UNITED STATES OF AMERICA

DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOOD AND DRUG ADMINISTRATION

+ + +

CENTER FOR DEVICES AND RADIOLOGICAL HEALTH

+ + +

PUBLIC WORKSHOP - EVOLVING ROLE OF ARTIFICIAL INTELLIGENCE IN RADIOLOGICAL IMAGING

+ + +

February 26, 2020
9:00 a.m.

Natcher Conference Center, Building 45
Ruth Kirschstein Auditorium
NIH Main Campus
National Institutes of Health
9000 Rockville Pike
Bethesda, MD

Free State Reporting, Inc.
1378 Cape Saint Claire Road
Annapolis, MD 21409
(410) 974-0947

WORKSHOP PARTICIPANTS:

MARJAN NABILI, Ph.D.
Biomedical Engineer, Division of Radiological Health
Office of Device Evaluation and Quality
CDRH/FDA

ANTHONY SAMIR, M.D.
Director of CURT and Service Chief, Body Ultrasound
Department of Radiology
Massachusetts General Hospital

RANDY KING, Ph.D.
Program Director, Ultrasound, Division of Applied Science and Technology
National Institute of Biomedical Imaging and Bioengineering (NIBIB)

BEHROUZ SHABESTARI, Ph.D.
Acting Director, Division of Health Informatics Technologies
Division of Applied Science and Technology (Bioimaging)
National Institute of Biomedical Imaging and Bioengineering (NIBIB)

SHAHRAM VAEZY, Ph.D.
Biomedical Engineer, Division of Radiological Health
Office of Device Evaluation and Quality
CDRH/FDA

HA HONG, Ph.D.
Investigator/Co-founder
Caption Health, Inc.

BENNY LAM, Ph.D.
Principal Regulatory Affairs Specialist, AI/DH Regulatory Intelligence
Philips Ultrasound

JOHN MARTIN, M.D., M.B.A., FACS
Chief Medical Officer
Butterfly Network

MIKE WASHBURN
Chief Engineer
GE Healthcare - Ultrasound

BRIAN GARRA, M.D.
Physician, Division of Imaging Diagnostics and Software Reliability
Office of Science and Engineering Laboratories
CDRH/FDA

HELEN FELTOVICH, M.D., M.S.
Attending Physician, Intermountain Healthcare
Associate Research Scientist, Medical Physics
University of Wisconsin, Madison

JEAN LEA SPITZ, M.P.H., CAE, RDMS
Executive Director
Perinatal Quality Foundation

JOSHUA BASILE
Founder
Determined2Heal and SPINALpedia

HARSH THAKKAR
Program Coordinator
MedStar NRH
President
United Spinal Metro D.C. Chapter

ISABELLA O'BRIEN and AMY O'BRIEN
Youth Member and Parent
International Children's Advisory Network (iCAN)

ROB TRAHMS
Artificial Intelligence Leader
Philips Healthcare

RICHARD FRANK, M.D., Ph.D.
Chief Medical Officer
Siemens Healthineers

JESSICA LAMB, Ph.D.
Assistant Director (Acting), Mammography, Ultrasound, and Imaging Software Team
Division of Radiological Health
Office of Product Evaluation and Quality
CDRH/FDA

GARRETT ASTARY, Ph.D.
Biomedical Engineer, Division of Radiological Health
Office of Device Evaluation and Quality
CDRH/FDA

SCOTT PAULSON
Senior Director, Regulatory Affairs and Quality Systems/Compliance Officer
FUJIFILM SonoSite

BERKMAN SAHINER, Ph.D.
Physicist, Division of Imaging Diagnostics and Software Reliability
Office of Science and Engineering Laboratories
CDRH/FDA

XIN FENG, Ph.D.
Human Factors and Reliability Engineering Team
Drug Delivery and General Hospital Devices and Human Factors
Office of Device Evaluation and Quality
CDRH/FDA

ANDY MILKOWSKI
Senior Director, Ultrasound R&D
Siemens Healthineers

SAM SURETTE
Caption Health, Inc.

TONY RODER
Regulatory Affairs Executive
GE Healthcare

PUBLIC PRESENTERS:

DAVID RITSCHER
Senior Consultant
Cambridge Consultants

AALPEN A. PATEL, M.D., FSIR
Chair, Department of Radiology
Medical Director, Artificial Intelligence
Vice Chair, Imaging Informatics
Medical Director, 3-D Imaging
Geisinger

ANANTH ANNAPRAGADA, Ph.D.
Professor/Director, Basic Research in Radiology
The E.B. Singleton Department of Pediatric Radiology
Texas Children's Hospital and Baylor College of Medicine

HERSH SAGREIYA, M.D.
Assistant Professor of Radiology
University of Pennsylvania

ENHAO GONG, Ph.D.
Founder/CEO
Subtle Medical, Inc.

DAVID B. LARSON, M.D., M.B.A. (On behalf of Greg Zaharchuk, M.D., Ph.D.)
Associate Professor of Pediatric Radiology
Vice Chair, Education and Clinical Operations
Department of Radiology
Stanford University School of Medicine

SHANDONG WU, Ph.D.
Director, Intelligent Computing for Clinical Imaging (ICCI) Lab
Department of Radiology
University of Pittsburgh

ARMEN A. VIDIAN
Partner
DCVC

BRIAN SCARPELLI
Senior Global Policy Counsel
Connected Health Initiative

HOUSTON BAKER, Ph.D.
Imaging Technology Development Branch
National Cancer Institute

GEORGE REDMOND, M.S., M.B.A.
Imaging Technology Development Branch
National Cancer Institute

INDEX

INDEX

M E E T I N G

(8:00 a.m.)

DR. NABILI: Good morning, everyone and welcome to Day 2 of FDA public workshop, Evolving Role of Artificial Intelligence in Radiological Imaging. My name is Marjan Nabili, I'm a Lead Reviewer in the Division of Radiological Health in Center for Radiological Health at the FDA. Today's workshop focuses mostly on AI-guided image acquisition, targeting ultrasound applications.

This is a disclaimer.

