David Benkeser who is an assistant professor at the department of biostatistics and bioinformatics at Emory University. Dr. Benkeser got his PhD in biostatistics from University of Washington and had his post-doctoral fellowship from University of California at Berkeley. Dr. Benkeser is an expert in methods for machine learning and non-parametric statistical inference. He has made important contributions to integrate machine learning methods to draw causal inferences with observational data. He also has interesting work on preventative vaccines and HIV prevention, which he’s going to share with us today. Welcome David, the floor is yours. Thanks, yeah, it’s a great pleasure to be here today. Well, here today, but with you guys today giving this talk. So I did see that I think Tony Fauci spoke at Yale yesterday, so it was very nice of you Fan to book Tony Fauci as my opening act and I’ll try not to disappoint him with my followup. So the talk I’m giving today is a very high-level talk.
So the title is statistics and COVID-19 vaccine development, but it’s really a talk mostly about COVID-19 vaccine development. There’s not math until maybe slide like 29 out of 30. So really these are sort of just the high-level issues that have come up as I’ve worked with companies and government organizations on COVID-19 vaccine development. So I think there’s a lot of really interesting stuff here and really, really glad to share it with you today. So if you want to kind of slide along with the slides they’re available on GitHub so there’s a link at the bottom there, and you can click on that and that’ll pull up the HTML slide back, and I have sort of references hyperlinked in there. So that’s an easy way to access the references there as well. Okay so I’m going to start just kind of talking about the biology a little bit of SARS-CoV-2, about how we can think about developing vaccines that will prevent an infection and COVID-19 disease. And so this is a nice little graphic that I ripped off from The Washington Post, who’s very much better at making these cutesy little graphics.
than I am using PowerPoint or something.

So let’s kind of walk through this.

And the goal here is to try to understand,

you know, how SARS-CoV-2 is infecting your cells,

how it’s replicating,

and then to understand what the mechanisms

that immunological mechanisms of the vaccine are

that can potentially block that infection

and prevent clinical disease.

So we’ll just go quickly through this

and this is sort of the story for most viruses, right?

Is that viruses are really just genetic material

in this case RNA that’s wrapped up in the glycoprotein.

So it’s genetic material wrapped up in a protein.

And so for SARS-CoV-2 you may have heard of a couple

of these proteins in particular, the spike protein will play

a large role when we talk about a vaccine development

and why is this spike protein so important?

Well, that’s the guy that sort of latches

onto your cell and it does that through this ACE2 pathway

and it grabs onto your cell and insert itself inside

you cell and once it’s inside

it releases its genetic material, right?

It releases its RNA and kind of tricks

your cell into replicating the virus, right?
So that your cell is producing new copies of this virus, they’re pieced together out of proteins that are released into your bloodstream to go infect more cells and more people.

Okay so that’s sort of the infection process and where along the lines do you know vaccines sort of halt this? So I’ll walk through a few different of the major vaccine constructs that are being used for SARS-CoV-2 vaccines, and the details aren’t super important here, but I do think it’s sort of helpful to have a high level overview in comparison, right? Because there’s so many vaccine products being developed, at least having some point of biological comparison of how they’re working is useful. So to walk through these slides, all of these slides are basically going to be the same on the right hand side of the slide and how they’re gonna differ is what goes into the vaccine on the left-hand side. So let’s actually start on the right-hand side, right? And talk a little bit about immunology, right? And how your body tries to fight off infection. And we have a couple of different mechanisms
So there’s a kind of T-cell responses, cytotoxic T-cells. Those are T-cells that recognize cells in your body that have been infected with a pathogen and destroy those cells, right? Because the cells are producing copies of the virus, releasing in the bloodstream. So if we’re able to destroy infected cells, we can potentially stop infection, prevent disease.

And that’s sort of what’s on the bottom here. And what those antibodies do is they basically grab onto these surface proteins, right? So remember we talked about the spike protein, and what antibodies do is basically just bind onto that and sit there and so neutralizing antibodies. So there’s two classes of antibodies that are kind of relevant for vaccines.

