**PROFESSOR:** So welcome, everyone. Today is the first of what will be a series of four guest lectures throughout the semester. There will be two guest lectures, starting the week from today, and then there'll be another one towards the end of the semester. And what Pete and I decided to do is to bring in people who know a lot more than us about some area of expertise.

In today's instance, it's going to be about cardiovascular medicine, in particular about how to use imaging and machine learning on images in that context. And for today's lecture, we're very excited to have professor Rahul Deo to speak. Rahul's name kept on showing up, as I did research over the last couple of years.

First, my group was starting to get interested in echocardiography, and we said, oh, here's an interesting paper to read on it. We read it, and then we read another paper on doing subtyping of ejection fraction which is a type of heart failure, and we read it. I wasn't really paying attention to the names on the papers, and then suddenly, someone told me, there's this guy moving to Boston next month who's doing a lot of interesting work and interesting machine learning. You should go meet him.

And of course, I meet him, and then I tell him about these papers I read, and he said, oh, I wrote all of those papers. He was a senior author on them. So Rahul's been around for a while. He is already a senior in his field.

He started out doing his medical school training at Cornell, in Cornell Medical School, in New York City, at the same time as doing his PhD at Rockefeller University. And then he spent the first large chunk-- after his post-doctoral training, up here in Boston, at Harvard Medical School-- he spent a large chunk of his career as faculty at UCSF, in California. And just moved back this past year to take a position as the chief data scientist-- is that right-- for the One Brave Idea project which is a very large initiative joint between MIT and Brigham and Women's Hospital to study cardiovascular medicine. He'll tell you more maybe.

And Rahul's research has really gone the full spectrum, but the type of things you'll hear about today is actually not what he's been doing most of his career, amazingly so. Most of his career, he's been thinking more about genotype and how to really bridge that genotype-phenotype branch, but I asked him specifically to talk about imaging. So that's what he'll be focusing on today in his lecture. And without further ado, thank you, Rahul, for coming here.

[APPLAUSE]

**RAHUL DEO:** So I'm used to lecturing the clinical audiences, so you guys are by far the most technical audience. So please spare me a little bit, but I actually want to encourage interruptions, questions. This is a very opinionated lecture, so that if anybody has sort of any questions, reservations, please bring them up during lecture. Don't wait till the end.

And in part, it's opinionated because I feel passionately that the stuff we're doing needs to make its way into practice. It's not by itself purely academically interesting. We need to study the things we're doing. We're already picking up what everybody else here is already doing.

So it's OK from that standpoint, but it really has to make its way. And that means that we have to have some mature understanding of what makes its way into practice, where the resistance will be. So the lecture will be peppered throughout with some opinions and comments in that, and hopefully, that will be useful.

So just a quick outline, just going to introduce cardiac structure and function which is probably not part of the regular undergraduate and graduate training here at MIT. Talk a little bit about what the major cardiac diagnostics are and how they use them. And all this is really to help guide the thought and the decision making about how we would ever automate and bring this into-- how to bring machine learning, artificial intelligence, into actual clinical practice. Because I need to give enough background so you realize what the challenges are, and then the question probably every has is where's the data? How would how would one get access to some of this stuff to be able to potentially do work in this area?

And then, I'm going to venture a little bit into computer vision and just talk about some of the topics that at least I've been thinking about that are relevant to what we're doing. And then talk about some of this work around an automated pipeline for echocardiogram, not as by any means a gold standard but really just as sort of an initial foray into trying to make a dent into this. And then thinking a little bit about what lessons-- David mentioned that you talked about electrocardiogram last week or last class, and so a little bit of some of the ideas from there, and how they would lend themselves to insights about future types of approaches with automated interpretation.

And then my background is actually more in biology. So I'm going to come back and say, OK, enough with all this imaging stuff, what about the biology? How can we make some insights

there?

OK. So every time people try to get funding for coronary heart disease, they try to talk up just how important it is. So this is still-- we have some battles with the oncology people-- but this is still the leading cause of death in the world. And then people like I, you're just emphasizing the developed world. There's lots of communicable diseases that matter much more. So even if you look at those, and you look at the bottom here, this still, if this is all causes of death age-adjusted, cardiovascular disease is still number one amongst that. So certainly it remains important and increasingly so in some of the developing world also.

So it's important to think a little bit about what the heart does, because this is going to guide at least the way that diseases have been classified. So the main thing the heart does is it's a pump, and it delivers oxygenated blood throughout the circulatory system to all the tissues that need it-- the brain, the kidneys, the muscles, and oxygen, of course, is required for ATP production. So it's a pretty impressive organ.

It pumps about five liters of blood a minute, and with exercise, that can go up five to seven-fold or so, with conditioned athletes, not me, but other people can ramp that up substantially. And we have this need to keep a very, very regular beat, so if you pause for about three seconds, you are likely to get lightheaded or pass out. So you have to maintain this rhythmic beating of your heart, and you can compute what that would be, and somewhere around two billion beats in a typical lifetime.

So I'm going to show a lot of pictures and videos throughout this. So it's probably worthwhile just to take a pause a little bit and talk about what the anatomy of the heart is. So the heart sits like this, so the pointy part is kind of sitting out to the side, like that. And so I'm going to just sort of describe the flow of blood.

So the blood comes in something called the inferior vena cava or the superior vena cava, that's draining from the brain. This is draining from the lower body, and then enters into a chamber called the right atrium. It moves through something called the tricuspid valve into what's called the right ventricle. The right ventricle has got some muscle to it.

It pumps into the lungs. There, the blood picks up oxygen, so that's why it's shown as being red here. The oxygenated native blood comes through the left atrium and then into the left ventricle through something called the mitral valve. We'll show you some pictures of the mitral valve later on. And then the left ventricle, which is the big workhorse of the heart, pumps blood

through the rest of the body, through a structure of the aorta. So in through the right heart, through the lungs, through the left heart, to the rest of the body.

And then shown here in yellow is the conduction system. So you guys got a little bit of a conversation last class on the electrical system. So the sinoatrial node is up here in the right atrium, and then conduction goes through.

So the P wave on an EKG represents the conduction through there. You get through the AV node, where there's a delay which is a PR interval, and then you get spreading through the ventricles which is the QRS complex, and then repolarization is the T wave. So that's the electrical system, and of course, these things have to work intimately together.

Every single basic kind of cardiac physiology will show this diagram called the Wiggers diagram which really just shows the interconnectedness of the electrical system. So there's the EKG up there. These are the heart sounds that a provider would listen to with the stethoscope, and this is capturing the flow of sort of the changes in pressure in the heart and in the aorta.

So heart fills during a period of time called diastole. The mitral valve closes. The ventricle contracts. The pressure increases. This is a period of time called systole.

Eventually, something called the aortic valve pops open, and blood goes through the rest of the body. The heart finally starts to relax. The atrioventricular valve closes. Then, you fill again. So this happens again and again and again in a cyclical way, and you have this combination of electrical and mechanical properties.

OK. So I have some pictures here. These are all MRIs. I'm going to talk about echocardiography which is these very ugly, grainy things that I unfortunately have to work with. MRIs are beautiful but very expensive. So there's a reason for that.

So this is something called the long axis view of the heart. So this is the thick walled left ventricle there. This is the left atrium there, and you can see this beautiful turbulent flow of blood in there, and it's flowing from the atrium to the ventricle.

This is another patient's. It's called the short axis view. There is the left ventricle and the right ventricle there. So we're kind of looking at it somewhat obliquely, and then this is another view called the physical. It's a little bit dull there. I'm sorry. We can brighten it a little bit.

This is the what's called the four chamber view. So you can see the left ventricle and right

ventricle here. So the reason for these different views is, ultimately, that people have measures of function and measures of disease that go along with these specific views. So you're going to see them coming back again and again.

OK. So the way that physicians like to organize disease definitions really around some of these same kind of functions. So failures of the heart to pump properly causes a disease called heart failure, and this shows up in terms of being out of breath, having fluid buildup in the belly and in the legs, and this is treated with medications. Sometimes, you can have some artificial devices to help the heart pump, and ultimately, you could even have a transplant, depending on how severe it is. So that's the pump.