In this presentation, I would like to discuss, very generally, why ultrasound could be the candidate for AI-guided image acquisition, what are the challenges, what the -- future looks like, and what could be the benefit-risk of using ultrasound in AI-guided image acquisition in a home environment.

What are the advantages of ultrasound compared to other imaging modalities? Ultrasound is one of the core diagnostic imaging modalities which is used routinely in clinical practice to evaluate internal body structures. Some of the advantages are there is no ionizing radiation involved and it is low cost compared to other modalities. Portability is one of the great advantages; real-time imaging, and the fact that the user is able to receive real-time feedback makes it a good tool for guiding different procedures, especially in minimally invasive procedures such as needle biopsies. Ultrasound imaging could cover a range of clinical applications such as fetal monitoring, cardiac applications, bone density, and so on. These advantages could make ultrasound an attractive modality for AI/ML space, both for image acquisition and post processing.

There are, of course, challenges associated with using ultrasound. Image quality can be greatly influenced by how the operator uses a handheld probe. There is variability across different manufacturers. There are many presets which could be confusing for some

of the new users. The presence of noise and artifacts could also be challenging during acquisition. Machine learning could promise to play an important role in addressing some of these challenges by helping with automatic ultrasound data analysis such as guidance, segmentation, and so on.

You have seen AI/ML applications in everyday life and this technology made its way through healthcare and put advanced medicine into a new realm. What could be the benefits of using AI-guided image acquisition? Reproducibility, procedures are more standardized, and it could improve efficiency and reliability. Clinical AI applications may assist in acquisition of a standardized image independent of the operator, guiding both sonographers and non-experts to acquire images with diagnostic quality.

Recently the FDA authorized marketing of software to assist medical professionals in the acquisition of cardiac ultrasound images. We have seen development of AI in other areas of ultrasound imaging, for example, auto-ejection fraction, in which the ultrasound machine automatically tracks the endocardial border and calculates the ejection fraction of the left ventricle. Auto-ejection fraction reduces the interaction from the user, which not only speeds up examinations, but also improves reproducibility, particularly for less experienced echocardiographers.

Here is what we learned while reviewing the first AI-guided image acquisition. The addition of such clinical AI applications and the potential for new users of these devices similarly affect the benefit-risk profile and the expectation for the safety and effectiveness of these devices.

There could be two types of errors-- device error or failure to provide guidance on acquiring diagnostic-quality images or signals, and there are user errors or operator failure to follow the guidance provided by the device to acquire diagnostic-quality images. Both errors can lead to delay, prolonged examination or additional unnecessary procedures. Of

course, intensive testing is required to investigate the safety and effectiveness of such technology.

As the computers go through a big evolution, the ultrasound systems are changing, too. Just like computers, medical ultrasound imagers have been getting smaller. They are now more portable. This advancement opens the door for ultrasound to be used in different intended use environments, such as ambulances and patients' homes.

Future direction. The advantages of ultrasound imaging make it an attractive modality as an in-home monitoring tool with real-time feedback. Of course, safety and effectiveness should match the professional use environment.

There are benefits and risks associated with in-home use of ultrasound. Benefits could be that these indications for use will provide valuable information to patients who are, for example, on bed rest or live in an area of limited access to clinics or healthcare facilities.

Please note that this device should be used under supervision of healthcare professionals.

There are also risks associated with in-home use. There may be unnecessary prolonged exposure to ultrasound energy if the user doesn't follow the instructions. If the user -- the user may incorrectly interpret the information if the image quality is bad -- for example. Or sometimes there may be no oversight of how the device is used by the patient.

Ultrasound has advantages and could be a great candidate for AI-guided image acquisition, which could help professional and inexperienced users. We would love to hear your ideas about innovations in this space, so come discuss these ideas with us through the pre-submission program. We will discuss the most appropriate regulatory pathway for a specific intended use of your device, we could discuss performance testing to support the intended use and if testing and labeling could address any applicable special controls. We

encourage you to take advantage of this free service and come talk to us early.

Through this workshop, FDA is seeking to engage with you to explore benefit and risk of evolving AI in radiology, and today we focus on AI-guided image acquisition using ultrasound.

We will talk about benefits and risks of AI in ultrasound application and as the intended use and benefit-risk changes, it is important to use methods to evaluate and characterize the performance testing of this technology for intended use.

In this workshop, FDA is also seeking innovative and consistent ways to leverage existing methods and to develop new methods for validation of these AI-based algorithms, and explore opportunities for stakeholder collaboration in these efforts.

The first session this morning is about opportunities and challenges in AI-enabled healthcare. Our first speaker is Dr. Anthony Samir, Director of CURT and Service Chief from the Body Ultrasound Department of Radiology, Mass General Hospital. His topic of the talk is No-Human-in-the-Loop AI-enabled Healthcare: Risk, Rewards, and Regulation. Please welcome Dr. Samir.

(Applause.)

DR. SAMIR: Thank you so much, Marjan.

Good morning, everyone. I'll briefly introduce myself. I wanted to also acknowledge Laura Brattain of MIT and Masoud Baikpour, one of the post docs in my group who worked on this presentation.

I have a number of disclosures but none of them are pertinent to this talk. I don't have specific conflicts associated with this talk. I have several roles. I'm also here representing- just go back to the slide- the AIUM. The AIUM has recently formed an artificial intelligence community of practice to facilitate the growth of artificial intelligence applications in medical ultrasound across all specialties. I am the incoming chair of that

group and for those of you who will be attending the AIUM conference, I encourage you to attend the first AI session, which will be on the Monday of the conference.

My group does bench-to-bedside research in medical ultrasound. We do signal processing for ultrasound physics and machine learning with approximately 10 machines to any scientist, hardware development and clinical translation and clinical trial design and I'm really thrilled to be here and very honored in a way because the -- and I'm a physician and radiologist-- but the NIH, through the mechanism of a K award gave me 4 years of supported time which I diverted to the engineering sciences and to learning the underlying quantitative methods of machine learning.

This talk is going to be partly clinical, partly technical, partly high level but funneling down to specific applications of machine learning, and I'm going to try, as I go through this, to take a systems approach.