So neutralizing antibodies really, you’re just doing that. They’re gonna sit on all of those spike proteins and because they’re sitting there now the virus can’t grab...
onto your cells to infect them. There's also binding antibodies, which are somewhat considered to be less important in this context, but what those guys do is bind onto those surface proteins, they don't neutralize the virus itself, but they send out chemical signals to other cells in your body that say, hey, here's a virus. Please come eat it for me. So those are the sort of antibody classes response that you can have. So there's these two sort of immune mechanisms that we have to neutralize infections by viruses. How do they learn to neutralize them? Well, there's this sort of middleman. So we're moving just to this middle panel here with these APC cells, so these antigen presenting cells, right? Those are the guys that what they're doing is basically digesting little bits of the virus in this case of the surface protein, right? And they're teaching or training your immune system to recognize that pathogen, right? So they're the ones that go and talk to the T cells, talk to the B cells and say, here's that how this virus looks,
please go produce some antibodies or please recognize cells
that have been infected with this
and neutralize them for me.
So really again, the whole right side of this plot
is about your immune system.
This is the way your immune system fights off infection.
And what’s different between this slide and the next few slides is basically how we present
pieces of the pathogen pieces
of the virus to these APCs, right?
So how do we get these APCs, the material that they need
for you to mount an immune response against SARS-CoV-2?
And so the first class of vaccines
I’ll describe are nucleic acid vaccines.
And so I’m talking about first
because they’re sort of the first wave of vaccines
that are in phase three trials in the US.
So Moderna and Pfizer, who are probably the most advanced
candidates for US licensure are both mRNA vaccines.
And so how are those vaccines made?
Well, we take a little bit of messenger RNA, a little bit of viral genetic material,
and wrap that in a lipid shell, right?
That’s the construct of the vaccine.
And when you’re injected that lipid shell
latches onto your cell, right?
Delivers that mRNA into your cell, just like a natural infection, right?
Remember the SARS-CoV-2 grabbed onto your cell
and inserted itself and then made copies of itself.
So what is the mRNA doing once it’s in your cell,
it’s actually just making copies of the spike protein itself, right?
So you’re manufacturing this protein within your own cells
that are then released for these APCs to detect.
So this is how we’re getting these APCs,
spike protein with an mRNA vaccine.
We’re basically using your cells as a warehouse
to produce the antigen of the vaccine
and so this is a really cool idea and a new idea, right?
So, am mRNA or DNA vaccine has never been licensed before
and that’s not to say that we tried many times and failed.
It’s just to say that this is a very new technology,
it’s sort of interesting that it’s kind of come
to the forefront in this context.
So why do we like mRNA vaccines?
Well, they’re very fast to manufacturer.
We’ll talk about some of the other vaccine constructs
where we’re making this spike protein in a lab, and that is a long and arduous. It needs to be very careful process and when we’re thinking about scaling up vaccine manufacturing, mRNA vaccines are very appealing in that sense, you can manufacture them very quickly at scale. They don’t require a cold chain and so that’s another great advantage these vaccines enjoy in terms of thinking about vaccine deployment, particularly in developing world settings. But again, this is a brand new technology. We don’t have any safety data. We don’t have any efficacy data. So it’s sort of an open question in the field as to how well these things are gonna work. So moving to sort of more classical, constructive vaccines and viral vector vaccines. So again, the right side of this picture is exactly the same. The story is how do we get an APC the right antigen? How do we show an APC a little bit of the spike protein? So a viral vector vaccine, right? Is going to take a different virus and splice
a little bit of SARS-CoV-2 into that virus, okay.
So for example, AstraZeneca, that’s the Oxford that you may have heard of, they take a chimpanzee adenovirus, that’s like, it’s a virus that causes the common cold in chimpanzees and they splice in a little bit of SARS-CoV-2 into that and so that sort of host virus, that adenovirus holds genetic material infects your cells and your cells then produce the antigen. They produce the spike protein of SARS-CoV-2. So AstraZeneca and Janssen are using this construct again, both with adenoviruses, a very common virus vector. And again, we like these types of vaccines because they’re quick to manufacturer, but a challenge of them is that your body can sort of develop separate immune responses against the vector itself, right? So you can develop a separate immune response against say an adenovirus right? Such that your body neutralizes those adenoviruses before they’re able to infect your cells and produce the SARS-CoV-2 antigen. So we do see tendency a kind of faster waning vaccine effects with this class of vaccines.
So moving on to subunit vaccine. This is NovaVax and Sanofi’s vaccine will be subunit vaccines and this is where I kind of mentioned before actually what happens here is these spike proteins or whatever the antigen is, created and purified in a lab. So they actually use insect cells that they infect with SARS-CoV-2, those insect cells then produce the antigen that’s purified and that’s what goes into the vaccine are those protein subunits, right? So there we’re just directly giving you the spike protein that we’ve grown outside of the host and that’s how we’re getting these APCs, those antigens. And so this is a commonly used vaccine construct. So the hep B vaccine is highly effective. HPV vaccine is highly effective. That’s the construct of these, but the downside of course to it, so it’s a well-trodden way of developing vaccines. But the downside is that they’re slower to manufacture. There’s this whole process where we have to cultivate and grow these viruses in a lab, we have to purify them,
and moreover they often also require an adjuvant.
So that’s really just sort of adding something a little bit extra that stimulates a better immune response in your body.
So basically at the site of injection, it’s something that increases your inflammatory response actually to kind of stimulate your immune system into recognizing those antigens.

So there’s subunit vaccines.

So the fourth class here is a weakened/inactivated vaccine.
And so this is, I think, what most people like what my grandparents probably think all vaccines are, is basically we take a pathogen and we weaken it in some way, or we kill it, right?
And then that’s the construct of the vaccine and that’s what’s injected into you.
And we go through this similar process there that literally mimics natural infection, right?
Where your cells are infected by this weakened form of the virus, the virus replicates, and that’s how we get antigens to the APCs.
So this is the construct used in of course some classic vaccines like MMR, polio vaccine, but again, it’s slower manufacturing, right?
Because we have to cultivate the virus in the lab and then it also requires adjuvants. So I don’t think there’s currently any plans to have US phase three trials of weaken inactivated vaccines, but there are in China.

So Sinopharm and Sinovac vaccines were using this construct. So that’s just a bit of a background in immunology and how all this works and how we think about preventing infection with SARS-CoV-2 and hopefully preventing clinical disease COVID-19 disease.

So now we’re gonna segue to talk a little bit about the vaccine development process, right? ’Cause this has all happened extremely fast. So let’s talk about sort of the process whereby vaccine products are typically brought to market, right.

And what looks a little bit different about the COVID-19 vaccine development process? So this is a figure from a nice New England journal paper that’s referenced at the bottom. That’s just talking about sort of what’s different in terms of how are we accelerating the vaccine development process.

