of our positive test results were false positives, and hence, public health authorities could safely ignore our bombshell result.<sup>14</sup> This claim was wrong because it overlooked a statistical adjustment we used that corrected the false positive rate of our test kit. Unfortunately, the details of this calculation were buried in our paper's appendix. Nonetheless, the blog post persuaded many people that our estimates were somehow wrong.

Some serious statisticians also weighed in with negative reviews. Columbia University's Andrew Gelman posted a hyperbolic blog that we should apologize for releasing the study.<sup>15</sup> He incorrectly thought we had not accounted for the possibility of false positives. He later recanted that harsh criticism but wanted us to use an alternative method of characterizing the uncertainty around our estimates.

In parallel, the Star Chamber circulated our preprint to epidemiologists at Stanford. One Stanford professor offered an alternative method of measuring the uncertainty around our estimates. We implemented a modified version of this in the paper's second version.

Robert Tibshirani, a distinguished Stanford statistician, harshly criticized our paper. Unfortunately, he was not an expert in our study's subject matters. He never made clear the substance of his disagreement nor offered any substantive suggestions to improve our study. (Spoiler: After a rigorous peer review of our statistical methods, the *International Journal of Epidemiology*, a journal with the highest methodological standards, ultimately published our paper.)<sup>15</sup>

Some weeks earlier, the Star Chamber had told us to seek Tibshirani's statistical help, but Tibshirani demanded that we share the confidential health records of the patient volunteers with him. This would have been unethical as he was not a study coinvestigator. At some point after we had written our paper, Bendavid offered to include him as a paper coauthor so he could provide Tibshirani the data access he had initially sought. But by then, Tibshirani had become hostile to us and declined.

During a Stanford seminar about the Santa Clara study, Tibshirani demanded that we "retract" our preprint. Tibshirani's hyperbolic demand misunderstood the preprint process. Preprints are not represented as finished science; a preprint publication is supposed to lead to comments and suggestions from the scientific community. The proper response to valid criticism is to revise the paper and upload the updated version that addresses it.

Tibshirani later wrote an op-ed in the *New York Times* calling for the press to censor reporting of studies that did not meet his approval, using our Santa Clara study as a prime example of the sort of study that the press should bury.

One non-Stanford statistician quietly emailed us about a mistake we had made in the formula for the standard error calculation. This error did not alter our main finding but widened the confidence bound around it. Another UC Berkeley statistician proposed yet another novel method of calculating uncertainty around our estimates.<sup>16</sup> He took to Twitter and told the media that we made several statistical errors—errors that we had not made.

<u>Sample</u>	<u>Gold standard N</u>	<u>Kit agreement N</u>	<u>Sample type</u>	<u>Notes/provenance</u>
<i>Specificity (negative agreement)</i>				
1	371	368	Serum	Manufacturer's data based on pre-COVID-19 era sera
2	30	30	Serum	Pre-COVID-19 era healthy adults from New York
3	70	70	Serum	Adults tested negative for Lyme disease in a hospital in La Crosse Wisconsin pre-COVID-19 + adults admitted in early 2020 before first case in county
4	1,102	1,102	Serum	Adults admitted to hospitals pre-COVID-19
5	300	300	Plasma	Pre-COVID-19 plasma donors, in heparin
6	311	311	Fingerstick blood	Healthy adults, company staff of Biotest, the kit manufacturer
7	500	500	Venous whole blood	Children and adults admitted to hospital pre-COVID-19, on heparin
8	200	198	Serum	Pregnant women pre-COVID-19
9	99	99	Serum	Pre-COVID-19 serum of patients with <450 U/ml rheumatoid factor
10	31	29	Serum	Pre-COVID serum of patients with >600 U/ml rheumatoid factor
11	150	146	Serum	China CDC data collected from COVID-19-era hospital patients negative by PCR in Jiangsu province
12	108	105	Plasma	Pre-COVID-19 plasma donors
13	52	50	Serum	COVID-19-era PCR-negative specimens, some of which were positive for another respiratory virus
14	80	80	Serum	Pre-COVID samples tested at NCI and FDA (including 10 HIV+)
<b>Total</b>	<b>3,404</b>	<b>3,388</b>		

Figure 11

We had a long, fruitful email discussion with the Berkeley professor and tested his new method carefully. His proposed change was overly sensitive to small differences between studies in the false positive rate of clinical laboratory tests. Following his approach would invalidate nearly all clinical laboratory tests if applied at scale. To date, his criticism only exists on his blog, and his method has not been published in a peer-reviewed journal.

Conversely, Bendavid and Ioannidis, working with other colleagues, published a new, improved method for estimating uncertainty when seroprevalence is low, in the peer-reviewed *Journal of Applied Statistics*.<sup>17</sup> Again, the conclusions of our study would not change.