Blood supply to the heart ultimately can also be blocked, and that causes a disease called coronary artery disease. If blood is completely blocked, you can get something called a heart attack or myocardial infarction. That's chest pain, sometimes shortness of breath, and we open up those blocked vessels by angioplasty, stick a stent in there, or bypass them altogether.

And then the flow of blood has to be one way. So abnormalities of flow of the blood through valves is valvular disease, and so you can have either two type valves, so that's called stenosis. Or you can have leaky valves. That's called regurgitation. That shows up as light-headedness, shortness of breath, fainting, and then you've got to fix those valves.

And finally, there's abnormalities of rhythm. So something like atrial fibrillation which is a quivering of the atrium, so too slow heartbeats, which would look like cardiac, can present as palpitations, fainting, or even sudden death. And you can stick a pacemaker in there, defibrillator in there, or try to burn off the arrhythmia.

OK. So this is like the very physiology-centric view, but the truth is that the heart has a whole lot of cells. So there's a lot more biology there than simply just thinking about the pumping and the electrical function. Only 30% of the cells or so are these cardiomyocytes. So these are the cells that are involved in contraction.

These are cells that are excitable, but that's only 30% of the cells. There is endothelials in the cell. There's fibroblasts. There's a bunch of blood cells in there too, certainly a lot of red blood cells in there too.

So you have lots of other things. So we're going to come back to here a little bit when talking

about how should we be thinking about disease? The historic way is to think about pumping and electrical activation, but really, there's maybe a little bit more complexity here that needs to be addressed.

OK. So there's a lot of different-- so cardiology is very imaging-centric, and as a result, it's very expensive. Because imaging costs a lot of money to do, and so I have dollar signs here reflecting the sorts of different tests we do. So you saw the cheapest one last week, electrocardiogram, so one dollar sign, and that has lots of utility.

For example, one could diagnose an acute heart attack with that. Echocardiography, which involves sound waves, is ultimately more used for quantifying structure and function, can pick up heart failure, valvular disease, high blood pressure in the lungs. So that's another modality.

MRI, which is just not used all that much in this country, is very expensive. It does largely the same things, and you can imagine, even though it's beautiful, people have not had an easy time and able to justify why it's any better than this slightly cheaper modality. And then you have angiography which can either be by CAT scan or by X-ray. And that visualizes the flow of blood through the heart and looks for blockages which are going to be stented, ballooned up and stented.

And then you had these kind of non-invasive technologies, like PET and SPECT that use radionucleotides, like technetium, rubidium, and they look for abnormalities in blood flow to detect whether or non-invasively there's some patch of the heart that isn't getting enough blood. If you get one of these, and it's abnormal, often, you go over there, and you take a trip to the movies-- as my old teachers used to say-- and then you may find yourself with an angioplasty or stent or bypass. So one of the sad things about cardiology is we don't define our diseases by biology. We define our diseases often related to whether the anatomy of the physiology is abnormal or normal, usually based on some of these images or some of these numbers.

OK. So we have to make decisions, and we often use these very same things too to be able to make some decisions. So we have to decide whether we want to put a defibrillator, and to do so, you often need to get an echocardiogram to look at the pumping function of the heart. If you want to decide on whether somebody needs angioplasty, you have to get an angiogram. If you want to decided to get a valve replacement, you need an echo.

But some of these other ones actually don't involve any imaging, and this is sort of one of the

challenges that I'm going to talk about is that all of the future-- you can imagine building brand new risk models, new classification models. You're stuck with the data that's out there, and the data that's out there is ultimately being collected because somebody feels like it's worth paying for it already. So if you want to build a brand new risk model for who's going to have a myocardial infarction, you're probably not going to have any echocardiograms to be able to use for that, because nobody is going to have paid for that to be collected in the first place.

So this is a problem. To be able to innovate, I've got to keep on coming back to that, because I think you're going to be shocked by the small sample sizes that we face in some of these things. And part of it is because if you just want to piggyback on what insurers are going to be willing to pay for to get your data, you're going to be stuck with only being able to work off the stuff we already know something about. So much of my work has been really trying to think about how we can change that.

OK, so just a little bit more, and then we can get into a little bit more meat. So sort of the universal standard for how imaging data is stored is something called DICOMs, or Digital Imaging and Communications standard, and really, the end of the day, there is some compressed data for the images. There's a DICOM header, which I'll show you in a moment. It's lots of nice Python libraries that are available to be able to work with this data, and there's a free viewer you could use too.

OK. So where do I get access to this? So this has actually been an incredible pain. So hospitals are set up to be clinical operations. They're not set up to make it easy for you to get gobs of data for being able to do machine learning. It's just not really there.

And so sometimes, you have some of these data archives that store this data, but there's lots of reasons for why people make that difficult. And one of them is because often images have these burned in pixels with identifiable information. So you'll have a patient's name emblazoned in the image. You'll have date of birth. You'll have kind of other attributes.

So you're stuck with that, and not only is it a problem that they're there, the vendors don't make it easy to be able to get rid of that information. So you actually have a problem that they don't really make it easy to download in bulk or de-identify this. And part of the reason is because then it would make it easy for you to switch vendors and have somebody else take over.

So they make it a little bit hard for you. Once it's in there, it's hard for you to get it out, and

people are selling their data. That's certainly happening too.

So there's a little bit of attempts to try to control things that way, and many of the labels you want are stored separately. So you want to know what the diseases of these people. So you have the raw imaging data, but all the clinical stuff is somewhere else. So you have to sometimes link that, and so you need to get access there.

And so just to give you a little bit of an idea of scale, so we're about to get all the ECGs from Brigham and Women's which is about 30 million historically, and this is all related to cost. So positron emission tomography, you can get about 8,000 or so, and we're one of the busiest centers for that. Echocardiograms are in the 300,000 to 500,000 range in archives. So that gets a little bit more interesting.

OK. This is what a DICOM header looks like. You have some sort of identifiers, and then you have some information there, attributes of the images, patient name, date of birth, frame rate. These kind of things are there, and there's some variability. So it's never quite easy.

OK. So these different modalities have some different benefits to them which is why they're used for one disease or the other. And so one of the real headaches is that the heart moves. So the chest wall moves, because we breathe, and the heart moves too. So you have to image something that has enough temporal frequency that you're not overwhelmed by the basic movement of the heart itself, and so some of these things aren't great.

So SPECT or PET acquire their images, which are radioactive counts, over minutes. So that's certainly a problem when it comes to something that's moving like that, and if you want to have high resolution. So typically, you have very poor spatial resolution for something that ultimately doesn't deal well with the moving aspect.

So coronary angiography has very, very fast frame rates. So that's X-ray, and that's sort of very fast. Echocardiography can be quite fast. MRI and CT are not quite as good, and so there's some degradation of the image.

As a result, people do something called gating, where they'll take the electrocardiogram, the ECG, and try to line up different portions of different heartbeats. And say, well, we'll take this image from here, line it up with this one from there, this one-- I'm going to talk a little bit about that, about registration, but ultimately, that's a problem that people have to deal with. So it's a computer vision problem of interest.

OK. Preamble is almost done. OK. So why do we even imagine any of this stuff is going to be useful? So it turns out that the practice of interpreting involves a lot of manual measurements. So people like me, and people who have trained for way too long, find themselves getting little rulers and measuring various things.

So for example, this is a narrowing of an artery. So you could take a little bit of calipers and measure across that and compare it to here and say, ah, this is 80% narrowed. You could measure the area of this chamber, the left ventricle, and you can measure its area is, and you can see, ah, its peak area is this. It's minimum area is this. Therefore, it's contracting a certain amount. So we do those things. We measure those things by hand.

And the other thing we do is we actually diagnose things just by looking at them. So this is a disease called cardiac amyloid characterized by some thickening. I'll show you a little bit more about that and some sparkling here.

So people do look and say, ah, this is what this is. So there's kind of a classification problem that comes either at the image or video level. So we'll talk about whether this is even worth doing.

**AUDIENCE:** I have a question.

**RAHUL DEO:** Yes.

**AUDIENCE:** Is this with software, or do you literally take a ruler and measure?

**RAHUL DEO:** So the software involves clicking at one point, stretching something, and clicking another point. So it's a little better than pulling the ruler out of your back pocket, but not that much better. OK. So we're going to talk about or three little areas, and again, this is not-- I got involved in this really in the last two years or so. It's nice of David to ask me to speak here, but I think there are probably people in this room who have a lot more experience in this space.