And so I think as biostatisticians,
anyone who works on clinical trials is fairly familiar with the traditional paradigm for bringing products to market, right. It involves sort of a lot of R&D in the lab, preclinical work and you start doing human trials in phase one, these are small dose finding safety trials, checking whether these vaccines generate any immune response. And then what we’ll often do is run a small randomized trial. That’s a phase two trial, right? We’re we’ll have a placebo control, maybe pick out a particularly high risk population and start to see if we’re getting any efficacy signal, right? Phase one typically advances very slowly. Phase two, we think very hard about whether the efficacy signal was really worth it to advance a candidate to phase three and it’s a very deliberate process, right? To get to this phase three licensure trial, right? So the phase three trial is the big one involving the most participants. It’s a randomized controlled trial, right?
Enrolling many, many subjects that’s well powered to detect efficacy signals and based on the results of that phase three trial and safety data that’s been accumulated throughout this whole process, right. We’re able to provide licensure ideally for a product. And so that’s sort of the clinical development process, but also in the context of COVID vaccines it’s important to think about the manufacturing process, right. And how that looks a little bit different. So typically, companies are very sort of hesitant to scale up manufacturing before they know that they have a product that will be licensed, right. Which makes sense, you know, they’re sort of risk averse. We don’t want to start manufacturing a product that may ultimately be shot down by the FDA. So really large scale manufacturing is not happening until after product licensure. So what’s happening with COVID vaccine is basically this whole long deliberate timeline is being compressed into a shorter time period. And so how do we do that? Well, basically what happens is we’ve collapsed
the phase one and phase two trials, right?
So we’re doing small safety studies.
We’re checking whether these vaccines are generating immune responses,
but we’re really not doing that smaller efficacy study
that is typical of vaccine development.
And so we’re collapsing the phase one and two process,
the phase three process is where we’re at, right.
We’re doing these large scale trials, right?
Because we need robust efficacy data and we need robust safety data to gain licensure,
but a big thing that has changed, so the clinical process
yeah a little bit compressed, but mostly the same,
the big thing that’s changed
is the manufacturing process, right.
Is we wanna make sure that once a vaccine is licensed
and is proven to be safe and effective that we’re able
to start distributing that vaccine immediately.
So that means that manufacturing needs to start ramping
up right before we ever have a signal of efficacy
and that’s a huge risk for companies to take.
So, I’ll talk in a couple of slides about sort of how
the government has come in to try to remove
some of that risk from these companies
and then the next slide I think is just showing sort of that it’s really impressive that we’re even talking about potentially having a COVID vaccine available this year or early next year, just given the timelines that are required to bring effective vaccines to market. And so here’s just a few, you know, polio, measles, chickenpox, mumps, malaria on this list. It took about 30 years to get a partially effective malaria vaccine to market. So this is typically a very long process, right? And for COVID, we’re looking at hopefully doing this in just under a year or two. So how is the US government playing a role in this? Well, it’s through this program that you may have heard of called Operation Warp Speed, which is this huge convoluted mess of an amalgamation of programs across the government from DOD to many branches of NIH, BARDA, NIAID, so it’s sort of all over the place. And this is really just the same figure that I showed you from the New England journal paper.
Just maybe a slightly more confusing if you ask me, I don’t think Edward Tufte, he would be a big fan of graphic but the point here I want to mention is how is the government responding to COVID vaccine development? How are they contributing to that process? Well, there’s really two ways that they’ve offered to accelerate the process. The first is through funding of phase three clinical trials, right? So a number of companies, six of the major companies, basically every company that’s running a phase three trial in the US besides Pfizer that you’ve heard about is contracting with BARDA. That’s an arm of the NIH, they’re contracting with the government to have the government fund their phase three trials. So it’s a joint agreement between the government and these companies where the government, you the taxpayer, right, are paying for these phase three trials that will eventually lead to licensure. So that’s the first way that the government is sort of throwing money at this problem. It’s through design and paying for these phase three trials.
The second way is that they’re paying for manufacturing, right?

They’re removing that risk for these companies by basically committing to buy a certain number of doses before we ever have any efficacy data. So we’re in the hole basically to all of these companies for a fixed number of doses right. But that motivates the companies then to scale up their manufacturing ahead of the time that efficacy data are available. And that type of agreement has been entered into with Pfizer as well.

So all of these companies that OWS Operation Warp Speed is running the phase three trials for also have this manufacturing agreement. Pfizer has that manufacturing agreement as well.

So what role have I played in any of this big messy thing? I work with a great group of scientists in the COVID-19 Prevention Network. This was a clinical trials network established by National Institute of Allergies and Infectious Disease so that’s an arm of NIH, and it’s basically anyone who works in clinical trials with these clinical trials networks, right?
It’s an amalgamation of researchers and study sites, laboratories, people who focus on recruitment and retention of trial participants, statisticians. So it’s researchers who are really experts in running clinical trials, designing clinical trials and ensuring their robust conduct. The CoVPN was formed by basically leveraging four existing clinical trials networks. One of which I was a part of, which is the HIV vaccine trials network. And so from our group, we’ve really brought a great group of statisticians, many of whom are at the Fred Hutch in Seattle as well as great groups of laboratories at U Dub. And so what are the roles that we’re playing in these trials? So in our statistical group, there’s a couple of statisticians who are designated as like CoVPN representatives for each of these phase three trials. So I sit on calls with these trials and advise on their design and analysis approaches for their efficacy monitoring, for their safety monitoring.

We help them address DSMB and FDA comments and sort of that’s all happening in conjunction with...
with both government statisticians, right.

Representatives of BARDA and NIAID as well as company statisticians.

And so we get on these calls and, you know, nerd out over clinical trials,

statistical decision-making, and it’s a good old time.

Another aspect that we really contribute a lot on,

or that CoVPN has sort of been tasked with taking

the lead on is the development of immune correlates.

And so that’s the part of my talk

where I’ll get a little bit into statistics

and talking about what immune correlates are,

some of the types of analytic approaches

we use to study those and the idea of immune correlates

just to give you a teaser

so you don’t, you know, sign off Zoom early.

So immune correlates are really the idea there is

we’re looking for immune responses that are predictive

of the vaccines working, right.

So what we’d really like to be able to do is understand,

okay, if we’re able to generate this level

of neutralizing antibody,

then that will lead to this level of protective effect

of the vaccine, right?
So that’s the whole goal there is identifying what are these immune responses that are responsible for providing protection?

Okay so I’m gonna walk through just a few of the design and analysis questions.

And so these are things that have come up as we’ve worked with these company statisticians, as we thought about sort of the whole OWS vaccine program, what are some of the issues that statisticians are kicking around and people who have worked on clinical trials, right, on clinical trials, right,

a lot of these issues aren’t gonna be new and one thing that I think is sort of interesting about this whole pandemic and operating as a public health professional

in this and a clinical trial statistician in particular, is that a lot of things that we take for granted as scientists are either very confusing or sort of counterintuitive for a lot of the lay public.