Before releasing the first version of the preprint, we asked several senior scientists from diverse institutions and members of national academies to review our work. These scientists, among the best in the world in statistics, epidemiology, and survey research, made helpful suggestions and made us feel more secure that our main findings were solid.

Once we published the preprint, responding to the voluminous correspondence we received became impossible. However, the public peer review gave us more information

about the test kit's accuracy and identified ways we might improve our statistical methods, so we decided to focus on updating the paper.

Despite some negative feedback we thought was unfair, this was precisely how open-source science was supposed to work. Our paper received more constructive comments than I had ever had for all 160+ peer-reviewed articles I had published in my career.

We updated the paper and sent the second version to the preprint site medRxiv within two weeks of the first version. The update implemented many statistical suggestions. Most importantly, we incorporated the dozen independent additional measures of the false positive rate of the test kit. Results from these laboratories confirmed the manufacturer's false positive rate of 0.5%. Though we had worried endlessly about that number, it turned out the manufacturer had gotten it right.

While the substantive results of the revision were the same, the second version was more methodologically sound.

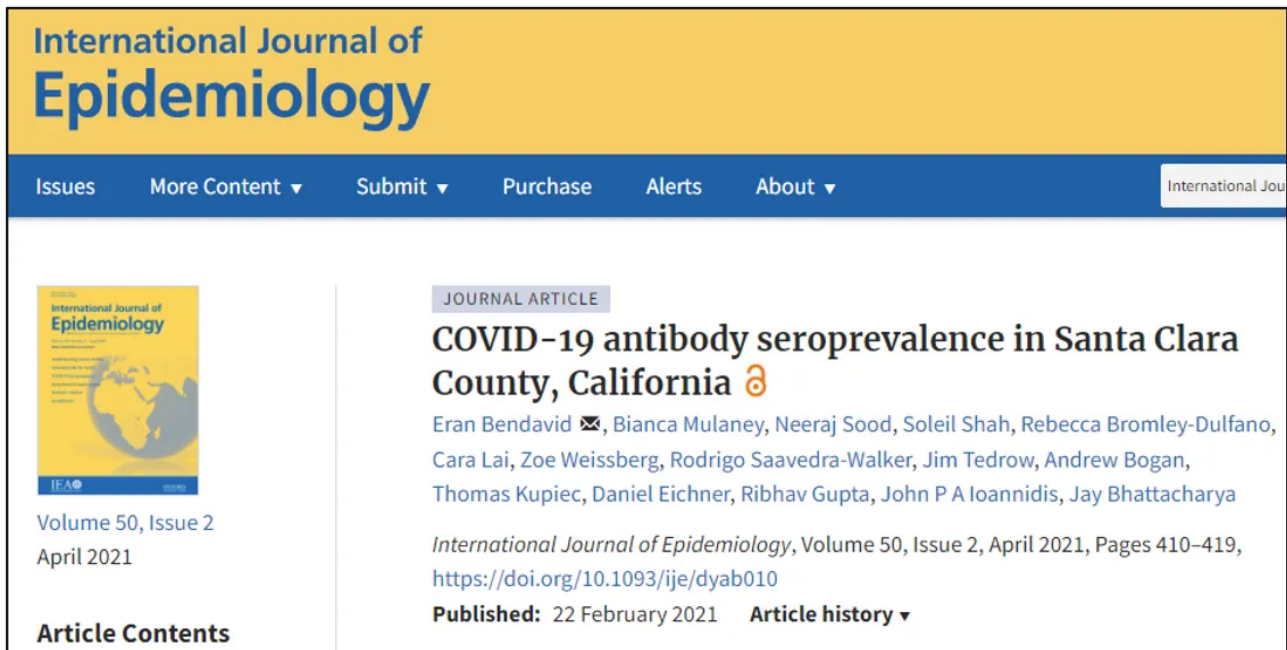


Figure 12

## STAR CHAMBER REDUX

We had not originally informed the Star Chamber that we were planning to release our paper as a preprint because Stanford faculty, like all other academics, do not need university permission to do science. At least, that is what I thought at the time. But when the Star Chamber found our paper published online, they were livid. The first sign was a call Bendavid received from Bonnie Maldonado. She informed him that Stanford's leadership was unhappy with us because we had released the paper.

We later learned that Maldonado, around this time, was seeking a grant from Facebook billionaire Mark Zuckerberg's foundation to conduct an antibody seroprevalence study, just like the one we conducted. This study was also to take Santa Clara County. It is a grant that she ultimately won for \$13 million, with a promise to produce a seroprevalence estimate by December 2020.<sup>18</sup> The study's design was eventually published in November 2021 in the *Annals of Epidemiology*, although they had no results.<sup>19</sup> Their team took five months to recruit the same number of individuals as we recruited in 2 days.