But the areas that have been relevant to what we've been doing has been image classification and then semantic segmentation. So image classification being assigning a label to an image, very great. Semantic segmentation, assigning each pixel to a class label, and we haven't done anything around the image registration, but there are some interesting problems I've been thinking about there. And that's really mapping different sets of images onto one coordinate system.

OK. So seems obvious that image classification would be something that you would imagine a physician does, and so maybe we can mimic that. Seems like a reasonable thing that happens. So lots of things that radiologists, people who interpret images, do involve terms of recognition, and they're really fast.

So it takes them a couple of minutes to often do things like detect if there's cancer, detect if somebody has pneumonia, detect if there's breast cancer in a mammogram, tells there's fluid in the heart, and then even less than that, one minute often, 30 seconds, they can very, very fast. So you can imagine the wave of excitement around image classification was really post-image net, so maybe about three years, four years, or so ago. We're always a little slow in medicine, so a little bit behind other fields.

And the places that they went were the places where there are huge data sets already, and where there's simple recognition tests. So chest X-rays and mammograms are both places that had a lot of attention, and other places have been slowed down by just how hard it is to get data. So if you can't get a big enough data set, then you're not going to be able to do much.

OK. So David mentioned, you guys already covered very nicely, and this is probably kind of old hat. But I would say that prior to convolutional neural networks, nothing was happening in the image classification space in medicine. It was just not. People weren't even thinking that it was even worth doing.

Now, there's a lot of interest, and so I have many different companies coming and asking for help with some of these things. And so it is now a very attractive thing in terms of thinking, and I think people haven't thought out all that well how we're going to use that. So for example, if it takes a radiologist a minute to two minutes to read something, how much benefit are you going to get to automate it? And the real problem is you can't take that radiologist away.

They're still there, because they're the ones who are on the hook. And they're going to get sued, and it's among the most sued profession in medicine. So there's lots of people who can read an X-ray. You don't need to have all that training.

But if you're the one who's going to be sued, it ends up being that there really isn't any task shifting in medicine. There isn't that kind of, oh, I'm going to let such and such take on 99%, and just tell me when there is a problem. It just doesn't happen, because they ultimately don't

feel comfortable passing that on. So that's something to think about.

So you have a task that's relatively easy for a very, very expensive and skilled person to do, and they refuse to give it up. OK. So that's a problem, but you can imagine there is some scenarios-- and we'll talk more about this-- as to where that could be.

So let's say it's overnight. The radiologist is sleeping comfortably at home, and you have a bunch of studies being done in the emergency room. And you want to figure out, OK, which one should we call them about? So you can imagine there could be triage, because the status quo would be, we'll take them one by one.

Maybe you could imagine sifting through them quickly and then re-prioritizing them. They'll still be looked at. Every single one will still be looked at. It's just the order may change.

So that's an example, and you could imagine there could be separate-- someone else could read at the same time. And we'll come back to this in terms of whether or not you could have two streams and whether or not that is a scenario that would make some sense. And maybe, in resource-poor settings, where we're not teaming with the radiologist, maybe that makes sense too. So we'll come back to that too.

OK. So here's another problem. So almost everything in medicine requires some element of confirmation of a visual finding, and some of the reasons are very simple. So let's say you want to talk about there being a tumor. So if you're going to ask a surgeon to biopsy it, you better tell them where it is. It's not enough to just say, this image has a tumor somewhere on it.

So there is some element of that that you're going to need to be a little bit more detailed than simply making a classification with a level one image, but I would say beyond that. Let's say, I'm going to try to get one of my patients to go for valve surgery. I'll sit with them, bring up their echo, sit side by side with them, and point them to where it is. Bring up a normal one and compare, because I want them to be involved in the decision.

I want them to feel like they're not just trust-- and they have to trust me. At the end of the day, they don't even know that I'm showing-- I'll show them their name, but ultimately, there is some element of trust. They're not able to do this, but at the same time, there is this sense of shared decision making. You're trying to communicate to somebody, whose life is really at risk here, that this is why we're doing this decision. So the more you could imagine that there is

obscuring, the more difficult it is to make that case.

So medicine is this-- I found this review by Bin Yu from Berkeley, just came out, and it talks about this tension between predictive accuracy and descriptive accuracy. So this is of the typical thing we think about that matters, and there's lots of people who've written about this thing. Medicine is tough in that it's very demanding in this space here, and it's almost inflexible in this space here. So it's a tough nut to crack in terms of being able to make some progress, and so we'll talk more about when that's likely to happen.

OK. So this again may be something that's very familiar to you. So we had this problem in terms of some of the disease detection models, and I didn't find this all that satisfying in terms being able to successfully localize. So just digging through the literature, it looks like this idea of being able to explain what part of the image is driving a certain classification. That field is modestly old. Maybe it goes back before that.

But ultimately, there's two broad ways. You can imagine finding an exemplary image that maximally activates the classical work, or you can take a given image and say, what aspect of it is driving the classification? And so in this paper here did both those things. They either went through and optimized-- starting from an average of all the training data-- they optimized the intensities until they maximized the score for a given class. So that's what's shown here.

And then another way to do it is in some sense you could take a derivative of the score function relative to the intensities of all the pixels and come up with something like this. But you could imagine, if you showed this to a patient, they wouldn't be very satisfied. So it's very difficult to make a case that this is super useful, but it seems like this field has progressed somewhat, and I haven't tried this out.

This is a paper by Max Welling and company, out by a couple of years, and maybe you guys are familiar with this. But this ultimately is a little bit of a different approach in the sense that they take patches, the sort of purple-like patch here, and they compare the final score, or class label, relative to what it-- so taking the intensity here and replacing it by a conditional result sampling from the periphery. And just comparing those two things and seeing whether or not you either get activation, which is the red here. This is the way that they did the conditional sampling, and then blue would be the negative contributors.

And there, you can imagine, there's a little bit more distinction here, and then something a little bit more on the medical side is this is a brain MRI. And so depending on this patch size, you

get a different degree of resolution to localizing some areas of the image that are relevant. So this is something that we're going to expect a lot of demands from the medical field in terms of being able to show this. And at least our initial forays weren't very satisfying doing this with what we were doing, but maybe these algorithms have gotten better.

OK. So next thing that matters. OK. So this is what people do. So I did my cardiology fellowship in MGH, and I just traced circles. That's what I did.

I just trace circles, and I stretched a ruler across, and then fed that in. At least the program computed the volumes for me, the areas and volumes, but otherwise, you have to do this yourself. And so this is like a task that's done, and sometimes you may have to-- here's an example of volumes being computed by tracing these sorts of things and much radiology reports just involve doing that. So this seems like a very obvious task we should be able to improve on.

So medicine tends to be not the most creative in terms of trying a bunch of different architectures. So if you look at the papers, they all jump on the U-net as being the favorite architecture for semantic segmentation. So maybe familiar to people here, really, it just captures this encoding or contracting layer.

Where you're downsampling, and then there's a symmetric upsampling that takes place. And then ultimately, there's these skip connections, where you take an image, and then you can catonate it with this upsampled layer, and this helps get a little bit more localization. So we used this for our paper, and we'll talk about this a little bit, and it's very popular within the medical literature.

One of the things that was quite annoying is that what you would find for some of the images, you'd find, let's say, a ventricle. You'd find this nicely segmented area, and then you'd find this little satellite ventricle that the image would just pick. The problem is that this pixel-level classification tends to be a problem, and a human would never make that mistake. But that tends to be something that sounds like it is common in the-- this is a common tension is that this sort of focusing on relatively limited scales ends up being problematic, when it comes to picking up the global architecture.

And so there's lots of different solutions it looks like in the literature. I just highlighted some of these from a paper that was published from Google a little while ago. One of the things that's captured is these ideas of dilated convolutions, and so that you have convolutions built on

convolutions. And so ultimately, you have a much bigger receptive field for this layer, though you haven't really increased the number of parameters that you have to learn.

So there is some. It seems like there's lots. This is not just a problem for us but a problem for many people in this field. So we need to be a little bit more adventurous in terms of trying some of these other methods. We did try a little bit of that and didn't find a gains, but I think, ultimately, there still needs to be a little bit more work there.