And so it’s been sort of interesting to have that laid bare.

In some of these issues, some of these things that we think are no-brainers like doing interim analysis for example are kind of highly controversial and have ended up being, you know, sort of areas of huge disputes.
And so I just want to run through some of these issues that I think are quite fascinating, a lot of which, you know, really don’t have a correct answer and they’re really just sort of food for thought the types of things that we’re thinking about when we’re designing these trials. So I’ll start by just giving a sort of more specific idea of what these trials look like and how they’re conducted.

I’ve picked AstraZeneca because that’s the one I’ve worked on for the longest and most closely, but all of the trials sort of follow this similar design.

And so the first thing I’ll note is that you can read these trial protocols. So one of the interesting things that’s happened in this COVID-19 development processes is there was a huge public push led by like Eric Topol and others to have the protocols of these trials made public, which when it happened was I guess when that push started happening, you know, I emailed all my colleagues and said, really do we not usually make protocols public? And that was just sort of interesting disconnect for me as an academic who’s used to sort of everything.
being open science and that’s a no brainer right.

Working in this setting, right, where these protocols are really seen as trade secrets for pharmaceutical companies. So it’s really unusual that actually these protocols have been made public. So it’s sort of neat, but one of the things that happened is all of these protocols went public and reporters got their hands on them and said, wow, these are really dense documents, right? These are very sort of detailed documents that are kind of hard for the public to parse. So it’s been sort of a be careful what you wish for thing in terms of releasing these protocols, but that’s an aside.

So let’s talk about actually what these trials look like. Here’s a schematic, and again, this is AstraZeneca in particular, but this is basically the design of most of these trials.
will look something like this.

So who is the population?

Most of these trials are gonna be primarily in adults.

I think Pfizer has now started to talk about including children.

I'm not exactly sure where that’s happening, but adults for the most part,
these are mostly healthy individuals.

That don’t have, you know, chronic diseases that are at risk or high risk of death.
And we’re really looking at targeting individuals who are at an increased risk for SARS-CoV-2 acquisition.

and severe COVID disease and so the idea there is number one these are the people that are bearing the brunt of the pandemic, right?

So we want to be able to get a product to those people as fast as possible.

But number two also, right, that means that we’ll accrue from a sort of cold hearted and statistician point of view that means we’ll accrue end points faster.

We’ll observe more cases of COVID-19 disease and potentially get an efficacy signal a little bit faster.

So there’s a lot of interest in sort of recruiting and retaining individuals at high risk for COVID-19.
So you can go onto the COVID-19 prevention trials network and fill out a survey, right. Then we’ll basically under the hood assess your risk for COVID-19 and if you’re found to be at high risk, we’ll aggressively email you and try to get you enrolled in one of these trials. If you’re at low risk, we’ll say, thanks for taking the survey, we’ll be in touch and likely you won’t hear from us anytime soon. Okay so that’s the trial population. So how does the actual trial conduct look? AstraZeneca is using a two to one randomization scheme. So you have two chances of getting the active vaccine versus one chance of getting a placebo. And in this case, it’s a true placebo, just a saline dose and then most of the vaccines, most all with Janssen being the accepted are two dose vaccines. So you receive the first dose at day one and the second dose about a month later. And in the interim, we take a couple of measurements. We have a phone call to assess reactogenicity right. Does your arm hurt, or have you experienced any adverse side effects?
effects of the first dose of vaccine?

And then there’s also an immune response measurement that happens after a couple of days.

So we get an early signal of how immunogenetic these vaccines are.

And so then individuals come in for their second dose.

of vaccine and it’s a similar story, right?

Did you have any reactions?

We measure your immune response and after that,

that’s sort of when the clock starts for active follow-ups.

So this day 57, that’s two weeks roughly after,

am I doing that math right?

Well, it looks like roughly two weeks after the second dose.

of the vaccine is typically when this clock is gonna start and we’re gonna start counting COVID events.

And then it’s sort of just the standard sort of game we play in clinical trials.

We wait for events to accrue.

We have certain monitoring plan for when we’re gonna check for efficacy and we’ll talk about some of that.

So, I just want to note that there’s sort of two ways that we’re ascertaining events

that are happening here, right?

The first is passive monitoring.
What that means is we basically wait for individuals to present with symptoms of COVID-19, right? So you get a cough, you lose taste, right? You call the study site, right? So I am having these symptoms. They say, come on in. You get a PCR test to see whether you’re infected. In that case, you would count as a COVID-19 endpoint, right? If you check off some check boxes for symptoms you’d go down as a COVID-19 endpoint.

There’s also these sort of active follow-up visits. So these like day 90, day 180 and day 360, and at those visits we’ll do a serology check.

And what that means is we basically take a blood draw and we measure whether you have antibodies that are distinct against SARS-CoV-2, right, antibodies that are generated in response to the vaccine.

So we’re basically able to tell whether you were infected in this sort of interim period, when you show up for these visits. So that’s active follow up and so there you’re gonna be able
to pick up sort of asymptomatic cases, right? 'Cause if you never have symptoms, you’ll never come in and be captured by passive followup. So we have to wait for these set clinic visits to do the serology testing, to ascertain it asymptomatic cases. And so this is gonna actually play a role in a little bit, when I started talking about, you know, what are the end points that we’re thinking about measuring? Like, what do we want to know how well the vaccine works at preventing? Is it asymptomatic infection? Is it disease? Is it severe disease and so forth? So we’ll talk through some of those issues, but just want to note already that the design has started to inform some of the challenges that we might see when we want to talk about how well the vaccine works against certain forms of infection and disease. And so I think if you read the newspaper and you’ll see the term vaccine efficacy tossed around a lot. So the first thing I want to talk about is right, what is the primary hypothesis that these trials are trying to test? And what is the parameter?
What is the estimate, right, that they’re going after in these trials and for whatever reason nobody consulted me when they decided that VE would be measured in this way. But for whatever reason, we studied this that we quantify the efficacy of a vaccine in a sort of weird way. So a vaccine efficacy, we describe as the percent reduction in relative risk comparing vaccine to placebo. So it’s this one minus a risk ratio. There’s a one minus a risk ratio where you take the risk in the vaccine and the numerator and the risk in the placebo and the denominator. So, I mean, we can just play a quick little intuitive game, right? How do we get a VE close to one that would be a perfect vaccine? Well, we would make the risk in the vaccine close to zero, right? If you have a perfectly effective vaccine, there’ll be no risk of infection and or disease amongst the vaccinated. So you would get VE close to one. But on the other hand, how do we make VE zero? Well, we would take the risk in the vaccine
and set it equal to the risk in the placebo, right.