Bendavid later received an email from Robert Harrington, the Chair of the Department of Medicine, excoriating him for allegedly being more interested in generating a scientific scoop than in the well-being of the patients in our study. Bendavid followed a script approved by the Stanford Human Subjects review committee, making this charge incredibly unfair. Dr. Bendavid is a board-certified infectious disease doctor with decades of patient care experience, and he is good at counseling patients.

The press gave overwhelming attention to our results. I gave TV interviews with CNN, NPR, and a host of other outlets. I found the CNN and NPR interviews odd.<sup>20</sup> The reporters were very concerned that our results would some-

how convince people to hold covid parties (analogous to chicken pox parties of yore) aimed at intentionally spreading the virus. I emphasized that covid parties were a bad idea as the virus was still dangerous for high-risk people like the elderly. Bendavid and Neeraj Sood gave an interview on Good Morning America to a national audience.

Harrington emailed me and Bendavid demanding that we stop giving public interviews and stop talking with the press. This order was entirely inconsistent with Stanford's nominal commitment to the free speech rights of professors. Harrington (the Chair of the Department of Medicine) later apologized to me for overstepping boundaries with this order, though he has yet to apologize to Bendavid. He told me privately that his order came from someone above him in the university hierarchy. As chair of medicine, that narrows the set to the very highest officials at Stanford.

Ignoring the fact that our test kits' accuracy had been confirmed by a Stanford lab and other independent researchers, the Star Chamber ordered us to alter our study protocol. They told Bendavid we might falsely inform patients of their antibody status since the false positive rate was so high for our test kits. We had already worked through this issue with the Stanford human subjects board, and the script we devised to inform the positive patients meticulously emphasized the possibility of a false positive. The Star Chamber ordered us to bring back all the patients who had tested positive for retesting using Boyd's antibody test—a test we had not verified nor sought approval for use in our study.

In retrospect, we should have told the Star Chamber to pound sand. They had no evidence that Boyd's experimental ELISA test was more accurate than our test kit and they provided us with no data. On the contrary, the U.S. Food and Drug Administration eventually granted an emergency use authorization to our test kit, while Boyd's experimen-



tal ELISA platform never got such authorization. We agreed to the Star Chamber's demand because we wanted to take every conceivable measure to protect our patient volunteers.

Bendavid informed the Stanford human subjects committee about the change in protocol, which they approved.

We invited the 50 antibody-positive study volunteers back, and 47 agreed to be retested. Since there were about 3k people in the study and the false positive rate was about 0.5%, we expected about 15 ( $= 3,000 * 0.5\%$ ) false positives or about 31 or 32 who were positive on Boyd's ELISA. And that assumed that Boyd's test had a 0% false positive rate, which was certainly not the case.

Eventually, Boyd emailed the Star Chamber and us with his results. He had found that of those 47 positive volunteers, 31 were positive on his ELISA – precisely the number we were expecting. To estimate the false positive rate, we need to start with a sample that is known negative and then estimate what fraction tested positive. Boyd had done the opposite. He had started with a known positive sample (those who tested positive on our test) but interpreted our sample as being all negative. He incorrectly divided by the number of antibody-positive participants rather than the total number of participants!

In fact, his retest provided yet another confirmation that we knew the correct false positive rate of the test we had used.

The news from Boyd could not have come at a better time. We were still receiving grief from online critics convinced that it was possible – even likely – that all our positives were false positives. Upon submitting the paper for publication to a formal scientific journal, we wanted to tell the journal editors that an entirely independent assessment showed that our test kit was accurate. An independent confirmation of our result by a critic (even better, a hostile critic) would have convinced many people.

We wrote to the Star Chamber and Boyd that we would publicize this confirmatory result undertaken at the behest of the Star Chamber. Boyd immediately refused us permission to use his independent confirmation in our published paper. We were shocked by this breach of collegiality. The Star Chamber told us we must abide by Boyd's wishes and suppress these confirmatory data. We obeyed because we were intimidated by a direct order from Stanford Medical School leadership.

## BUZZFEED NEWS, AN EMAIL, AND A PHANTOM CONFLICT OF INTEREST

In late April, while we were updating the paper and dealing with the Star Chamber, I received an email from a BuzzFeed News journalist named Stephanie Lee. She had obtained an email that my wife, Dr. Catherine Su, had sent to my younger son's middle school email list without my knowledge. My wife is a practicing radiation oncologist in private practice. She sent the email to alert her friends about the Santa Clara study's existence to gather more participants.

In her email, my wife wrote that a positive antibody result implied that a person was immune to reinfection



Figure 13

and could probably resume work and normal life. My wife is not a clinical researcher and did not know that these prophetic thoughts were forbidden at the time. Nor did she understand that alleging infection provided immunity was deemed misinformation by the press. Her email was a small ripple in a vast flood of unauthorized online talk, pointing people to our study.