OK. So the last thing I'm going to talk about before getting into my work is really this idea of image registration. So I talked about how there are sometimes some techniques that have limitations, either in terms of spatial resolution or temporal resolution. So this is a PET scan here, this sort of reddish glow here, and in the background, we have a CAT scan of the heart. And so clearly, this is a poorly registered image, where you have the PET scan kind of floating out here, when it really should be lined up here. And so you have something that's registered better there.

I also mentioned this problem but gating. So ultimately, if you have an image taken from different cardiac cycles, you're going to have align them in some way. It seems like a very mature problem in computer vision world. We haven't done anything in this space, but ultimately, it has been around for decades. If not, I would just at least touch it, touch upon it.

So this is sort of the old school way, and then now people are starting to use conditional variational autoencoders to be able to learn geometric transformations. This is the Siemens group out in Princeton that has this paper. Again, nothing I'm going to focus on, just wanted to bring it up as being an area that remains of interest. OK. So I think we're doing OK, but you said 4:00.

**PROFESSOR:** 3:55

**RAHUL DEO:** 3:55. OK. All right, and interrupt. Please, interrupt. OK? I'm hoping that I'm not talking too fast. OK.

As David said, this was not my field, but increasingly, there is some interest in terms of getting involved in it, in part because of my frustrations with clinical medicine. So this is one of my frustrations with clinical medicine. So cardiology has not really changed, and one of the things it fails at miserably is picking up early-onset disease. So here's the typical profile, a little facetious.

So people like me in our early 40s, start to already have some problems with some of these numbers. So I like to joke that, since I came back to the Harvard system from California, my blood pressure has gone up 10 points which is true, unfortunately. So these changes already start to happen, and nobody does anything about it.

So you can go to your doctor, and you're also saying, no, I don't want to be on any medicine. They're like, no, no, you shouldn't be on any medicine. So you kind hem and haw, and a decade goes by, 15 years go by. And then finally, you're like, OK, well, it looks like at least my coworkers are on some medicines, or maybe I'll be willing to do that. And so they've got lots of stuff you can be treated, but it is often very difficult, and you see this at the doctor level too. Yes.

**AUDIENCE:** For the optical values, how much personal deviation is there for the values?

**RAHUL DEO:** So the optimal value is fixed and is just like a reference value. And you can be off-- so blood pressure, let's say. So people consider optimal to be less than 120 over less than 80. People are in the 200s.

So you'd be treated in the 200s, but there'll be lots of people in the 140s and the 150s, and there'll be a degree of kind of nihilism about that for some time. And my patients would be like, oh, I got into the fight with the parking attendant. I just had a really bad phone-- there's like countless excuses for why it is that one shouldn't start a medication, and this can go on for a long time. Yes.

**AUDIENCE:** [INAUDIBLE]. How can you assess the risk [INAUDIBLE] for blood pressure? Is that, like, noise [INAUDIBLE]?

**RAHUL DEO:** Yeah. So OK. So that's a great point. So yeah. So the question is that many of the things that we're seeing as risk factors have inherent variability to them. Blood sugar is another great example of those things.

If you could have a single-point estimate that arises in the setting of a single clinic visit, how much do you trust that? So it's a couple of things related to that. So one of them is that people could be sent home with monitors, and they can have 24-hour monitors. In Europe, that's much more often done than here. And then, the thing is that often they'll say that, and then you go look at like six consecutive visits, and they all have something elevate, but it's true.

This is a noisy point estimate, and people have shown that averages tend to do better. But at the same time, if that's all you have-- and the bias is interesting. Because the bias comes from some degree of stress, but we have lots of stress in our life. I hopefully am not the most stressful part of my patient's life, and so I think that ultimately there are-- and the problem with that is it's a good reason for someone to talk you out of them starting them on anything. And that's what ends up happening, and so this can be a really long period of time.

OK. So this is the grim part. OK? So it turns out that once symptoms develop for something like heart failure, decline is fast. So 50% mortality in five years, after somebody gets hospitalized for their first heart failure admission, and often the symptoms are just around that time. So unfortunately, these things tend to be irreversible changes that happen in the background, and largely, you don't really have any symptoms until late in the game.

So we have this problem, where we have this huge stretch. We know that there is risk factors, but we have this huge stretch, where nobody is doing anything about them. And then we have sort things going downhill relatively quickly after that. And unfortunately, I would make a case that probably responsiveness is probably best did this phase over there. Expense is really all over there.

So we really want to find-- and this is what I consider to be missing in medicine. I'm going to come back to this again a little bit later on-- but really, we want to have these-- if you're going to do something in this asymptomatic phase, it better be cheap. You're not going to be getting MRIs every day or every year for people who have no symptoms. The system would bankrupt if you had that. So we need these low cost metrics that can tell us, at an individual level, not just if we had 1,000 people like you, somebody would benefit.

And this is what my patients would say is that they would be so excited about their EKG or their echo being done every year, because they want to know, how does it look like compared to last year? They want some comparison at their level, not just some public health report about this being a benefit to 100 people like you. And so it shouldn't be both low cost, should be reflective at something an individual level, should be relatively specific to the disease process, expressive in some way, and should get better with therapy. I think that's one of the things that's pretty important is if somebody does the things you ask them to do, hopefully, that will look better. And then that would be motivating, and I think that's how people get motivated is that they get responses.

So I would make a case that even simple things like an ultrasound-- and I have one showed here-- really does capture some of these things, and not all those things, but they have some of those things. So you have, for example, that in the setting of high blood pressure, the left ventricular mass starts to thicken, and this is a quantitative, continuous measure. It just thickens over time, and the heart starts to change. The pumping function can get worse over time.

The left atrium, which is this structure over here, this thin-walled structure is amazing in the sense that it's almost this barometer for the pressure in the heart. Oh, that's a horrible reference. OK, but it tends to get kind of bigger and bigger in a very subtle way before any symptoms happen. So you have this, and this is just one view. Right? So this is a simple view acquired from an ultrasound that captures some of these things at an individual level.

So this gets to some of my thoughts around where we could imagine automated interpretation benefiting. So if you want to think about where you're less likely. So with these very, very difficult, end-stage, or complex decisions, where you have a super skilled person even collecting the data in the first place. They've gone through training. They're super experienced.

You have a very expensive piece of hardware used to collect the data. You have an expert interpreting it. This is done late in the disease course. You have to make really hard decisions, and you don't want to mess it up. So probably not good places to try to stick in an automated system in there, but what would be attractive would be to try to enable studies that are not even being done at all.

So move to the primary care setting. Use low cost handhelds. So there's even now companies that are starting to try to automate acquisition of the data by helping people collect it and guide them to collecting the right views. Early in the disease course, no real symptoms here. Decision support just around whether you should start some meds or intensify them, low liability, low cost. So this is a place where we wanted to focus in terms of being able to introduce some kind of innovations in this space.

OK. So this comes back to this slide of I talked about where you could imagine some of these things being low hanging fruit, but maybe those aren't the ones that we should be focusing on we should instead be focusing on enabling more data at low cost, getting more out of the data that we're collecting, and helping people even acquire it in the first place. So that's one

category of things, and that's the one I just highlighted in the previous slide. You can imagine something running in the background at a hospital system level and just checking to see whether there's anybody who was missed in some ways.

And then triage I'm going to talk about in the next slide. I'll come back to that, and then really-- and this is, again, one of the reasons I got into this-- we want to do something that elevates practice beyond just simply repeating what we already do. And so this idea of quantitative tracking of intermediate states, subclasses of disease, which is actually the real reason I got into this space is because I wanted to increase scale of data to be able to do this, and this is where you potentially would like to go.

So the ECG example is an interesting one, because automated systems for ECG interpretation have been around for 40 or 50 years, and they really got going around the early 2000s, when people realized-- there's a pattern called an ST elevation. I'm not sure if you guys talked about that. This is a marker of complete stoppage of blood flow to the heart. So muscle starts to die.

And then the early 2000s, there was a quality movement that said, as soon as anybody sees that, you should get to somebody doing something about it within an hour and a half or so. And so the problem was that in the old days and the old way to do this-- and even this was around the time I was a resident-- you would have to first call the cardiologist. Wake him up. They would come.