In which case basically saying the vaccine’s not doing anything and then on the other hand, a VE is negative, right?

That’s indicating that there’s actually higher risk in the vaccine.

So just to give you sort of a few reference points, right?

So that VE of one is perfect, VE of zero is nothing and what we’re really hoping for with these COVID trials is a VE of at least 50%.

And that’s sort of the cutoff that FDA guidance has stipulated is that you need to show a point estimate of VE for your primary end point.

And again, we’ll talk about what these primary end points are but we need a VE against a primary end point.

of at least 50% and we need to definitively rule out the possibility that the vaccine efficacy is less than 30%.

So basically we have to reject the null hypothesis that VE is less than 30% along with having a point estimate of VE being greater than 50%, right.
And we need to do that while controlling type one error at two and a half percent.

Okay and so here, just one final note, since this is a statistics talk, I’ll talk a little bit more about what I mean by risk, right?

So risk here can be quantified in a number of ways and it often is.

So we can quantify this using hazards, for example, like you can imagine fitting a Cox model, right.

A proportional hazards model, right.

That only adjusts for vaccine, right.

And presenting like one minus a hazard ratio from a Cox model, that’s something that’s commonly done.

You can also think about cumulative incidents, right?

So like mapping,

maybe one minus a survival probability as a way

of quantifying risk, incidents rate ratios.

So they’re all sort of used for different vaccines.

And usually we like to sort of argue about which one of these is better and I’ve thought a lot about that in my career.

And in this setting, it turns out because COVID is such a rare event that all of these ways of quantifying rates are basically the same and you end up
with almost identical operating characteristics of a trial.

So it’s really not worth sort of losing sleep over whether we’re talking about VE in terms of hazard or incidents or incidents rate and so forth.

So how are folks going about estimating this VE?

Here’s just a quick table of the four most advanced phase three trials, the ones that have released their protocols at least.

So we see for Moderna, AstraZeneca, and Janssen, they’re using pretty kind of the standard approaches.

Moderna a Cox model as I describe, AstraZeneca a Poisson regression model, it’s like, okay, that’s basically a Cox model, and then Janssen is using a sort of exact binomial test with this sequential probability ratio rest.

Pfizer is a little bit of the oddball. So they have stipulated a bayesian approach wherein they’re basically specifying a prior for vaccine efficacy and are using sort of a beta-binomial bayesian approach to evaluate the posterior probability of the vaccine efficacy is greater than 30% and so at the end of the day, there’s four different statistical methods here.
Again, if you do a simulation study with parameters that are approximately similar to what we expect to see in these COVID trials, you’re really not gonna see much difference in terms of operating characteristics across these. So it’s interesting to notice that assertions that there’s these different approaches, but at the end of the day, we’re basically talking about how many vaccinated people get infected, how many placebo people got infected, and almost all of these methods are gonna yield very similar inference. When it comes down to brass tacks, how many numbers fall into those categories? That’s a little bit about sort of how we quantify VE in these settings but one of the big things I haven’t described yet is VE against what, right? What is the end point that we’re measuring here? And so here’s a figure from a paper we just had come out in Annals of Internal Medicine, the link’s here. This is where we were spending a lot of time, you know, earlier this summer, thinking about,
What's the right endpoint for a primary analysis of the clinical trial? And it's complicated for something like SARS-CoV-2, right? Because we know we can start up here with the SARS-CoV-2 infection, right? That's sort of the base, you can become infected and then a number of things can happen, right? You can go on to be infected but develop no symptoms. So we would call that an asymptomatic infection, or you can develop symptoms, right. In which case we don't call you a SAR-CoV-2 infection anymore, we call you a COVID-19 disease endpoint. You have a clinical manifestation of your infection. But even beyond that, right, amongst people who exhibit symptoms some of them, maybe many of them are quite mild, right. So we have this kind of category of non-severe COVID, whereas others we know that are extremely adversely impacted by infection and end up with severe COVID disease. So you have all of these choices of sort of which end points you might want to talk about
and so I'll kind of walk through some what I see as the positives and negatives of this and then I'll also talk about this burden of disease very briefly end point that we’ve put together and so that’s kind of a composite end point that we’ve suggested that could kind of bring all of these different end points together. Okay so starting with SARS-CoV-2 infection, right? Asymptomatic, symptomatic don’t care, let’s count any infection as an event and measure VE against preventing infection. Okay and so that’s definitely relevant, right. It’s relevant the context of a pandemic. We’re preventing infections, we’re preventing spread of the disease, we’re bringing our knot down, we’re impacting the pandemic. And moreover, we’re going to see many more infections than we will cases of symptomatic disease. We know that many people who were infected never go on to develop symptoms so thinking about having an answer faster, right. SARS-CoV-2 infection is a nice endpoint, but then the question is, is it a clinically relevant endpoint?
So it’s really not describing an impact on patients at all. So we could kind of question its relevance from that perspective. The other thing, right, is that we remember going back to the study design, we’re only able to ascertain asymptomatic infections sort of very coarsely in time and moreover you have this phenomenon that happens is that when you’re testing many, many individuals, right.

It’s sort of the classic biostat one-on-one problem that we give people, right. You’re testing many individuals, but the prevalence is low. So even if you have high sensitivity and high specificity, you could end up with low positive predictive value. And the effect of that when you come to the time to analyze the data is that you’ll be biasing VE towards the null.