Lee's story in BuzzFeed News created a "scandal" around my wife, putting her in the national news for a day. The story created a false sense among the public that the Santa Clara study's methods were flawed without explaining why. For me, the story was the source of tremendous anxiety. I worried that I had exposed my family and loved ones to vicious and unfair attacks for the crime of being associated with my research. I felt helpless.

Lee went on to write at least two more hit pieces about the Santa Clara study or our study team. One tried to make a scandal out of Ioannidis' attempt to organize a team of prominent scientists with relevant public health experience, all with faculty positions at prominent universities and a broad range of views regarding the wisdom of lockdowns. Ioannidis's crime was failing to secure a meeting with President Trump for this diverse group of scientists.

For Lee and BuzzFeed News, non-approved scientists trying to inform policy was scandalous.

Ioannidis' view in March 2020 about the lockdown policy was equivocal – he had expressed skepticism and the need for better data.<sup>21</sup> He had provided consent for a short lockdown in a widely viewed video, but this video was censored by YouTube after the uproar from the BuzzFeed News hit pieces.

In one email, Ioannidis ironically joked about how his efforts had hit a wall and were ignored by the Trump White House. Taking this joke out of context, Lee falsely presented the endeavor as proof of a lobbying effort to influence the White House. The hit piece was accompanied by a photo of Ioannidis next to a picture of Donald Trump. Not surprisingly, at a time when emotions ran high and the press was turning against Trump's pandemic policies, the

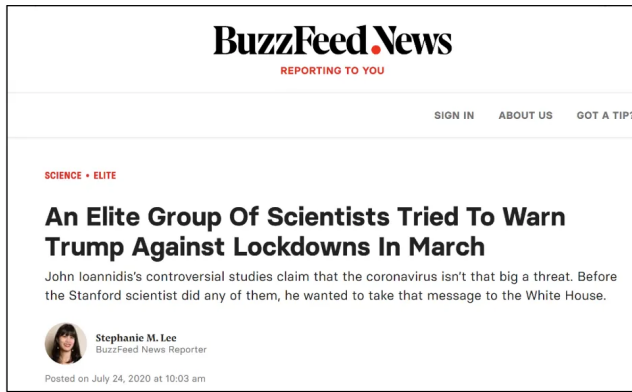


Figure 14

Buzzfeed article put the lives of Ioannidis and his family at risk, leading to death threats and media hoaxes.

The other Lee piece was a long article littered with false allegations about the Santa Clara study. To begin, she aired false doubts about the accuracy of our antibody test kit, based on statements from Boyd, without disclosing his conflicts of interests, as he had a competing test.

Second, she created a phantom conflict of interest regarding a \$5,000 donation by Jet Blue founder David Neeleman. Neeleman donated the money to a Stanford philanthropic gift account, not to any of the study authors. The donation constituted less than 5% of the total study expenses, and none of us were compensated for our work on the study. To make matters worse, Lee's article ran a photo of Neeleman next to Ioannidis, implying he was involved in the donation, even though he did not know about it until after Neeleman had sent the money to Stanford.

But Ioannidis is an internationally well-known scientist, and slander resulted in a bigger buzz for the aptly named BuzzFeed.

I was still upset with Lee about her targeting my wife for a hit piece, so I did not respond to her long and tendentious email regarding her new "investigation." But Ioannidis and Bendavid spent hours on the phone with Lee, explaining the facts I have written here. Neeleman's donation did not reflect a conflict of interest, and the test kit we used was sufficiently accurate for the purposes of our study. Ioannidis sent meticulously written replies to questions that Stephanie Lee sent. Little to nothing of this material made it into Lee's inaccurate buzzy hit piece.

In recognition of her work and with applause for her methods, Stephanie Lee was even given an award by fellow journalists in the summer of 2022.

## THE INQUISITION

Lee's yellow journalism induced Stanford leaders to take more action against us, starting a process that they labeled a "fact-finding" review. Resorting to gallows humor, we called the process an "inquisition."

Almost immediately, we got a bit of good news. While Lee had hyperbolically written of unspecified "whistleblowers" casting dirt on our study, Stanford leaders informed us

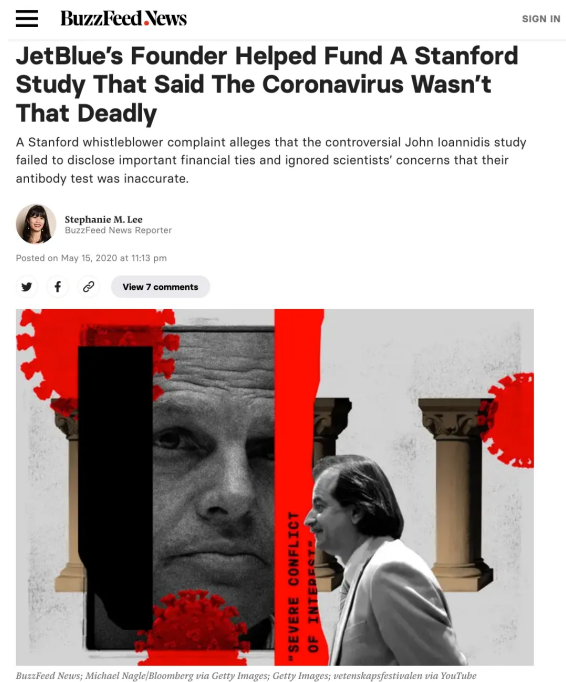


Figure 15

that no whistleblowers existed. We were also told that the "fact-finding" review was being done to better inform how Stanford should manage studies like ours and to see if mistakes were made from which Stanford and we could learn.