You'd send them the image. They would look at it. Then, they would decide whether or not this was the pattern they were seeing, and then they would activate the lab, the cath lab. They would come in, and you were losing about an hour, hour and a half in this process.

And so instead they decided that automated systems could be used to be able to enable ambulance personnel or emergency room docs, so non-cardiologists, to be able to say, hey, look, this is what we think is going on. Let's bring the team in, and so people would get mobilized. People would come to the hospital. Nobody would do anything in terms of starting the case, until somebody confirmed it, but already, the whole wheels were turning.

And so you have this triage system, where you're making a decision. You're not finalizing the decision, but you're speeding things up. And so this is an example where you could imagine it's important to try to offload this to something. So this is an example, and there's going to be false positives.

And people will laugh and mock the emergency room doctors and mock the ambulance drivers and say, ah, they don't know what they're doing. They don't have any experience. But ultimately, people were dying, because they were waiting for the cardiologist to be available to read the ECG. So you've got to think about those in terms of places where there may be cost for delay.

OK. So coming back to echoes. OK. So why does an echo get studied? Because this is probably not something that is typical. It's a compilation of videos, and there are about 70 different videos typically in the studies that we do at the centers that we're at. And they're taken over multiple cycles and multiple different views, and often it takes somebody pretty skilled to acquire those views. And they take about 45 minutes to an hour to gather that data, multiple different views, and the stenographer is changing the depth to zoom in on given structures.

And so you can understand that there's already somebody who was already very experienced in this process even collecting the data which is a problem. Because you need to take them out of the picture, because they're expensive to be able to do those things. So we were doing at UCSF 12,000 to 50,000. Brigham was probably a little busier at 30,000 to 35,000. Medicare back, in 2011, had seven million of these perform, and there's probably hundreds of millions of these archives, so lots of data.

So we published a paper last year trying to automate really all of the main processes around this, and part of the reason to do all is it doesn't help you to have one little bit automated. Because at the end of the day, if you have to have a cardiologist doing everything else and a stenographer doing everything else, what have you really saved by having one little step? So the goal here was to start from raw study, coming straight off the machine, and try to do everything. And so that involves sorting through all these different views, coming up with empirical quality score with it, segmenting all the five primary views that we use. Directly detecting some diseases, and then computing all the standard mass and volume types of measurements that come from this.

So we wanted to do it all, and this was, I think, it wasn't strikingly original in the algorithms that were used. But at the same time, it was very bold for anybody in the community to try to take this on, and of course, in general, all the backlash you could imagine when, you try to do something like this. I still hear it, but there's excitement. And certainly on the industry side, there's really excitement in that this is feasible.

So I was running biology lab, back in 2016 or so, and then decided-- so my cousin's husband is the Dean of Engineering at Penn, and I emailed him and said, do you know anyone at Berkeley? I live near there. I have a very long commute, and I was like closer to there. Is anybody you know there?

So he's like, yeah. I know Ruzena Bajcsy there. She used to be a Penn, and I know Alyosha Efros. And so he just emailed them and said, can you meet? [INAUDIBLE]. And so I met some of them, and then I tried to find some people who were willing to work. So I just spent a day a week there for about two years, just hanging out, writing, code and try to get this project off the ground. So we have a few different institutions.

Jeff Zhang was a senior undergraduate at the time. He's at Illinois right now as a graduate student. It's interesting, because it's hard to get grad student level people excited over stuff that's applications of existing algorithms, but they're happy to advise. So I ended up having to write a lot of the code myself.

And undergraduates are, of course, excited to do these kind of things, because it's better than homework, and I can pay. But I think, ultimately, it's interesting to try to find that sweet spot and also find things that ultimately could be interesting from an algorithmic standpoint too. So I'm trying to do more of that these days. OK.

So we aren't the first to even do something around classifying views. So somebody already had publish something, but we wanted to be a little bit more nuanced than that. In that we wanted to be able to distinguish, for example, whether this structure, the left ventricle, is cut off. Because we don't want to measure it if it's cut off, and we don't want to measure the atrium if it's completely cut off here.

So we wanted to be able to have a classifier able to distinguish between some of those things. It's not an easy task, and a lot of these labels were me riding the train in my very long commute from East Bay, in California, to UCSF. And so I did a lot of labeling, and I did a lot of segmentation too. So I could fly a lot.

And that's the other thing that's kind of interesting is that you often need-- even to do the grunt work-- you may need somebody fairly specialized to do it which is OK, but yeah, so that ended up being me for a lot of this. So I traced a lot of these images, and then I got some other people to help out. But you're not going to get a computer science undergraduate to trace art

structures for you, nor are you going to get them excited about doing this. So we didn't end up having that much data, and I think we could probably get better than that.

But we had the five main views, and we implemented a modified version of unit algorithm. We imposed a bit of a penalty to keep this problem of, for example, a little stray ventricle being out there. We imposed a penalty to say, well, if that's too far away from the center then, we're going to have the loss function take that into account. That helped somewhat, but so that was our approach to-- this is a pretty substantial deal to be able to do all these things that normally would be very tedious.

And as a result, when we start to analyze things, we can segment every single frame of every single video. The typical echo reader will take two frames and trace them. That's it. That's all you get. So we can do everything over every single cardiac cycle, because there's amazing variability from beat to beat. And so it's silly to think that that should be the gold standard, but that is the gold standard.

So we had thousands of echoes. So that's the other thing. So it turns out that it's almost impossible to get access to echoes, so I wrote a keystroke encoder that sat at the front end and just mimicked me entering in studies and downloading them. So that was the only way I could get.

So I had about 30,000 studies built up over a year, but there's no way to do bulk download. And so again, you've got to do some grunt work to be willing to play this space. So we had a fair number of studies we could use in terms of where we had measurements and decent values in terms of that. I think it's interesting in terms of thinking about how good one can-- how close one can get. And one of the things we found is that, when there were big deviations-- these are Bland-Altman plots-- almost always the manual ones were wrong.

**AUDIENCE:** Why is that?

**RAHUL DEO:** Oh, OK. OK. So Bland-Altman plots, so people don't like using correlations in the medical-- so Bland and Altman published a paper in the *Lancet* about 30 years ago complaining that correlations and correlation coefficient are ultimately not good metrics. Because you could have some substantial bias, and really you want to know, if this is the gold standard, you need to get that value. So it really is just looking at differences between, let's say, the reference value and the, let's say, automated value, and then plotting that against the mean of the two. So that's it.

I did it as percentages here, but ultimately, it's just that. It's that you're just taking the mean of, let's, say the left ventricular volume. You have a mean of the automated versus the manually measured one, and then you compare what the difference is of one minus the other, and so you'll be on one side or the other. So ideally, you would just be sitting perfectly on this line, and then you're going to look and see whether or not you're clustered on one side or the other. So that's just the typical thing.

People try to avoid correlation coefficients, because they kind of consider them to be not really telling you whether or not-- there really is a gold standard, and there truly is a value here, and you want to be near that value. And so that's the standard for looking at comparison of diagnostics. So we had about 8,000 things. The reviewers gave us a hard time for the space up here, and there are not that many studies up here, but ultimately, there are some.

And when we manually looked at a bunch of them, always the manual ones were just wrong. Either there is a typo or something like that, so that was reassuring, but we were sometimes very wrong. And you'd find that the places we'd be wrong would be these ridiculously complex congenital heart studies that we had never been given examples like that before.

So that's a lesson to be learned is that, sometimes, you're going to be really off in these sorts of approaches, and you have to think a little bit. And what we ended up doing is having an interative cycle, where we would identify those and feed them back and of keep on doing that, but that still needs to be improved upon.

OK. So function, again, there's, a couple of measures a function. There's a company that has something out there in this space, got FDA approved for having an automated ejection fraction. So I think we're better than their numbers, overall, but yeah. I think that that's just one of those things you're expected to be able to do.

And then here's a problem that we run into. So we're comparing to the status quo which, like I said, is one person tracing two images and comparing them. That's it. So we're processing potentially 200, 300 different frames per study and competing median, smoothing across. We're doing a whole lot more than that.

So what do we do about that in terms of the gold standard? And if you just take into observer variability, you're going to have up to 8% to 9% in absolute compared to 60% of the reference. So that's horrible.