Let's speak a little about my main personal research interest, which is no-human-in-the-loop artificial intelligence. I'm going to speak about the process of ultrasound imaging and machine learning, some examples from our center, the Center for Ultrasound Research & Translation, and then at a high level, I'll speak about what I think the implications are for the regulation of machine learning technologies.

There are a few key points that I'm going to start with, but I hope you'll leave the talk with these key points at the end. The first is that no-human-in-the-loop AI will be ubiquitous. In fact, I think that no-human-in-the-loop AI is going to be more common and more ubiquitous than human-in-the-loop AI in medicine and in fact, in society in general.

The second is that the risk profile of no-human-in-the-loop AI is fundamentally different, not worse, just different to conventionally engineered systems. And we're going to speak a little about some of the underlying statistical and mathematical principles that inform that insight, and it flows into the notion that postmarket surveillance is going to

acquire increased importance.

So what is this idea; what is no-human-in-the-loop AI? Well, for many applications and many kinds of machines, a human cannot practically be involved in decision making or from a medical imaging context, you might think about image reconstruction. We've had a lot of conversations and discussions in the last day about how a human might sit at the workstation and check the classification output of an algorithm, but that doesn't really address the fact of where the image came from in the first place and I don't believe that humans are going to be checking CT sinograms or ultrasound channel data, you know, to determine whether images have been reconstructed accurately.

It is highly probable that most ultrasound devices are going to be using machine-learning technologies for echo location and beam formation in the coming 15 years. The images that you're going to be looking at to do classification with the assistance of AI may well be generated using AI technologies. It's not possible to have humans in the loop for this any more than it's possible for you to really get involved in the middle of the highway if your Tesla is barreling down there and you're going to sort of override it in some way at the last minute. And this is a very, very common thing that is being deployed throughout all of society and all of medicine.

So really, what this looks like is circumstances where a human cannot practically be involved in decision making. It's very common. It's happening now. And it's also important to understand this doesn't mean humans will replaced, but it does mean that the notion that we're going to be able to check everything is not a real or true notion; it's not an accurate description of what happens.

Now, if you're a radiologist or another kind of imaging physician, you might have the impression from your perspective that the process of imaging or the system of imaging is something that a technologist does and a request comes in and then the imagery is

presented to you and that is a valid perspective. But from the perspective of patients referring physicians and healthcare systems, medical imaging is not separate from the radiologist, the radiologist is part of a system that takes a request and turns it into a report-- a useful report. In other words, we have a system that takes incoming data and converts into an outgoing data or outcoming data that can be used to positively influence a patient's outcome; and that outcome uses a largely currently deterministic system to produce image data and a natural intelligence, not an artificial intelligence, a natural intelligence in the form of the radiologist to produce the report. We currently have a system that has a deterministic element in the machinery and a stochastic element in the radiologist to produce the report.

And you see this in the process of ultrasound imaging, which I've depicted here. So you have an operator, the operator is holding the transducer, and the operator has control of that transducer and what you can see in front of you is a diagram that depicts the conventional six degrees of motion in a classic Newtonian space that every mechanical engineer will know. You can move things six possible ways, and we have words for that in ultrasound, fan the transducer, translate the transducer, rotate the transducer, six degrees of freedom.

So the operator looks at the imagery, the imagery is integrated in the operator's mind to perform analysis, intervention, and diagnostics but also to understand whether the transducer is in a desired position or the desired state and that state is position and pressure, and then to move the transducer to a new position that provides better data. One of the areas, and the area of great interest today, is the use of machine intelligence in order to guide the process in which the six degrees of freedom motion of the ultrasound transducer that would typically be guided by the natural intelligence of the user in the imaging system could be replaced by some form of artificial intelligence.

Another way to think of this is like, so in front of you, you have some ultrasound imagery that's taking place and if you were to just look at the bottom images for the moment, what you can see on the left is the urinary bladder and on the right, the gall bladder. But if you were to ask a user who only had the information on the imagery at the bottom what the gall bladder was and what the urinary bladder was, you might have a hard time deciding, right?

But you can decide and the way you decide is context. Some of the context is on the imagery located above each of the images where you can see liver or you can see pelvic anatomic structures, and part of it is the fact that you're holding the transducer in your hand and you know where it is and your brain integrates the positional information with the image information in order to provide useful interpretative data. And really, when you're talking about AI-guided image acquisition, what you're talking about is the combination of sensor data and image data in order to produce a similar level of spatially tagged contextual awareness to make a diagnosis and that's what this looks like, okay.

You can also use AI-guided image acquisition in the analytic phases, so this is an example of a shear wave elastogram. In this particular case it's breast tissue, but shear wave elastography is widely used in both the breast and the liver. And the main point that I'll make and I think will be obvious to everyone in the audience is that the input data is very complex, it has quantitative components, it has spatial components, there's a color map, it's quite hard to work with and typical current ultrasound systems simply measure a series of pixel values and provide an average and show a velocity or estimated Young's modulus and then that's used. The kind of statistical richness of the data is collapsed in order to provide useful diagnostic information.

But it turns out that you can build, and we have built, systems that will automate the entire process, will place the region of interest for you, will winnow out the imagery that

isn't valuable and after having done all of that, can perform machine learning-based computational diagnostics on the rich imagery and this is another example of AI-guided image acquisition and analysis and what this looks like in this case is the automatic region of interest placement. I'll use the mouse, thank you. I'm pressing the wrong thing. To use the automatic feature maps in order to automatically place regions of interest and then, you know, the obligatory ROC curve. But what this ROC curve shows is that using one-tenth of the amount of data that's used with conventional ultrasound systems, we can improve the AUC from 0.74 to 0.89 for the distinction of moderate or greater liver fibrosis, which is a very common pattern.

The other thing you can do is you can do simple things like hepatorenal index estimation, another set of algorithms that we've developed, and this is an important and useful thing to do and in fact, has pretty good results in many papers that have been published but it's not widely done, and it's not widely done because it is laborious and hard to do reliably across different technologists from a methods perspective. So what you can do is automatic real-time segmentation of the liver and kidney, you can automatically select regions of interest, you can compare organ echogenicity, and you can perform tissue typing diagnosis.