Nevertheless, the inquisition took months and was emotionally taxing. For the first time in my life, I had trouble sleeping and often forgot to eat. I lost nearly 30 pounds over the course of a couple of months due to the anxiety. Bendavid and I hired lawyers to advise us in our dealings with Stanford.

The inquisition concluded that we had conducted the study honorably and made no finding of fault. The letter summarizing the inquisition emphasized that the university still had full confidence in us as faculty members in good standing and that we could continue our teaching and research activities with full university support. It included some criticisms of the study and Stanford's oversight that we thought missed the mark. But since the letter concluded the process favorably to us, we made little fuss over it.

Of course, we were eager to make this finding public, but the inquisitors insisted that we request permission before releasing it. Furthermore, the university told us that it would make no public statement exonerating us, even though the inquisition had done precisely that.

We protested that the university administrators defend our reputation – to no avail. This "fact-finding" started with a university press release, putting a cloud over our heads, and ended in silence once they absolved us of wrongdoing.

Many months later, some media still mentioned that we were being "investigated" as if we were suspected criminals. I even heard the slander in a major anti-lockdown court case in Manitoba, where I testified as an expert witness. The distortions of what happened were sometimes

more laughable than lamentable. For example, one utterly confused paper published in the journal *Cultural Studies of Science Education* even mistook Ioannidis to be the billionaire founder of JetBlue: “The study was in fact partially funded by John Ioannidis, the founder of JetBlue airlines, ... a corporate businessman ... Ioannidis had a significant conflict of interest in the study, wherein his primary interest was in the profitability of his business...”<sup>22</sup>

To not rock the boat any more than we already had, and because I found myself in other covid policy related fights in the midst of which this story would have been a distraction, I've kept this part of the story out of the public eye until now.

## EPILOGUE

Despite all the controversy, acrimony, and university interference in the Santa Clara study, the scientific process ground its way slowly and ultimately won out. We sent the revised preprint paper to a scientific journal for peer review and, behind closed doors, heard many of the same criticisms of the paper we'd heard on Twitter and elsewhere. Because we had already heard all of these arguments before, we could counter them to the satisfaction of the journal editors effectively, and the *International Journal of Epidemiology* – one of the top epidemiology journals in the world – finally published the paper in early January 2021.

With nearly 800 citations at the time of this writing, the study is among the four most highly cited serology studies worldwide and the most highly-cited one done in the USA.

Publication, peer review, and citations are not the ultimate tests of the validity of novel scientific findings. Unless independent research teams can replicate a scientific result, it stands little chance of being widely accepted by the scientific community. By that test, the Santa Clara study has passed with flying colors.

A week after we collected community samples for the Santa Clara study, a study team that included Bendavid, Sood, and me (all authors on the Santa Clara study) repeated the exercise in Los Angeles County, California. The LA county study differed in some ways from the Santa Clara study. Rather than Facebook sampling, we hired a polling firm with an existing representative community sample of LA County residents. The other difference is that we ran the study under the auspices of the University of Southern California rather than Stanford University.

The study result was nearly identical to Santa Clara. By April 11, 2020, we found that over four in a hundred L.A. County residents showed evidence of infection and recovery from covid. This was nearly forty times more infections than previously identified cases. The infection fatality rate estimate was also very close to our Santa Clara estimate, about 0.2%.

Our experience in Santa Clara informed how we wrote up the results. We sent the L.A. County paper to the *Journal of the American Medical Association*, which published the piece in May 2020.<sup>23</sup> It has received nearly 500 citations from other researchers to date.



Figure 16

The biggest contrast between the L.A. County and Santa Clara County studies is how the University of Southern California (USC) dealt with us. USC awarded Sood, the study's lead author, an enormous grant to conduct further studies on covid, which he used to great effect during the pandemic. The LA County Department of Public Health announced the study results with Sood present at a public press event.

Unlike at Stanford, there was no Star Chamber and no inquisition.

I also helped investigators associated with Major League Baseball (MLB) to study the spread of the virus among MLB non-athlete employees. This was the first seroprevalence study of nationwide scope, and we found a prevalence rate of 7 in 1,000 in this population. That this number was lower than the community-dwelling populations in Santa Clara and LA counties was unsurprising. The MLB had suspended spring training with the lockdowns, and most were working from home. The MLB study also showed a 0% infection fatality rate since no MLB employee had died up to then.