So it’s actually, while it seems like maybe a nice end point from the perspective of observing many infections, it’s a very challenging endpoint to analyze quantitatively. So moving down we could talk about COVID. So again, COVID is just infection, PCR confirmed infection with clinical symptoms.
So that’s of course more clinically relevant, right.

Because we’re starting to talk about an impact, excuse me, the endpoint that impacts patients.

All right so that’s more clinically relevant and moreover we’ll expect to have a reasonable number of cases, right.

By including more mild cases, for example, in this endpoint definition.

But then on the other side of that coin is it really that clinically relevant if we’re just talking about mild symptoms?

We’re talking about a disease where you get it and you end up with a little cough for a couple of weeks and that’s it.

We want to be protecting the most vulnerable individuals so we should be quantifying how well our vaccines work towards preventing those most severe endpoints.

And so most clinically relevant, and also there’s sort of a long history of vaccine development where really we see the best VE against severe cases of disease.

So that’s really where we expect the vaccines
to have the most impact is maybe we are not preventing you from being infected but we’re lessening the symptoms once you become infected.

So we’re not totally blocking transmission but we’re making a clinical impact on disease and that’s sort of been seen for a number of vaccines in the past. The downside of this end point of course is that there’s very few cases expected to be observed.

So amongst all infections, only a fraction have any symptoms. Amongst those with any symptoms, only a fraction develops severe symptoms. So we’re really whittling away the number of end points. So we need to do larger trials or have longer follow-up to evaluate this endpoint. And so in that paper, I’m sort of pressed for time so I won’t spend too much time talking about this, we also proposed this burden of disease measure where you’re sort of scoring these these outcomes, right?

So maybe you would get a score of zero if you’re an asymptomatic infection ‘cause it’s really no burden on you as a patient, right? You don’t have any symptoms.
And then we’re sort of assigning arbitrarily a score of one for non severe COVID so that’s like mild cases of COVID and a score of two for severe cases of COVID and this end point actually has some nice operating characteristics we think, but of course it’s subject to controversy, anytime you start talking about an ordinal scoring system, right, you start to raise questions about how you’re assigning the burden of disease score, right? Why should severe cases be a two versus a three versus a five and so forth? So you can kind of get bogged down in some of the specifics of that. So what has FDA said about this? So FDA guidance documents states that either the COVID end point or SARS-CoV-2 infection is an acceptable primary endpoint and then somewhat ironically OWS has been telling companies that infection alone is not acceptable as a primary end point. So we had one company that was interested in including that as co-primary and for whatever reason we told them please don’t do that, and then beyond that so COVID has sort of won out as the end point of choice.
But beyond that, FDA guidance states that companies should consider powering efficacy trials for the severe COVID endpoint as a co-primary or at least as a key secondary endpoint in the trial. So far, only Janssen has taken up that offer of making severe COVID primary. And that’s why, if you look at the number of individuals planning to enroll in their trial, it’s twice as many as any of the other OWS trials. So like AstraZeneca is planning for 30,000, Janssen is planning for 60,000 in their trial. And that’s the power, to see more cases of severe disease to be sufficiently powered to detect VE against that.

This is a controversial slide. Or this is virtual topic I found, something that clinical trials statisticians sort of take for granted is doing interim analyses, right? If the treatment is working and we have enough evidence we’d like to stop that trial early to get that treatment to patients, right. One would think that that’s true here and so many of these trials...
were designed with aggressive sort of interim looks, right?

Because we’re in the middle of the pandemic and we’d like to get a vaccine to individuals as quickly as possible.

So I have a table, we won’t go through it all here,

just sort of the planned interim analysis for these different trials.

I would say Pfizer seems to be the most aggressive so far.

They have five interim looks or four interim looks

and a final look at their data, right?

So that’s fairly aggressive.

OWS again, the trials that we’re running,

we’re really encouraging companies to be a bit more conservative in the approach to this

and only maybe two or three

so you see what’s been adopted by Moderna and AstraZeneca

and so this was really a big point of contention

I think when these protocols were made public is this idea

that like, can you really know that a vaccine works

based on 32 data points, right?

We’re talking about a vaccine that’s going to be given to billions of people around

the world based on these results

and you’re gonna do that based
on the results in 32 individuals? And like, so I can stare at the math and say that like, yes, that appropriately controls type one error and so forth, but it still makes me just feel a little bit uncomfortable. There's a bit of dissonance between sort of my life as a statistician and just me being a human and saying 32 data points is probably not enough to decide to vaccinate billions of people. And so a lot of people I think sort of shared that viewpoint and in response FDA has now been sort of backpedaling in a way and asking companies to provide more data in order to grant an emergency authorization. FDA has now said that companies need to have 50% of participants complete at least two months of follow-up for safety signals and that you need to have at least
six COVID cases in the oldest age group. Of course, that’s an age group of particular interest in terms of severe cases and at least five cases of severe COVID in the placebo group. So they want to be able to see some data, even if you’re not specifying severe COVID as a primary end point, they want to be able to see some data, some signal of efficacy against that in order to grant licensure. So I’ll sort of, I won’t go through this slide. It’s just to say that like, sort of when Pfizer released their protocol, everyone was like, ooh a bayesian analysis and got very sort of skeptical, right? Because the Pfizer CEO has been out there sort of chest thumping and saying they’re gonna have a vaccine before the election and so forth and then they came out with this bayesian design that was a little atypical and so everybody was asking the question, well, are they trying to hide something? So I sort of did a quick analysis and found that really it doesn’t look that different than a classic kind of post hoc monitored design. And if you want to read more about that,
I have some slides up on my GitHub about it.

So let’s see, I’m running low on time so I’m gonna skip over sort of the question of what happens if efficacy is declared early.

So I have some reasons that we should be excited, right?

If one of these trials stops earlier, I can get a vaccine.

There’s good data that the vaccine works and that’s nice.