It's not only confined to what you might do in a clinic or what a sonographer might do, you can give a patient an ultrasound device and you can guide the patient to scan themselves and do simple tasks like detect and localize ascites, which you can do with the unit and again, you can guide a patient to scan their own abdomen, detect pockets of ascites and provide you with information regarding whether that ascites could be drained. And when you look at an expert user's estimation and AI-guided estimation of whether fluid is small or large in amount, you find that both of them are roughly equally inaccurate; nonetheless, clinically very useful.

This is a project in my laboratory that has been done for the Department of Defense. Again, a lot of the conversation is about artificial intelligence. Will artificial intelligence replace humans, but the truth is that artificial intelligence is going to perform functions in circumstances as I've spoken about earlier where humans just aren't available, and when you're in a field on a hillside in Afghanistan and people are shooting at you and you need to urgently cannulate a deep blood vessel, you may not have the luxury of a trained operator to do so.

Using handheld devices, you can automatically, in real time, detect blood vessels and if you have the appropriate algorithms and levels of clinical expertise, you can automatically categorize those blood vessels, detect them, and categorize them as arteries in legs. This is work that we've done in collaboration with MIT Lincoln Laboratory sponsored by the Department of Defense and you can use this and build this into hardware-based systems to perform automated interventions.

There are a lot of rewards with AI in healthcare. There are also some risks and it's useful to speak about these briefly. Whatever you may or may not think of Donald Rumsfeld, this was a truly wonderful quote and it's a long quote, I'm not going to read all of it, but I think the thing that really summarizes -- it really summarizes the challenges here very well, there are things we do not know and there are things that we don't know that we don't know, and the big challenge in the regulation of medical AI at the moment and ultrasound in particular is we don't know what we don't know. And a couple of slides that illustrate the trade-offs that we're making.

Deterministic processes are processes in which you know exactly what you're doing. The device is engineered using known specifications, the behavior of all the components is clearly understood, there are no black box components and you can engineer these. And this is a sort of example, if you will, of a conventionally engineered system and these

systems are wonderful because they're explainable and well understood. They're also typically characterized by relatively inferior model performance.

And then you have stochastic systems where a lot of things are happening at once and it may be hard to know exactly where each item that's being washed is at any given point in time but nonetheless, you end up with superior model performance.

And the problem here is that we don't quite know our unknown unknowns. We don't quite know what the deviations from predictable outcomes will be, that lead to superior model performance. The tradeoff is not properly characterized mathematically or scientifically and this is the big challenge. And the big challenge with regulation is unknown unknowns, in my opinion, dealing with the unknown unknowns.

And I think from that perspective a key thing to recognize is that there's a fundamental limitation and we need to recognize that premarket, pre-characterized regulatory processes are probably not going to be able to perfectly capture all of the risk and postmarket surveillance is going to be really important.

And the interesting thing is, there's a really good analog for this because if you think about what we do with humans and when you have the true imaging system which incorporates the human operator and the human interpreter, we already do surveillance and we call that surveillance continuing medical education and for those radiologists in the room, online, has anyone done their OLA cases, Online Longitudinal Assessment? We surveil radiologists and other physicians postmarket through M&M meetings, through recertification and other mechanisms, it's highly analogous.

So back to the key points at the beginning of the meeting. No-human-in-the-loop AI will be ubiquitous and I would contend it will be more ubiquitous than human-in-the-loop AI.

The risk profile is fundamentally different. It is not proven, but I believe that the

systems will ultimately prove to be less risky than deterministic systems, not more risky. But the issue is that the pre-specification of that risk is not likely to be completely tractable and because of that, I believe postmarket surveillance will be key. Thank you so much.

(Applause.)

DR. NABILI: Thank you, Dr. Samir.

The next presenter is Dr. Randy King, Program Director of Ultrasound, Division of Applied Science and Technology, National Institute of Biomedical Imaging and Biotechnology.

DR. KING: Hi. Thanks for having me. As Marjan said, I'm Randy King, I'm a program director at NIBIB here at NIH. I concentrate on the ultrasound and optics portfolios. Mainly, a lot of my talk will be on ultrasound, we will touch a little bit on MR, but I want to come at this in a little different aspect from the FDA talks you've heard before. I'm going to talk mainly about preclinical research and some of the research that we're currently funding to give you an idea of where we think the field is going.

To start with, here at NIH -- I'm going to assume everyone knows what we do at NIH because you're all here today, but we fund biomedical research for the good of not just America, but for human healthcare throughout the world.

Specifically, at NIBIB we have a unique role at NIH. We are the technology development institute and what does that mean? We're one of the only institutes that's not specifically focused on a specific single disease, an organ or a tissue type. We're interested in the convergence of engineering, computational technologies, and imaging with life sciences and we look for projects that are applicable to a broad spectrum of disorders and diseases rather than limited to that single use. NIBIB does not develop projects with the understanding of basic biological functions or mechanisms and that's a little different from the rest of NIH where NIH covers basic biological bench-top studies to

Phase 3 clinical trials. NIBIB is unique in that aspect that we are looking just for that technology development aspect.

Specifically, I'm from the medical imaging world, so we're going to talk about medical imaging and image guiding and how they're evolving. I think the next two slides are going to be really preaching to the choir of this crowd, but we moving from an anatomical and functional imaging science to more functional, looking at the physiology, looking at motion, looking flow dynamics. We're going from simply image-guided to information-guided applications and I'll show you specific examples of that later on.

And right now, at least in the preclinical research world, we're at the dawn of artificial intelligence and so the for the sake of this talk, I'm going to combine machine learning and deep learning neural networks, because it's all artificial intelligence, so I'll be combining them all under one umbrella, which I know some of you might be cringing at right now.

So again, showing you, just to give the basics to start with, data is where the truth lies. This came out of a DoD publication on artificial intelligence where they started with the image and everyone in the audience would agree that the first image is a panda. A deep neural network labeled it as a panda with about 60% confidence. Adding a slight amount of noise, 27% noise which the neural network actually classified it as a nematode, as worms, adding a slight amount of noise to that image, the deep neural network went on to assess that image as almost a hundred percent as a gibbon. And I venture to say almost everyone in the audience would still say that is a panda.