In May 2020, the MLB held a press conference where I publicly announced the study results. Having been burned by the vicious attacks from the Santa Clara study, and because the MLB study effectively duplicated the LA County and Santa Clara County studies, I decided not to release a formal paper.

But it wasn't just our team that replicated the results of our Santa Clara study. Teams worldwide ran thousands of seroprevalence studies using methods very similar to the ones we developed, though often with different antibody test kits. In December 2020, Ioannidis published a meta-analysis summarizing the results of dozens of such studies in the *Bulletin of the World Health Organization*. That analysis found an infection fatality rate from covid infection between 2 and 3 in 1,000, almost precisely the result we saw in the early Santa Clara study.

Our result of an 0.2% infection fatality rate in the community-dwelling population shocked many people, includ-



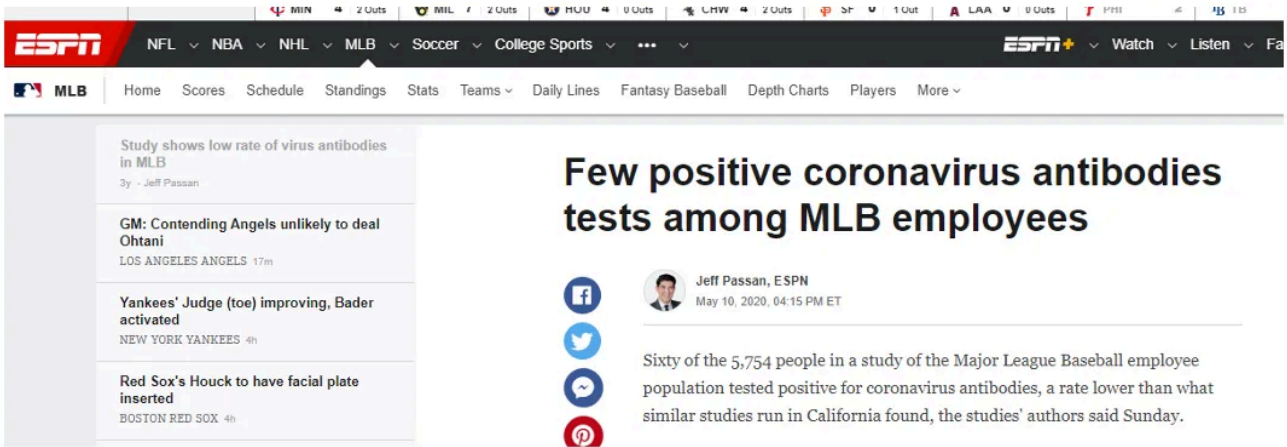


Figure 17

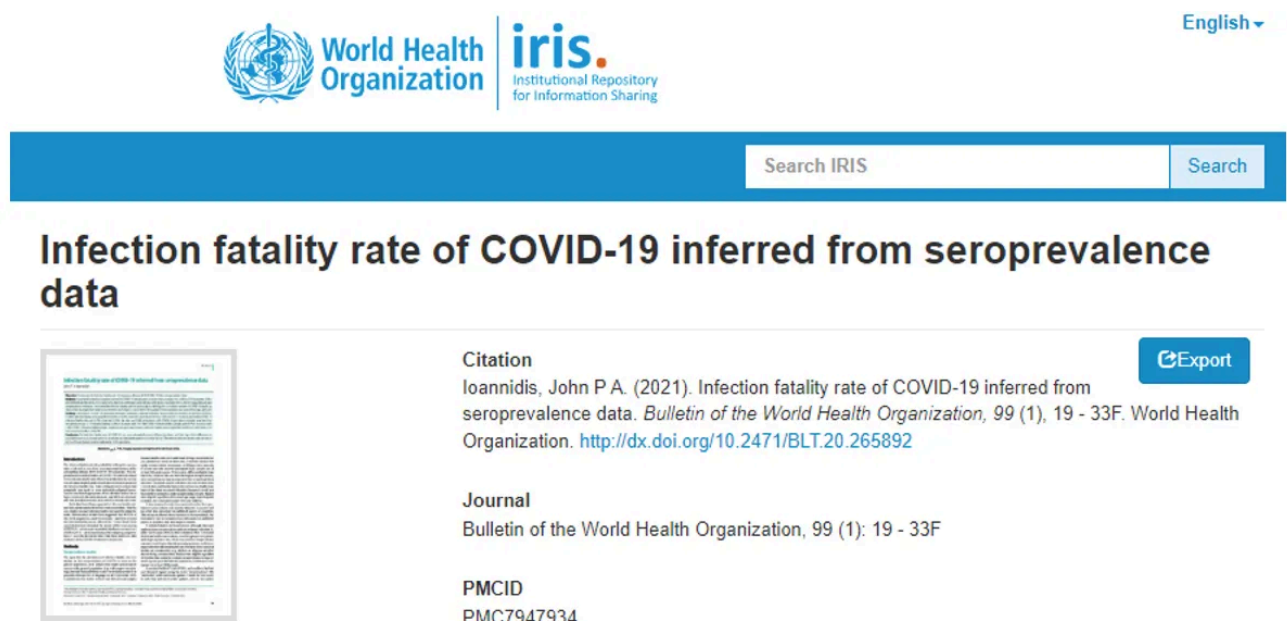


Figure 18

ing academics, journalists, politicians, and science bureaucrats. It seemed to us that many people were looking for reasons to disbelieve our results. Initial skepticism is itself not surprising nor inappropriate. But continued skepticism after replication by independent groups is not scientific.

The overwrought reaction to the Santa Clara study by some Stanford leaders, the press, and social media is, in some way, also easy to understand. Everyone has human emotions and personal motivations. By April 2020, public health authorities and governments had conditioned the public to think of covid as the second coming of the 1918 influenza pandemic. Many people – even professors at Stanford and public health authorities – were genuinely fearful that the virus might even kill them. At the time, public leaders had implemented lockdowns, forcing incredible sacrifices by ordinary people.

Embracing our results would have forced some leaders to admit error or reverse their policies, which would have meant a loss of reputation.

University administrators' actions to undermine the study and cast a cloud over the authors bear particular condemnation. Stanford University has a nominal commitment to academic freedom – the university's motto is "Let the Winds of Freedom Blow." By interfering with the Santa Clara study, specific Stanford administrators violated that commitment. While it is impossible for me to know the motives, some Stanford leaders seemingly prioritized commercial interests – developing and selling an antibody testing platform – over a commitment to academic freedom and the scientific endeavor.

Other Stanford leaders acted to protect a perceived threat to Stanford's brand. Perhaps they genuinely – but incorrectly – believed our study to be incorrect. Maybe they thought we were undermining public health messages,