I’d like to go back to something resembling normal as I’m sure you all would,

but of course there’s reasons to be concerned, right?

If efficacy is declared early in particular, if that means that blinded follow-up in a study stops, right?

Because that means we have no way to assess how durable the vaccine is.

We won’t be able to assess VE and key subgroups that we care about.

We might not be able to assess VE formally against severe end points.

So there’s real sort of concerns about stopping these trials too early, and what the implications of that are both for evaluating the vaccine in question, but as well as how it impacts the other clinical trials that are ongoing.

And of course in the current political climate,
everybody’s very concerned about the role political pressure might play in all of this. So yeah, so it’s kind of a double-edged sword in some sense in terms of what happens if efficacy is declared early,
but I want to save just a few minutes to talk about vaccine correlates ’cause I promised that I would show you some math and prove to you that I’m a real statistician. So let’s do a little bit of that.
So again, we’re kind of shifting gears here. So that’s the end of sort of talking about the primary analysis of these trials, what’s gonna lead to their licensure. And the correlates of protection is sort of a key secondary analysis and so why is it so important that we’re able to establish correlates of protection?
Well, because it’s gonna speed up the whole vaccine development process. So again, a correlative protection is really just, it’s an immune response and really an assay to measure that immune response that’s been validated to reliably predict vaccine efficacy. So why is that so important?
Well, basically what we’re hoping to achieve
is the establishment of a surrogate endpoint for COVID disease right?

So I’ve sort of mentioned the numbers that we’re talking about in these phase three trials,

enrolling 30,000 participants, 60,000 participants and ending up with one or two years of followup, right.

Just to be able to answer the primary question, right.

Does the vaccine prevent infection and/or disease?

So that’s a huge, expensive clinical trial.

It takes a long time to get an answer and so it would be very nice if all we had to do right

was give people the doses of vaccine that they need,

wait two weeks and measure their immune response and understand does that vaccine work or not.

That would be a much faster vaccine development process

than where we’re currently at

in having to run these enormous phase three trials.

So it’s valuable for establishing a surrogate endpoint.

It’s also valuable for accelerating approval of vaccines that have been licensed in certain populations, but not others.
For example, I mentioned that these phase three trials are mostly being conducted in adults. Well, what if we want to also obtain licensure for use of this vaccine in children? Well, if we had an established immune correlate we wouldn’t have to do a large randomized trial in children. We could do it just a small immunogenicity study and use the correlates results to bridge the VE that we observed from the phase three trial. That’s the immune response that we’ve observed in these children or pregnant women for example are another key population they’re being excluded from these phase three trials but we’d like to understand if these vaccines are safe and effective in those women as well. So really this is one of the key goals of this whole OWS program and the key role we’re playing in the CoVPN is developing the sampling plan and the statistical analysis plan for the immune correlate studies and so it’s just a little bit of the statistical issues that we’re dealing with in these trials, right, is that sort of running assays.
so running these immuno assays on 30,000, 60,000 individuals takes a long time, it’s expensive, and as it turns out, it’s really overkill in terms of statistical power. So we can actually be a little bit more clever about how we design these correlate studies in order to get answers faster and more cheaply. So the way we do this is we use a case cohort design. So we’re not gonna measure immune responses in all trial participants, we’re gonna measure them in a sub cohort and that sub cohort will consist of a stratified random sub cohort. So we’re gonna be sampling individuals randomly based on their baseline infection status. Were you infected with SARS-CoV-2 in the past? Based on your race, ethnicity, and based on age. And so based on that, we’ll take a random draw of the trial population, about 1600 individuals, excuse me and everyone so I should mention right everybody is having their blood drawn. And right now we’re talking about whose blood
are we gonna use to measure these immune responses?
So we’re gonna measure it in a random sample
and then we’re gonna wait until the trial is over
or until one of these interim analysis concludes efficacy
and we’re gonna measure immune responses
in all of the end points, right?
Remember that like power in these analyses is drive
by the individuals in which we observe end-points.
So we’re gonna make sure we get immune responses
in all the end point data, as in addition
to this random sub cohort and it turns out that that’s about
as statistically efficient as running the immune assays
on all 30,000 individuals in the trial.
So this is this kind of classic case cohort design
that Ross Prentice has been writing about for years
that Norman Breslow did some sort of pioneering work
in the 2000s and I’ll just talk a little bit about sort
of how this complicates our life as statisticians
and then maybe we’ll leave a few minutes for questions.
So here’s the math, we made it.
Well, the moment you’ve all been waiting for to see some math. So just introducing, you know, why is this sampling design challenging from a perspective of generating estimators, right? Well, we can sort of immediately see that this isn’t a totally random sample of the trial population, right? In particular we’ve over-sampled the individuals who end up getting diseased and it’s fairly obvious that those individuals have potential to be very different than a randomly selected individual in the population. So we have a bias sub sample. So we need some statistical methodology to try to back out, you know, whatever this parameter is. We want to be estimating it in the whole trial population, not just in this biased sub samples. So how do we do that?