And we're all familiar with this problem that can arise from artificial intelligence, especially when you're talking about something that first that could be a benign hyperplasia. An artificial image -- artificial system gets the image and diagnoses it as a HER2 positive breast tumor. Completely different organ, completely different situation. That

brings us to what is currently happening in the world of AI and how. I'm going to concentrate not just on how we're using AI to interpret the data, but specifically on how we gather the data.

To start, I'm going to show you an example of beamforming and speckle reduction, this is from a group at Stanford with Jeremy Dahl. And what you see on the right column is the reference, that's your phantom, it's a photograph of the phantom. The next column on the image is delay in some beamforming, so that's a standard beam-o (ph.) imaging from ultrasound, so that's what you're mainly going to see in the clinic.

The next column, SC, is Spatial Compounding. It's a classical processing technique where you have independently beamformed subapertures to come at targets from different angles for reduction in speckle. SC can give you reduction in speckle but you also lose some lateral resolution.

OBNL is a nonlinear post-processing method whereby you get significant speckle reduction and it works really well. It excels at sharp discontinuities, so where you go from specific dark to light, hypo- to hyperechogenistic areas and point targets, but it struggles to maintain resolution at broader targets and gradual changes in echogenicity.

And you'll see in the last column, that's the neural network beamforming technique that takes place inside the machine and you'll see a reduction in speckle while maintaining that nice lateral resolution and contrast.

And the two images on the right show that it can be done with hyper- and hypoechoic phantoms and images. This has also been done in a few patients in liver lesions.

We're also funding work which is redefining tissue characteristics. With standard ultrasound, the B-mode image on the left, you see three round circles. These circles are a vial of glass beads. They vary in size from the left to the right, increasing in size. With the standard B-mode image, you simply get an average of the intensity of the received

ultrasound signal at each region of interest. Based on the packing density of the glass beads and the size of the glass beads, you see no discernible difference in any of the three vials in the B-mode imaging.

Well, this work we're funding from the University of Illinois at Urbana-Champaign, Illinois. You'll see quantitative ultrasound and with quantitative ultrasound and the neural network processing, you can actually get an indication of the average scattering diameter, so you can see a difference in the three vials in the colored image on the right and you can actually see a change in the size of the scatters down to tens of microns. And this is pretty phenomenal for ultrasound because you're getting sub-wavelength resolution now with the addition of AI.

Moving on a little bit out of ultrasound, I'm going to look at the reconstruction of 7T-like images from 3T MR-acquired images. With this, on the left you'll see a standard, clinically available 3T scan of a patient in a clinically available magnet and on the right you'll see a 7T MR image of the same patient in the same area. And I think you'll agree, you get enhanced anatomical details and tissue contrast with the 7T. The problem is the 7T isn't clinically available.

This group with neural networks can reconstruct the acquired 3T MR data to the resolution and contrast of a 7T, it's called 7T-like image. This dataset was done with five healthy volunteers and eight patients with epilepsy and two with myocardial infarction, so this has been tried in patients. And you'll see the results. On the left you see the 3T data. On the far right, you'll see the 7T ground truth, so that's a 7T scan of the same person. And in the middle, you'll see the reconstructed image from the 3T data and it's very analogous to the 7T scanned image. And this is pretty groundbreaking, it has huge implications for rural disadvantaged, small and poor hospitals, where we're going to now allow any patient in the country to have access to state-of-the-art imaging techniques. A patient in a poor

small hospital gets scanned on the scanner they have available, whether it be 1.5 or 3T, and the images get reconstructed to the state-of-the-art 7T. So we're applying improved healthcare to anyone in the country and with cloud computing, the extensive computer infrastructure isn't necessary to be on site, either.

Many people also jumped on the bandwagon to use AI to help with the analysis of images as they come out of the scanner. Almost all of our applications have some AI component in one form or another now.

We're funding a group out of New York, Dan Sodickson of the P41 Center who's asking and others are asking the same question, if we're using neural networks to reconstruct the image and we're using the neural networks to help us with a diagnosis, and this group's also working and partnering with Facebook which is kind of an unexpected partner, but if we're using these neural networks, why can't we then just remove the middleman and actually use the neural networks to not only help us acquire the image but also do the diagnostic? And we're going to be working on this data in diagnostics.

The patient will go in the scanner and then they'll receive the information from the images without the actual images. The images are only necessary for the human to actually examine, and it takes a bit of computational power to do so.

We're funding this data diagnostic where we don't believe that the answer is going to be in removing the human. A lot of our projects are to aid the radiologist or aid the technician in their assessment of the data and this is one of those aspects where the AI will assist and they can do some prescreening and then the AI can -- then the radiologist can pull up images as they see fit later on for special interest.

As I am program director at NIH, I can't leave here without talking about some of our funding mechanisms. At the NIBIB we have the standard funding mechanism, the R01 for standard data, for standard projects with pulmonary data. There are three, which is a pilot

-- for pilot projects with supplementary data but much smaller than the R01.

And our specific R21 mechanism. At NIBIB we've gone back to the true nature of the R21, which we expect very highly innovative, high-risk/high-reward type of research. R21s come in two flavors; we have the trailblazer mechanism, it's for the new investigator, which means they haven't had substantial NIH funding before, like an R01 and it's for that novel -- that novel project. With the trailblazer we do a lot of supplementary data, it's limited to half a page which can include one figure, but to help an investigator we allow 3 years and $400,000.

For exploratory developmental R21, which we call an Omnibus R21, anyone can apply and establish they're a new investigator but it's specifically for established investigators. Again, since it's R21, the project has to differ substantially from the current thinking or practice.

On this mechanism, preliminary data is not allowed and applications are screened when they come in for preliminary data. If preliminary data is there, they will not go forward to review and they will be withdrawn. Preliminary data is defined as data that has not been published. If it's been published, it's just a reference, you're allowed to put as many references as you need, but don't delete the novelty of the project. These are for $275,000 for 2 years.