**To:** Hahn, Stephen (b) (6)  
**Cc:** Birx, Deborah L. EOP/NSC (b) (6), Redfield, Robert R (CDC) (b) (6)>, Kadlec, Robert P (OS) (b) (6), Caliguiri, Laura (b) (6)>, Lenihan, Keagan <(b) (6) Shuren, Jeff (b) (6)>, Lane, Henry C (NIH) (b) (6)  
**Subject:** RE: medRxiv: COVID-19 Antibody Seroprevalence in Santa Clara County, California

Steve:

Thanks for the quick reply. The data from the Stanford group (b) (5)

(b) (5)  
(b) (5)  
(b) (5)

Thanks,

Tony

Anthony S. Fauci, MD

Director

National Institute of Allergy and Infectious Diseases

Figure 19

which emphasized an erroneously high mortality risk from infection. Choosing Stanford's brand over Stanford's research mission damaged the university far more than our research, which is now published in peer-reviewed journals.

In the end, Stanford's leadership undermined public and scientific confidence in the results of the Santa Clara study. Given this history, members of the public could be forgiven if they wonder whether any Stanford research can be trusted. Is work published by Stanford faculty actually what the researchers think, or was there unreported and inappropriate interference by Stanford's academic bureaucrats in service of interests other than the truth?

High levels of government, including scientific bureaucrats like National Institute of Allergy and Infectious Disease head Tony Fauci, discussed our research.

These officials were the architects of the lockdown policies followed by many nations. We know this because of emails revealed by a Freedom of Information Act request.

In these emails, FDA Commissioner Stephen Hahn and Tony Fauci were apparently trying to understand what our results meant. The emails are almost entirely redacted, but it's clear policy officials in the nation's Capital were interested.

It is impossible for me not to speculate about what might have happened had our research been met with a more typical scientific spirit. For anyone with an open mind, the study's results implied that the lockdown-focused strategy of March 2020 had failed to suppress the spread of the disease. It also suggested that the disease had very likely been in circulation in California for longer than the late January 2020 official start date of the American pandemic. Finally, it suggested that fear-mongering about the fatality rate of the virus was irresponsible.

Of course, no single study— let alone the very first sero-survey in one county in the U.S. — should have been the sole guide to policy. However, as multiple independent sources replicated these results, the conclusion should have become common knowledge among public health officials: lockdowns had failed and likely harmed the well-being of children, the poor, and working-class people.

Open-minded acceptance of our study results might have helped the world avoid harm from both covid and covid mitigation policies. What damage might have been avoided had scientific bureaucrats actually followed the science in April of 2020?

This article was co-published by JoSPI and [The DisInformation Chronicle](#).