So what are you supposed to do? And I think so one thing people do is they take multiple readers and ask them to do that. But this is like, are you're going to get a bunch of cardiologists to do like 1,000 studies for you? It's very hard to imagine somebody doing that.

You could compare it to another modality. So we haven't done this yet, but you could, for example, compare it to MRI and say whether or not you're more consistent with another modality. And then this is indirect, but you can go to like outcomes in a trial and see whether or not you do a better job. So there are things you can do.

One of the things we decided to do is look for correlations of structures within a study itself and say, well, the mass-- so we know that, for example, thickened hearts lead to larger increases of pressure and left atrial enlargement. So we can look for correlations between those things and see whether we do a better job.

I'd say, for, the most part we're about on par with everything that's there. So I don't think we're any better. Sometimes we're better. Sometimes we're worse.

And I think, for the most part, this was another way to try to get at this, because we were stuck with this. How do you work with a gold standard that ultimately I don't think anybody really trusts as a gold standard? And this is a problem that just has to keep on coming up. This is just an example of where you could facilitate this idea of low cost serial imaging and point of care.

So these are patients who are getting chemotherapy, and so so Herceptin-- not herception, Herceptin, it's like inception-- is an EGFR inhibitor that causes cardiac toxicity, and so people are getting screening echoes. So you could imagine, if you make it easier to acquire and interpret that, all you want to care about is the function and the size. So you can imagine automating that.

So we just did this as proof of concept that you could imagine doing something like this. And for the last thing I want to talk about-- or sorry, the last thing in this space-- is that you could also imagine directly detecting disease. And so you have to say, well, why is that even worthwhile? Yes.

**AUDIENCE:** I was curious. I guess it's going back to the idea of if you look at blended models between human groud truth and maybe a biological ground truth, [INAUDIBLE] versus sort of what you could get from an MRI or something-- or maybe not necessarily an MRI, but what you were

saying based on the underlying biology, or if those two things are generally kept separate?

**RAHUL DEO:** Yeah. These are early days for a lot of this, and I think, anytime you make anything more complicated, then the readers will give you a hard time, but you can imagine that. And especially, you may want to tune things to be able to be closer to something like that. So yeah, I think, unfortunately, people are pretty conservative in terms of how they interpret, but it does make some sense that there's probably something that--

Ideally, you want to be able to have something that is useful, and useful may not be exactly the same thing as mimicking what humans are doing. So no, I think it's a good idea. And I think that this is going to be-- this next wave-- is going to be thinking a little bit more about that in terms of like how do we improve on what's going on over there, rather than simply dragging it back to that?

OK. So there are multiple rare diseases. I use to have a clinic that would focus on these, and they tend to get missed at centers that don't see them that often. So one place you could imagine is you can focus on trying to pick those up, and you could imagine, this could be just surveillance running in the background. It doesn't have to be kind of real time identification.

So there's a few diseases where it's very reasonable to do these things, where it's very obvious. So this is a disease called hypertrophic cardiomyopathy. I used to see it in my clinic. So abnormally thickened hearts, leading cause of sudden death in young athletes.

So Reggie Lewis, there's a bunch of people who've died suddenly from this condition. Unstable heart rhythm, sudden death, heart failure, it runs in families, and there are things you can do, if you identified it. And so it's actually a fairly easy task, in the sense that it tends to be quite obvious. So we built the classification model around this, and we tried to understand what it was doing in part.

And so we tried to do some of these kind of attention or saliency type things, and they were very unsatisfying, in part because I think there's so many different features across the whole image. So you're just getting this blob, but I think maybe we just weren't implementing it correctly. I'm not really sure, but you have a left atrium gets bigger. The heart gets thicker.

There's so many changes across the image. It was unsatisfying in terms of that. So we did something simple and just took the output of the probabilities and compared it to some simple things that we actually know about these things and found that there was some degree of

correlation. But I would like to make that a little bit better.

Cardiac amyloid, a very popular disease for which there are now therapies. And so pharma is very interested in identifying these people, and they really get missed at a pretty high rate. So we built another model for this. Usually, we had about 250 or 300 cases for each of these things and maybe a few thousand controls.

And then this one's a little interesting. This is mitral valve prolapse. So this is what a prolapsing valve looks like. If you imagine the plane of the valve here, it buckles back.

So it does this, and that's abnormal, and this is a normal valve. So you notice, it doesn't buckle back in. So it's a little interesting in that there's really only one part of the cardiac cycle that would really highlight this abnormality, at least that's the way that-- so the way that it's read clinically is people wait for this one part of the cardiac cycle where it's buckled back. They draw an imaginary line across, and they measure what the displacement is there, and so we built a reasonable model focusing. So we phased these images and picked the part of the cardiac cycle, those relevant, all in an automated way and built a model around that and pretty good, in terms of being able to do that, in terms of being in detect that. Yes.

**AUDIENCE:** And so is this model on images at a certain time? Like can you just go back? Because obviously, you weren't doing videos. Right?

**RAHUL DEO:** Well, so we would take the whole video. We were segmenting it. We were phasing it, figuring out what the part of the-- when was the end systole in that, and then using those as the-- so using a stack of those to be able to classify.

**AUDIENCE:** So how do you know the time point?

**RAHUL DEO:** Well, that's I'm saying. So we we're using the variation in the volumes.

**AUDIENCE:** The segmentation would allow you to know the time point.

**RAHUL DEO:** Exactly, because so a typical echo will have an ECG to use to gate, but the handhelds don't. So we want to move away from the things that involve the fanciness and all the bells and whistles. We're trying to use the image alone to be able to tell the cardiac cycle. So that's how we did it. Yes.

**AUDIENCE:** So you mentioned handhelds. With the ultrasounds [INAUDIBLE], are they different from

these?

**RAHUL DEO:** They look pretty similar. We got some now, and they look pretty similar in terms of the quality of the images, and you can acquire the very same view. So I think we haven't shown that we can do it off those, in part because there just isn't enough training data. But they look pretty nice, and I know at UCSF and at Brigham, all the fellows are using it.

It looks pretty much the same in terms of the-- the transducers are similar, and image quality is very good. Resolution is very good. Frame rate probably doesn't get up as high necessarily, but for the most part, I don't think it's that different. So that is the next phase. Yes.

**AUDIENCE:** Could you comment on-- so you mentioned how each of these three examples could be used within a surveillance algorithm.

**RAHUL DEO:** Yeah.

**AUDIENCE:** Could you comment on where along this true positive, false positive trade-off you would actually be realistic to use this?

**RAHUL DEO:** Yeah. That's a good point. I think it would vary for every single one of those, and you really want to have some costs on what the-- so I would typically err on the side of higher sensitivity and dump it on the cardiologists to be able to-- so I would work, but I think you have to pick some-- let's say, you're a product manager.

**AUDIENCE:** Just choose one of these three, and maybe--

**RAHUL DEO:** OK. Yeah. So this is a pretty rare disease. So your priors are pretty low in terms of these individuals. And so I think you probably would probably want to err somewhere along this area here, and so just working on what the-- so you probably will still be a relatively high rate of false positives even that space. But I would argue that it would take the treating cardiologist potentially just a few minutes to look at that study again, and if you picked up one of those patients, that would be a big win. So I think that the cost probably wouldn't be that high, and you just have to make the case.

So therapy for amyloid, for example, this is a nice sharp up stroke there. There's new drugs out there that are sort of begging for patients, and they're having a real hard time identifying them. So you could imagine again, it's sort of a calculus based on what the benefits would be for that identification and what burden you're placing on the individuals to have to over read

something. And you could probably tune that depending on what the disease is and who you're pitching it to. But you're right, you're going to crush people if like 1 in 100 ends up taking a true positive then you're not going to get many fans. Yes.

**AUDIENCE:** Could you comment on whether, for example, [INAUDIBLE] basis, the ones that you're able to predict very well at that point you just chose what distinguishes the ones that are defined well?

**RAHUL DEO:** So that's a good point, and I don't really know in the sense that I haven't looked that closely. But I'm going to guess, they're very thick and very obvious in that sort of sense. So we have a ECG model that may pick this up early. What you want is something to fix it up when it's treatable, not having something that's ridiculously exaggerated. So you may need multiple modalities some of which are more sensitive than others that can catch earlier stage disease to be able to do that. So there are interesting things about this disease in particular.