We recommend that you contact NIBIB, the program director specifically, to talk about your project before you submit, especially with the R21s because we have the unique R21s that can't be transferred. If you do not meet our requirements, they'll be withdrawn and there's not much we're going to able to do about it. So if you take nothing else from this talk, take home to contact your program director often and early.

As I said, I'm at the National Institute of Biomedical Imaging and Bioengineering. My expertise is in the ultrasound and optics field, but we do have program directors in the

audience. Behrouz Shabestari, Zhe Wang and Dr. Krish Kandarpa are there and they each have a portfolio in artificial intelligence and varying aspects of artificial intelligence and I encourage you to either find me later, e-mail me and I can put you in contact with the right person to discuss your project specifically. Thank you.

(Applause.)

DR. VAEZY: Good morning, folks. It's good to see you on this second day of our workshop and it started wonderfully, continuing the tradition that was set yesterday. Thank you, Jen, for organizing and chairing this wonderful workshop.

We're going to go right into the second session, basically the continuation of the first. In the first two talks we basically heard about the vision of AI-guided image acquisition. Now we're going to get into the development of these visions into reality, into products. And later, in the next session, where Dr. Garra will be moderating, we'll hear about the clinical applications of those technologies, if you will.

In this session, as I mentioned, we're talking about innovation, so innovation in AI-guided image acquisition. There are different goals, different objectives for an AI-guided image acquisition. One could be standardization of the image quality, images that are acquired by experts and we know there is a need for that. The other could be, of course, the assistance provided to somebody who is not as experienced, perhaps, not an expert in acquiring certain images, and the last one will be actually bringing novices up to a level where they could acquire images that could be of diagnostic quality. We hope to cover these areas in this session.

We're going to start off by basically having a presentation from a company that we have heard a lot in the last couple of days and that's Caption Health. Dr. Ha Hong is going to talk about Development and Validation of a Breakthrough AI-Guided Echocardiography

System. Dr. Hong.

DR. HONG: Okay. Good morning, everyone. Can you hear okay? All right, great. Good morning, my name is Ha Hong and I'm from Caption Health. It is my great honor to be here today and to talk about this great workshop on the new field of AI and medical imaging.

In this talk, I will talk about the development and validation of breakthrough AI-guided echocardiography system or heart ultrasound acquisition system. This talk is quite timely because, you know, as FDA just authorized this novel technology 3 weeks ago through a de novo clearance pathway. Today I'll talk about the innovation first and then focus on the validation of the technology, which is critical to demonstrating the safety and effectiveness, as well as confidently placing this new tool into the hands of doctors to improve patient care.

Let me first begin with my personal journey on how I got here. I grew up in South Korea, studying physics. One day I wanted to study the brain; fortunately, I was able to enter the Harvard MIT Health Sciences and Technology Ph.D. program. There, I studied the brain and I wanted to make computer models that can emulate the end-to-end visual perception processing happening in the brain.

Specifically, I wanted to make a computer program that takes an image and predicts neural responses to that image, I mean real neural responses, and then solves all different kinds of visual perception tasks such as visual object recognition and post-vision (ph.) tasks. These days everyone -- I mean, all of us know that, you know, computers can solve these kinds of problems, but when I first investigated 10 years ago, it was considered an unsolved problem. It was the holy grail ofmachine learning in AI. But as my first step to pursue this holy grail, I started collecting lots and lots of neural data. Specifically, I collected real neural

data from hundreds ofneurons from high-level visual areas for thousands and thousands of images, making one of the largest publicly available neural datasets so far. And then, I collected a similarly large human behavioral dataset for the same set of images.

With these rich datasets, my coworkers and I were able to create a computer model that was able to predict neural responses at a level far better than any other computer models at the time. At the same time, it was the first implementations, early implementations, of deep neural technology. I wasobviously we were quite happy with that.

And then we were also able to show that those computer models and real neurons that I collected shared similar representations and performance characteristics for certain core visual object recognition tasks.

Finally, we were able to create a computer model that was able to perform end-to-end visual perception model of the brain, essentially taking on image, predicting neural responses and then performing all sorts of, you know, visual perception tasks.

I was very fortunate to be a part of this discovery, contributing not only to the neuroscience and computer science, but also to the development of AI and deep-learning technology in general.

When I first began, you know, computers simply could not solve these problems but by the time that I was about to finish my Ph.D., you know, computers and especially deep-learning and AI technology starting to show promising results in, you know, visual perception, speech recognition, language translation, and also in radiology.

And then there was a serendipity. One day -- a good friend of mine and my coworker who -- Charles Cadieu came to me and he started to talk about a crazy but a very interesting idea. He started with one of the biggest public health problems, thatthere are too many people dying from heart disease. In fact, it is the number one killer in the U.S.

and worldwide. Many heart conditions are treatable or preventable. Early and accurate detection and diagnosis are the key and medical imaging is the gateway to this and in fact, ultrasound is one of the most effective methods for this with many advantages.

In ultrasound imagingyou hold a wand called the transducerand point it to the body to see what's going on inside. It has many advantages, it's very versatile, and can scan many different body parts. It is low risk and has no ionizing radiation and it can image in real time; it is very suitable forevaluating the pumping function of the heart and it can be made as a small and low-cost device. Because of these reasons, it's often called the stethoscope of the future.

Heart disease can be accurately diagnosed with echocardiography or heart ultrasound scanning, but this often happens too late. Let's suppose that you wake up in the middle of the night with severe chest pain and shortness of breath, symptoms consistent with heart disease. You are in severe pain and brought to the ER, the emergency room doctor performs physical exam on you using plain old stethoscope, but the doctor does not perform ultrasound on you. In fact, the doctor cannot. Although they clearly know all the benefits and advantages of using ultrasound in this particular case, there are not that many people who can scan. This bottleneck, CDR bottlenecking, image acquisition causes delay in diagnosis and treatment, costing you a life.