Submitted: September 12, 2023 EDT, Accepted: September 12, 2023 EDT

**From:** Hahn, Stephen (b) (6)>  
**Sent:** Saturday, April 18, 2020 10:31 AM  
**To:** Fauci, Anthony (NIH/NIAID) [E] (b) (6)>  
**Cc:** Birx, Deborah L. EOP/NSC (b) (6)>; Redfield, Robert R. (CDC/OD) (b) (6)>; Kadlec, Robert (OS/ASPR/IO) (b) (6)>; Caliguiri, Laura (FDA/OC) (b) (6)>; Lenihan, Keagan (FDA/OC) <(b) (6)>; Shuren, Jeff (FDA/CDRH) (b) (6)  
**Subject:** Re: medRxiv: COVID-19 Antibody Seroprevalence in Santa Clara County, California

Tony,

(b) (5)

(b) (5).

Steve

Figure 20



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

## REFERENCES

1. Lovelace B Jr. WHO says coronavirus death rate is 3.4% globally, higher than previously thought. *CNBC Health and Science*. 2020.
2. Centers for Disease Control and Prevention (CDC). Past Seasons Estimated Influenza Disease Burden. 2022.
3. Wong JY, Kelly H, Ip DKM, Wu JT, Leung GM, Cowling BJ. Case fatality risk of influenza A (H1N1pdm09): a systematic review. *Epidemiology*. 2013;24(6):830-841. doi:10.1097/ede.0b013e3182a67448
4. Flaxman S, Mishra S, Gandy A, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*. 2020;584(7820):257-261. doi:10.1038/s41586-020-2405-7
5. Ioannidis JPA. A fiasco in the making? As the coronavirus pandemic takes hold, we are making decisions without reliable data. *STAT News*. 2020.
6. Krieger LM. Coronavirus: New Stanford research reveals if you've been exposed. *The Mercury News*. 2020.
7. Selig K. Stanford researchers test 3,200 people for COVID-19 antibodies. *The Stanford Daily*. 2020.
8. Flores H. Gov. Newsom announces increased COVID-19 testing capacity in California. *KCRA 3*. 2020.
9. Johnson L. Newsom says he 'owns' California's coronavirus testing problems. *San Francisco Chronicle*. 2020.
10. Gutierrez M. Gov. Gavin Newsom said he 'owns' coronavirus testing lapses, announces task force. *Los Angeles Times*. 2020.
11. Scientists Focus On Antibody Tests That Could Be Key In Getting People Back To Work. *CBS News Bay Area*. 2020.
12. Terry M. Stanford Serology Test: Stanford Joins List of Organizations Offering COVID-19 Antibody Testing. *BioSpace*. 2020.
13. Bendavid E, Mulaney B, Sood N, et al. COVID-19 Antibody Seroprevalence in Santa Clara County, California. *International Journal of Epidemiology*. 2021;50(2):410-419. doi:10.1093/ije/dyab010
14. Srinivasan B. Peer Review of "COVID-19 Antibody Seroprevalence in Santa Clara County, California." *Medium*. 2020.
15. Bajak A. A Study Said Covid Wasn't That Deadly. The Right Seized It. *The New York Times*. 2020.
16. Fithian W. Statistical comment on the revision of Bendavid et al. *Berkeley Statistics*. 2020.
17. Cai B, Ioannidis JPA, Bendavid E, Tian L. Exact inference for disease prevalence based on a test with unknown specificity and sensitivity. *Journal of Applied Statistics*. 2022;50(11-12):2599-2623. doi:10.1080/02664763.2021.2019687
18. Stanford teams with UCSF, Chan Zuckerberg Biohub to study prevalence of COVID-19 in San Francisco Bay Area. *Stanford News*. 2020.
19. Lindan CP, Desai M, Boothroyd D, et al. Design of a population-based longitudinal cohort study of SARS-CoV-2 incidence and prevalence among adults in the San Francisco Bay Area. *Annals of Epidemiology*. 2022;67:81-100. doi:10.1016/j.annepidem.2021.11.001
20. Bhattacharya J. How Antibody Tests Can Inform Public Policies To Mitigate Coronavirus Pandemic. *NPR Health Care*. 2020.
21. Hartmann E. The Most Important Meeting in the History of the World That Never Happened. *Brownstone*. 2023.
22. Nasr N. Overcoming the discourse of science mistrust: how science education can be used to develop competent consumers and communicators of science information. *Cult Stud Sci Educ*. 2021;16(2):345-356. doi:10.1007/s11422-021-10064-6
23. Sood N, Simon P, Ebner P, et al. Seroprevalence of SARS-CoV-2-Specific Antibodies Among Adults in Los Angeles County, California, on April 10-11, 2020. *JAMA*. 2020;323(23):2425-2427. doi:10.1001/jama.2020.8279

## SUPPLEMENTARY MATERIALS

### Figure 10

Download: <https://www.jospi.org/article/88046-dr-jay-bhattacharya-reveals-stanford-university-s-attempts-to-derail-covid-studies/attachment/181152.png>

---