So cataracts sometimes happen before-- so ideally, the way you do this is-- and I'm actually consulting around something like this-- you ideally want a mixture of electronic health record, something from other findings-- mirror findings, eye findings, plus maybe something cardiac plus and have something that ideally catches the disease in the ideal most treated state. And maybe echo's not the best one, and I think that we'll come back to that at the end. We have a little bit of time.

OK. So UCSF is filing-- I don't know. I don't think this is actually patentable, but they are filing for a patent. I'm just filling the paperwork out today in terms of-- I don't know. But my code is all freely available anyway, for academic, non-profit use, and they're just trying to make it better.

I think, ultimately, my view as an academic here is to try to show what's possible. And then, if you want to get a commercial product, then you need people to weigh in on the industry side and make something pretty and make it usable and all that. But I think, ultimately, I'm trying to just show, hey, if we could do this in a scalable way and find out something new, then you guys can catch up and do something that ultimately can be deployed.

And what's interesting is I have a collaborator in New Zealand. There, they're are resource poor. So they have a huge backlog of patients. They don't have enough stenographers, and they don't have enough cardiologists.

So they're trying to implement this super ultra quick five-minute study and then have

automation. And so they want our accuracy to be a little bit better, but I think they're ready to roll out, if we're able to get something that has probably more training data. Yes. Are you from New Zealand?

**AUDIENCE:** No. I think you started talking about the trade-off between accuracy and-- so in academia, I get the sense that they're always chasing perfect accuracy.

**RAHUL DEO:** Yeah.

**AUDIENCE:** But as you said, you're not going to get rid of cardiologists in the diagnosis. So I have a philosophical question of are you chasing the wrong thing? Should we chase perfect accuracy?

**RAHUL DEO:** Yeah. So the question is around what should our goals be? So should we be just chasing after a level of accuracy that may be either very, very difficult to attain? And especially, if there's never a scenario where there'll be no clinician involved, should we instead be thinking about something that gets good enough to that next step? And I think that's a really good point.

And what's interesting is-- and also it's interesting from the industry side-- is the field starts with the mimicking mode, because it's much harder to change practice. It's much easier to just pop something in and say, hey, I know you have to make these measurements. Let me make them for you, and you could look at them and see if you agree. So that's what ECGs do. Right?

So nobody these days is measuring the QR rests width. Nobody does that. That's just not done. If you've got a number that's absurd, you'll change it. But for the most part, you're like, it's close enough, but you almost have to start with that.

To do something that's transformative is very hard to do. So I think something that involves-- and I talked to David about this. It's sort of like the man-machine interface is fascinating to think about how do we together come up with something better? But it's just much harder to get that adopted, because it requires buy-in in a way that's different than just you do my work for me, but more that we come together to do something better. And I think that's going to be interesting as to how to chip away at that problem.

OK. So a couple of musings, then I'm going to talk a little bit about One Brave Idea, if we have time, or I can stop and take questions instead, because it's a little bit of a biology venture. OK. So I do think that we should really look. People give me a hard time around echo, and I'm like,

well, ECG's been around for a long time, and there's automation there. So let's think about how it's used there, and then see whether or not-- it's not as outlandish as people think.

So I think a lot of these routine measurements are just going to be done in an automated way. Already in our software, you can put out a little picture and overlay the segmentation on the original image and say how good it looks. So that's easy. So you can do that.

And then this kind of idea of point of care automated diagnoses can make some sense around some emergency-type situations. So maybe you need a quick check of function. Maybe you want to know if they have a lot of fluid around the heart, and you don't necessarily want to wait. So those will be the places where there may be some kind of innovations around just getting something done quickly.

And then you always have somebody checking in the background, layer on, a little the heart attack thing I showed you, and I think this problem in echo is there. And so if you need skilled people to be able to acquire the data in the first place, you're stuck, because they can read an echo. A really good stenography can read the whole study for you.

So if you already have that person involved in the pipeline, then it's really hard to introduce a big advance. So you need to figure out how to take a primary care doc off the street, put a machine in their hand, and let them get the image and then automate all the interpretation for them. And so until you can task shift into that space, you're stuck with having still too high a level of skill.

So there are these companies that are in the space now, and there's a few that are trying. It's easy to imagine, if you can train a neural network to classify a view, you could get it to-- this gets to this idea of registration a little bit-- you can recognize if you're off by 10 degrees, or if you need a translation. You could just train a model to be able to do that. So I think that's already happening right now. So it's a question as to whether that will get adopted or not, but I think that, ultimately, if you want to get shifting towards sort of less skilled personnel, you need to do something in that space.

OK. So this is where it gets a little bit harder is to think about how to make stuff and elevate medicine beyond what we're doing. And this gets back to this problem I mentioned is, at the end of the day, you can't find new uses for echo, unless the data is already there for you to be able to show that there's more value than there currently is, sort of this chicken and egg thing.

So in some sense, what I hope to introduce in some way that we can get much bigger data sets, and they don't have to be 100 video data sets. They can be three video data sets, but we want to be able to figure out how to enable more and more of these studies.

So then you can sort of imagine learning many more complicated things. You want to track people over time. You want to look at treatment responses. So you've got to look at where the money is already and see who could do this.

So pharma companies are interested, because they have these phase II trials. They may only have three months or six months to show some benefit for a drug, and they're really interested in seeing whether there's differences after a month, two months, three months, four months. So that may be a place where you get-- and they're being frugal, but they have money.

So you could imagine, if you could introduce this pipeline in there and just have handheld, simple, quick to acquire, far more frequency, and you show a treatment response, and that's kind of transformative then. Because then, you could imagine, that can get rolled out in practice after that. So you need somebody to bankroll this to start with, and then you could imagine, once you have a use case, then you could imagine it getting much more. And this idea of surveillance, you could imagine that would be very doable, that you could just have something taking--

The problem is, you can even get the data in the archives anyway, but let's say you can get that. You could just have this system looking for amyloid, looking for whatever, and that would be a win too is to be able to imagine doing something like that. It's not putting any pressure on the clinical workflow. It's not making anybody look bad. I think, ultimately, it's trying to just figure out if-- well, maybe somebody may be looking bad if they miss something, but yeah. I think it is just trying to identify individuals.

And so this is an area I think that's hard, and so this kind of idea, this is where I started a little bit, around this kind of idea of this disease subclassification and risk models. And so that's like more sophisticated than anything we're doing. I think we're pretty crude at this kind of stuff, but one of the challenges is people just aren't interested in new categories or new risk models, if they don't have some way that they can change practice. And that becomes more difficult, because then you need to not only introduce the model, you need to show how incorporating that model in some way is able to either identify people who respond.

It always comes down to therapies at the end of the day. So can you tell me some subclass of

people who will do better on this drug, which means that you have to have trial data that has all those people with all that data. And unfortunately, because echoes are so expensive and places like the Brigham charge like $3,000 per echo, then you only have like 100 people who have an echo in a trial or 300 people have an echo.

You have a 5,000 person trial, and 5% of them have an echo. So you need to change the way that gets done, because you're massively underpowered to be able to detect anything that's sort of a subgroup within that kind of work. So yeah, unfortunately, the research pace of things outpaces the change in practice in terms of the space, until we're able to enable more data collection.

So I can stop there. I was going to talk about blood cells in slides.

**PROFESSOR:** We can take some questions.

**RAHUL DEO:** Yeah. Yeah. Yeah. OK. Why don't we do that. Yes.

**AUDIENCE:** When CT reconstruction started, I remember seeing some papers where people said, well, we know roughly what to the anatomy should look like, and so we can fill in missing details. In those days, the slices were run before, and so they would hallucinate what the structure looked like.

**RAHUL DEO:** Yeah.

**AUDIENCE:** And of course, that has the benefit of giving you a better model, but it also does risk that it's hallucinated data. Have you guys tried doing that with some of the--

**RAHUL DEO:** Yeah. That's a great point. So OK. So the question was so cardiac imaging has a very long history, and so there was a period of time where there's these kind of active modelers around morphologies of the heart. And so people had these models around what the heart should look like from many, many, many studies. And they were using that, back at the time, when you had these relatively coarse multi-slice scanners for a CT, they would reconstruct the 3D image of the heart based on some pre-existing geometric model for what the heart should look like.