So why this is such a bottleneck? Why aren't more people able to scan? The problem is that ultrasound is so difficult, it's very difficult to use, especially in cardiac space. It is an unnatural hand-eye coordination problem, a tiny movement in hand positioning can dramatically change the image on the screen and the images themselves are very unintuitive. You do not know where you are, where the target is, how to navigate, and even if you hit the target, you do not know whether that is good enough or not. And to compound this problem, everyone is really different. Obesity, pathologyand implanted

devices can introduce significant variations on the images and also, simply everyone has different heart shapes . So that's the reason why there's a bottleneck., It takes years and years to master this and to become an ultrasound imaging professional.

And I would like to emphasize that this is the main barrier to cardiac ultrasound scanning today. For example, if you give an ultrasound device for free for everyone, the problem will still persist because of this difficulty in learning. So that's basically what Charles and I talked about5 years ago in a small cafe at Cambridge and he asked the following question: So with AI, is it possible to make an ultrasound navigation system for everyone like, emergency room doctors? My answer to this was yes, obviously yes. I witnessed the full potential of AI and deep learning during my Ph.D., emulating the human abilities andall the expertise.

In the same way I thought that it should be possible to make an AI-based ultrasound navigation system thatrecognizes all the nuances of ultrasound imaging and alsoproviding the guidanceand capturing the expertise ofultrasound experts. At the same time I thought that working on this would be a perfect application of my expertisebenefitting not only the science and medicine but alsopublic health in general. And since then, I have been taking this journey for the last 5 years and it has been very, very rewarding.

So now, in 2020 today, I would like to introduce the AI-based heart ultrasound guidance technology that can empower healthcare providers to acquire high-quality ultrasound examinations. In a nutshell, it is a heart ultrasound navigation system.

The technology has been developed to emulate the intuition of imaging experts. The guidance technology predicts the deviation of the current transducer positioning from the ideal positioning in real time. The software then provides user instructions to efficiently minimize the deviation.

Specifically, the software recognizes the current transducer positioning purely based

on ultrasound imaging. Then it predicts the target ideal positioning and provides the user the instruction to reach the ideal. In this particular case, to slide down.

And finally, it recognizes when the user reaches the target, getting a diagnostic image, by the time the system automatically acquires diagnostic-quality images for you. It supports 10 different cardiac views or 10 different clinically relevantways of exposing different heart anatomies.

By the time we finished our prototyping, FDA accepted this technology into the breakthrough devices program. This classification is given to the device that is expected to provide a more effective diagnosis of a life-threatening or debilitating disease or condition, in this case heart disease, when there's no approved or cleared alternatives. And we are very grateful to FDA for their support on innovation.

Switching gears, now I'm going to talk about how we tested our innovation. Rigorous testing is very important because rigorous testing is very important for introduction of a new technology because of its potential impact on patients as well as, establishing a good precedence for all future devices in this new classification.

Let's take a look at one example verification test we performed. In this test we evaluated the performance of the system in terms of recognizing diagnostic-quality imaging. For each view, we recorded hundreds of ultrasound video clips of different imaging qualities. We recruited five differentexpert readers. At a given time they saw one clip and answered the following the question: Is the quality of the clip suitable for diagnostic purposes? We collected panel evaluations for all the clips across 10 views that the system supported and computed the performance in the metric ofarea under the curve. The performance was high and as a comparison, this is the human expert performance. As we can see, the two were comparable. We performed six more bench testings like this and got similarly a satisfactory result. This enabled us to ensure the proper operation of all the

components of the system.

With this in mind, we then tested the clinical efficacy of the system with a pilot study. In this pilot study we recruited 16 informed consent subjects with a wide range of body mass index orobesity and pathology. Specifically, we made sure to include more than 50% of obese or overweight subjects, as well as many subjects with, a cardio pathology including implanted devices such as pacemakers or artificial valves. People with these conditions tend to produce ultrasound images with reduced quality –(that's why we recruited them) to better test the safety and effectiveness of this technology.

Next, we recruited imaging experts, ultrasound experts, with more than 5 years of professional experience and they acquired exams of 10 views for each patient without the use of AI-guidance technology to establish the reference.

Next, we recruited nurses with no prior experience on echo scanning at all, one example group, user group thatour technology can potentially help, really help. They first received 2 days of minimal training, mainly abouthow to use our software, not about, traditional ultrasound echocardiography. And then they independently scanned exams of 10 views with the help of the AI guidance technology.

The exams collected were finally evaluated by the panel of five expert Level III cardiologists independently and blindly. At a given time, they saw one exam of 10 views and answered the following question: Can the exam be used to qualitatively assess the following clinical parameters such as left ventricle size, function, right ventricle size, and pericardial effusion? We were primarily interested in the percentage of nurse-acquired exams that were determined to bediagnostic quality by the panel.

So how well didthey perform? It would be an interesting exercise for you to pick a number for yourself. Everyone ready? This is the answer. The panel determined that the majority of nurse-acquired exams to be of diagnostic quality. Obviously, we are very happy

with this result and communicated this in major cardiology conferences.

And finally, with this positive result in pilot study, we proceeded to the validation study with a large-scale, prospective clinical trial. In this clinical trial we kept the basic study design the same, asking the same four clinical parameters as primary end points. One important addition is the performance target which was based onclinician surveys meaning that , on average, clinicians indicated that, if the non-experts can produce, clinical diagnostic images about 80% of the time, they were satisfied.

The trial was done at two leading institutions and based on the power analysis, we recruited 240 real-world hospital patients indicated forecho examinations and eight nurses who without prior experience on echo scanning, making one of the largest clinical trials done to obtainfor clearance in echo space. Each nurse received minimal training like before and then independently scanned 30 patients with the AI guidance.

Importantly, there was a stratification in terms of patient obesity and also cardiac pathology, meaning thateach nurse was required to scan a certain number of obese patients and also patients with cardiac pathology including implanted devices, again, to better test safety and effectiveness of our device.

Well, what was the conclusion? What was the result? This is the result. Basically, we produced the same result we obtained from the pilot study. The panel, again, the panel of expert readers concluded that the majority of nurse-acquired exams were diagnostic quality.

Finally, this series of R&D and validation activities culminated into a milestone event where FDA granted a de novo clearance on our technology named Caption Guidance. This is the regulatory pathway for novel, low to moderate risk devices and this is the first AI-guided image acquisition software to be cleared and approved.