So just a quick notation here, let’s call W baseline covariates, A is a binary vaccine assignment, Y is your binary COVID endpoint for example and then we’ll introduce this sort of indicators. Delta is one, if you’re selected into this immune response
sub cohort, either because you were a case, you were an end point or because you were randomly selected. And then we’ll call $S$ your immune response. And then we’ll just say, we’ll represent this as $\Delta S$, which just means we’ll arbitrarily set everybody who’s not in our sub cohorts immune response to be zero, that’s arbitrary doesn’t really matter. So let’s talk about how estimation would happen. So let’s pick a very simple parameter, right? Let’s just say that we want to know what’s the overall immune response in the whole population, not a particularly interesting parameter for actually measuring correlates, but just to motivate the types of statistical approaches we use in these settings. So how can we control for the bias of the sampling design? Well, one of the most straightforward ways is to use the tried and true Horvitz-Thompson or IPTW estimator, right. Where we’re just taking basically a sample mean but all our observations are sort of inverse weighted by their probability of being sampled into this sub cohort.
And so that’s, IPTW estimator if you’re in causal inference, you’re very familiar with this. If you’re in survey sampling, very familiar with this. Very classical way of adjusting for this selection bias. It turns out that there’s ways that we can be more efficient in doing this. We can use augmented estimators, AIPTW estimators. And the key idea there is that we take the IPTW estimator and we add a little bit of something to it and the key thing is that that little bit of something involves a regression of the immune response onto the covariates that were used to sample individuals. And so what’s the intuition as to why this is more efficient? Well, you can imagine what if we had a perfect predictor of S measured at baseline, right? Then this regression here is essentially imputing the correct value of S for every single person in the population. So it’s kind of like we’re getting more data in some sense, and the nice thing about these approaches, these AIPTW approaches is that they’re double robust and so again,
if you work in causal inference a very familiar idea, and it turns out because we know the sampling probability by design, this regression doesn’t have to be consistently estimated in order to obtain a consistent estimate of the parameter measures. So it’s this really nice sort of double robustness property that says, yeah, you might be turned off from this augmented estimator because you have to do a little bit of extra work, you have to fit a regression model say, and maybe you’re worried about misspecifying that regression well it turns out that because the sampling probabilities are known by design, you don’t have to concern yourself with that. So it turns out you can use any old regression estimator here and still end up with a consistent estimate of the parameter of interest. And so we’re applying this to much more interesting parameters. So we had a paper come out recently in biometrics that’s linked here where we’re starting to study a sort of causal inference flavored parameters in this context, things that we can really use to pin down,
you know, mechanisms of these vaccines working.

So, in this case, we’re studying sort of the effect of a stochastic intervention, we call it. So it’s sort of saying what would happen if we took everybody’s immune response, this particular immune response that we observed.

And so you ended up looking at a plot that’s kind of like this. So this is from an HIV vaccine trial. So at zero we’re saying that’s just the observed risk. And so we can see in this example, we found that the risk would be increasing, right.

And then if we’re moving to the right, what would happen if we increase your immune response. And so we’re kind of getting at something that’s like a controlled effects mediation type parameter.
with this approach and so we’re working out some details of the correlates plan currently and so when that’s done we’ll have it available for public comment. And again, we’re academics, right? So we’ll do it all open science. And then I’ll just say like two words of conclusion and I’ll shut up and leave some time for questions. So there’s been a big concern in the current political climate that we’re gonna sneak something through, that something’s gonna be approved without sort of the standard amount of evidence that would be required, right. That there’s political interference at the FDA and from where I sit, you know, I can say that the science behind the vaccine development program for COVID is extremely rigorous. These are exactly the type of people who you would want in charge of this decision making process and the type of people that will raise red flags as soon as sort of the process goes off the rails. So right now I feel good about where things stand. Of course, I watch presidential debates and hear, you know,
garbage science coming out and I get a little bit concerned.,
but from where I sit right now,
everything’s looking pretty good.
So overall, I’d say that the increased transparency
by releasing these protocols
has been good for scientists and consumers.
We want to bring vaccines to market,
but we also want people to trust those vaccine
so increasing transparency in whatever way we can is great.
And then finally, the final point is that a lot of these
issues that I’ve talked about,
how do we do interim monitoring, right?
What’s the right end point to be studying?
What’s the right S demand, right?
These are really hard decisions
and there are no right answers.
And so one of the things that’s been a little bit
disconcerting or disheartening to me
is the extent to which in the pandemic era,
academic debates have been made very much public
and I’m not against academic debates.
It’s just that most individuals aren’t used to seeing them.
And so what I’m worried is happening is that people
see high profile academics debating these challenging
problems where there’s no real right answer.
And they’re saying, well, these guys don’t know what they’re talking about.

So I think as academics and public health professionals in this pandemic, one thing that we can do is just to be very careful in how we’re presenting, you know, the science that we’re doing and acknowledge when there’s not a right answer, that you’re presenting your opinion. And that there is some validity, right? That this is very gray, unfortunately, that there’s nothing black and white here. So maybe that’s a controversial statement to end on, but I’ll end there and then thanks again to Fan for giving me the opportunity to talk and I’m happy to take questions as there’s time.

I don’t have anything scheduled after this, so I can stay a few minutes over as would be helpful. So thanks again.

Thank you David for this very nice talk. I think we do have three to four minutes for questions from the audience, if there’s any. Hi David, I have a question 'cause right now for COVID situation and because of the time
and the faster progress of the disease and it’s a hard to keep the standard method, but do you have other proofed vaccine for other disease and have a quick trial have a similar way as COVID and apply the method you’re using right now and we have standard results already and then compare to see how good the current method is. So that’s my question. Yeah it’s an interesting question.

So let me try to restate, so you’re saying, are there any lessons from vaccine development that we can try to draw from here to evaluate our methodology, whether it work? From other vaccines.

[David] So I guess what I would say is that at this stage, in phase three vaccines, these phase three trials look completely normal. So I would say the process of getting to the phase three looked very different and much more accelerated in terms of kind of squashing together and phase one and phase two in terms of the manufacturing, but in terms of what’s happening in a phase three trial, this is probably the phase three trial.
that would be done outside of the setting of a pandemic.

Maybe the interim analysis would be a little bit less aggressive for some of these companies, but really, I think the approaches that the companies are taking would be fairly standard even in any other vaccine context.

Even though for the established vaccine, there could be some field trial and that they also went through a phase three, but you can do the similar thing to enhance, to see whether it is possible to pass the current protocol and become some sort of false positive.

As someone who works in HIV vaccines, we’re very familiar with failure and learning from that. So again, I think the people who are running these trials are sort of the right people in terms of looking out for these false positive signals and so forth.

Thank you.

So I think we are just about the time and I’m sure that David is happy to take your questions afterwards by email.
So I’ll thank David more time. Again, thank you for sharing with us and we’ll see everyone again next week. [David] Thanks everybody.