And there's, of course, a benefit to that, but some risk in the sense that somebody may be very different in the space that's missing. And so the question is whether those kind of priors can be introduced in some way, and it hasn't been straightforward as to how to do that.

Whenever you look at these ridiculously poor segmentations, you're like, this is idiotic. We should be able to introduce some of that, and I've seen people, for example, put an autoencoder.

That's not exactly getting at it, but it's actually getting it somewhat with these coarser features. But no, I think in terms of using some degree of geometric priors, I think I may have seen some literature in that space. We haven't tried anything there. We don't have any data to do that, unfortunately, and I suspect, yeah, I just don't know how difficult that is.

**AUDIENCE:** You mentioned that you don't want to see a small additional atrium off at a distance. So that's, in a way, building in knowledge.

**RAHUL DEO:** Yeah. No. I remember when I was starting this space. I was like this is idiotic. Why can't we do this? Why don't we have some way of doing that?

We couldn't find at that time any architectures that were straightforward to be able to do that, but I'm sure there is something in that space. And we didn't also have the data for those priors ourselves. There's a long history of these de novo heart modelers that exist out there from Oxford and the New Zealand group for that matter who've been doing some of this kind of multi-scale modeling. It will be interesting to see whether or not there is anybody who pushes forward in that space, or is it just more data? I think that's always that tension.

**AUDIENCE:** Can I ask about ultrasounds?

**RAHUL DEO:** Yeah.

**AUDIENCE:** You didn't show us ultrasounds. Right?

**RAHUL DEO:** Yeah, I did.

**AUDIENCE:** Oh, you did?

**RAHUL DEO:** Yeah. The echoes are ultrasounds.

**AUDIENCE:** Oh, OK, but that's really expensive ultrasound. Right? Like there are cheaper ultrasounds that you could imagine that you constantly do. Right?

**RAHUL DEO:** Yeah. So there is a company that just came out with the $2,000 handheld ultrasound, the subscription model. Yeah. So I think that Philips has a handheld device around the $8,000

marker, so $2,000 is getting quite cheap. So that's I think the space for handheld devices.

**AUDIENCE:** We're talking about resource-poor countries.

**RAHUL DEO:** Yeah.

**AUDIENCE:** In a developing country, where maybe they have very few doctors per population kind of thing. What kind of imaging might be useful that we could then apply computer vision algorithms to?

**RAHUL DEO:** I think ultrasound is that sweet spot. It has versatility, and its cost is about where-- and I'm sure those companies rented it out for much lower cost in those kinds of places too. We're putting together-- or I put together-- actually, it may not have been funded. I'm not sure. But looking at sub-Saharan Africa and collaborating with one of the Brigham doctors who travels out to sub-Saharan Africa and looking to try to build some of these automated detection type of things in that space.

So no, I think there is definite interest in that, and then there may be a much bigger win there then the stuff I'm proposing. But yeah, no, I think that's a very good point, and that would be-- it's also, it's portable. You could have a phone-based thing. So it's actually very attractive from that standpoint.

**PROFESSOR:** [INAUDIBLE]

**RAHUL DEO:** All right. I feel like I'm changing the topic substantially but not totally. OK. So this is that slide I showed, and I pitched it in a way to try to motivate you to think of ultrasound. But I'm not sure ultrasound really achieves all these things, in the sense I wouldn't call it the greatest biological tool to get at underlying disease pathways. Some of these things may be late, like David said, or maybe not so reversible.

So we've been given this One Brave Idea thing $85 million now to make some dent in a specific disease, so coronary artery disease or coronary heart disease. It's that arrogant tech thing, where you just dump a lot of money somewhere and think you're going to solve all problems. And happy to take it, but I think that there are some problems. So this is what I wanted to do, so I've wanted to do this for probably the last five, six years, before I even started here, and this has motivated me in part for quite a while.

And so here's our problems. OK. So we're studying heart disease, so coronary artery disease or coronary heart disease is the arteries in the heart. You can't get at those. So you can't do

any biology. You can't do the stuff the cancer people-- do you can biopsy that. You can't do anything there. So you're stuck with the thing that you want to get at is inaccessible.

I talked about how a lot of the imaging is expensive, but all those other omic stuff is really expensive too. So that's going to be not so possible, and you're not going to be able to do serial $1,000 proteomics on people either. That's not happening anytime soon. And then everything I talked about, we were woefully inadequate in terms of sample size, especially if we want to characterize underlying complex biological processes. So we expect we're going to need high dimensional data, and we're going to need huge sample sizes. There's Vladimir Vapnik over there.

And then here's another problem. OK? So this stuff takes time. These diseases take time. So if I introduce a new assay right now, how am I going to show that any of this is going to be beneficial? Because this disease develops or 10 to 20 years. So I'm not going to talk about the solution to that, well, a little bit.

OK. So one of the issues with a lot of the data that's out there is it's not particularly expressive. It's a lot of that just the same clinical stuff, the same imaging stuff. So all these big studies, these billion dollar big studies, ultimately just have echoes and MRIs and maybe a little bit of genetics, but they really don't have stuff that is this low cost expressive biological stuff that we ideally want to be able to do. So this is really expensive and makes $85 million look like a joke, and it's not all that rich in terms of complexity.

So we wanted to do something different, and so this is the crazy thing. We're focusing on circulating cells, and so this is a compromise. And there's a reasonably good case to be made for their involvement. So there's lots of data to suggest that these are causal mediators of coronary artery disease or coronary heart disease. So you can find them in the plaques.

So patients who have autoimmune diseases certainly have accelerated forms after atherosclerosis. There are drugs. There's a drug called canakinumab that inhibits IL-1 one beta secretion from macrophages, and this has mortality benefit in coronary artery disease. There are mutations in the white blood cell population themselves that are associated with early heart attack.

So there's a lot there, and this has been going-- and there's plenty of mouse models that show that if you make mutations only in the white blood cell compartment, that you will completely change that the disease course itself. So there's a good amount of data out there to suggest

that there is an informative kind of cell type there. It's accessible. There's lots of predictive models already there that could be done with some of this, and they express many of the genes that are involved. And there's a window on many of these biological processes.

So we're focusing on computer vision approaches to this data. So we decided, if we can't do the omic stuff, because it costs too much, we're going to take slides and have tens of thousands of cells per individual. And then we can introduce fluorescent dyes that can focus on lots of different organelles. And then we can potentially expand the phenotypic space by adding all kinds of perturbations that can be able to unmask attributes of people that may not even be relatively there at baseline.

And I think I've been empowered by the computer vision experience with the echo stuff, and I'm like, hey, I can do this. I can train these models. So we're in a position now where we can-- this stuff costs a few dollars per person. It's cheap, and you can just keep on expanding phenotypic space. You can bring in drugs. You can bring in whatever you want here, and you're still in that dollars type range.

So we just piggy-back, and we just hover around-- just a couple of research assistants were hovering around clinics. And we can do thousands of patients a month, so tens of thousands of patients a year. So we can get into a deep learning sample size here, and so we want these primary assays to be low cost, reproducible, expressive, ideally responsive to therapy. So that's this space here, and there's lots of stuff that we have.

We have all the medical record data on all these people, and we can selectively do somatic sequencing. We can do genome associations. We have all ECG data. We have selective positron emission data.

So it's lots of additional thought, and we want to be able to walk our cheap assay towards those things are more expensive but for which there's much more historical data. So that's what I do with my life these days, and the time problem has been solved. Because we found a collaborary MGH who has 3 1/2 million of these records in terms of cell counting and cytometer data going back for about three years. So we should be able to get some decent events in that time.

I need to build a document classification model for 3 1/2 million records and decide whether they have coronary heart disease, but sounds like that's doable. We're fearless in this space. And then they also have 13 million images, so hundreds of thousands of people worth of

slides. So we can at the very least, get decent weights for transfer learning from some of this data, and we're doing this for acute heart attack patients.

So yeah, so this is what I'm doing, ultimately, and so it's this bridge between existing imaging, existing conventional medical data, and this low cost, expressive, serial-type of stuff that ultimately hoping to expand phenotypic space and keep the cost down. I think all my lessons from working with expensive imaging data has motivated me to build something around this space. So this is my it's my baby right now. And so lots of things for people to be involved in, if they want to, and these are some of the funding sources. All right. Thank you.

[APPLAUSE]