To summarize, the AI-guided image acquisition technology emulates the skills of

experts to empower frontline healthcare providers. It demonstrated safety and effectiveness with extensive tests. And we hope to see a substantial increase in access to echo in clinical settings where it used to be impractical or impossible, such as in emergency rooms or in rural clinics.

Hope you enjoyed this and thanks for listening.

(Applause.)

DR. VAEZY: Thank you, Dr. Hong. I liked the tutorial nature of your presentation.

And next we're going to move up the West Coast from San Francisco to the Seattle area. Dr. Benny Lam of Philips Ultrasound is going to talk about Deep Learning-enabled Ultrasound Imaging – Opportunities, Risks, and Challenges.

DR. LAM: Thanks, Shahram, and thanks for other speakers for setting up the stage for me. You guys' presentations already helped my presentation. And I would also like to thank FDA for giving me the opportunity to share similar view on AI-enabled image acquisition. Today I'm going to talk about some benefits of using deep learning in developing solutions for ultrasound image acquisition and associated technological and clinical challenges and some opportunity for us to work together. I hope that my presentation can stimulate further discussion, especially on the regulatory science development that will lead to more efficient ways to ensure the safety and effectiveness of these kind of devices.

My presentation is a team effort across several businesses in Philips. I would like to thank these people for their support.

Let's start with taking a look at what ultrasound currently can offer. Ultrasound has been known for its portability. Most of the commercial ultrasound systems are delivered –

which means that you can move them around in the hospital. The one on the right is an ultra-portable one. The whole ultrasound system is just a transducer and a tablet. You can bring it with you everywhere.

Ultrasound imaging is almost real time and highly interactive. You can see your image while you are scanning, and its high portability and real-time and interactive ability makes it an ideal modality for point-of-care diagnosis.

With the combination of different parts and software packages, ultrasound has become multifunctional. And ultrasound is relatively cheap compared to other types of imaging machines such as MR and CT. Furthermore, ultrasound does not have any concerns related to radiation. While ultrasound has quite a few clinically significant capabilities, there are still many places where the technology can be improved.

So according to a survey in 2009 (sic) from the Society for Radiologists in Ultrasound, the top clinical user demand is actually in image quality followed by workflow enhancement and clinical applications such as quantification and disease classification. In this context and the topics of discussion, the question you may ask would be related to how AI can satisfy some of these demands from an image acquisition perspective.

To facilitate discussion on how deep learning has been applied in ultrasound imaging, let's take a look at the ultrasound imaging pipeline. Ultrasound imaging starts with the ultrasound machine generating change of voltage pulses in various patterns to trigger the piezoelectric crystal in the transducer element where the ultrasound waves are generated. And ultrasound waves penetrate into the body, hit something inside and bounce back. The reverse process at the transducer element converts ultrasound waves back to analog RF signals, which are further processed. These analog signals are then converted to digital signals, which bring us to the digital world of opportunities.

With the signal being digitized, we can apply the -- computational technique which

other speaker has been talking about to achieve our goal, and deep learning is one of the many good technologies that we can use, which has also been pondered by a few speakers already.

As you can see here, deep learning has been applied to all of the processing on the right side, in the blue boxes. So deep learning has been applied to improving the beamforming reconstruction, speed and flexibility for better clinical workflow and expanding clinical application without sacrificing image quality, and these have been widely reported in the literature.

For Doppler signal processing, the deep learning solution also comes with an efficiency method while maintaining the structure, estimating accuracy and resolution. We have all seen this capability to learn in algorithm from end to end by examples. These provide important insights on how we can use deep learning.

In terms of application, for example, contrast-enhanced ultrasound or super-resolution imaging, deep learning has the flexibility for developing different solution for very different problems. Our experience also told us that the key is to understand and formulate a problem and to use deep learning in a meaningful way. And of course, workflow enhancement, as you guys have seen from other speaker, this is another great application of deep learning. I have further information about these examples with reference in the backup slides which can be downloaded from the FDA website, so you can read them on your own time.

This slide summarizes the benefit of using deep learning in developing ultrasound imaging solutions based on what we have seen so far. In terms of engineering benefit, deep learning solutions offer higher flexibility in development because the deep-learning model can be used for developing one component of the imaging chain or end-to-end, as you can see from the previous speaker. It can operate on RF data, beamformed data or even

reconstructed images. It is computationally efficient, robust, adaptive with minimal manual parameter tuning and can intrinsically exploit the hidden relationships in data. It can also replicate complicated calculations in shorter time by processing complex mathematical model learning by example such as singular value decomposition. If you come from the old school of machine learning, you must have experienced the process of handcraft feature extraction, so deep learning is a method that can save you from that.

With appropriate integration from human knowledge such as imaging physics and engineering, the development solution can become more efficient and data efficient, meaning it uses less data. By considering all these engineering benefits with using ultrasound, using deep learning in ultrasound imaging, the development of deep learning-enabled solutions are more flexible, robust, and generalizable than classical image and signal processing method.

This engineering benefit can actually develop into a clinical benefit such as faster and easier image acquisition without sacrificing image quality and emulating complex or expensive hardware with simple hardware and have a more streamlined workflow and a more sophisticated application. I believe, in essence, the clinical benefit here is in offering a better user experience for clinicians and a better quality of care for patients.

So same as many other computational method, deep learning has its own set of technological and clinical challenges. For technical challenges with ultrasound image acquisition using deep learning, I will bucketize them into three categories, as shown in the left column. I will also discuss how the existing regulatory framework has addressed it or covered this category and -- might need some additional attention or consideration. These questions are not, you know, specific to FDA, but they're something that we should think about together.

In terms of data quality and governance, I believe the FDA has provided a reasonable

amount of guidance and validation dataset origination and quality to match the intended use of the devices. Nevertheless, deep learning has its own set of unique challenges, especially in the need of loss of representative input and output datasets for training. Collecting, curating, and annotating this data requires lots of effort and investment. So the expectation of the system that we should put together in place to ensure adequate data governance and quality to support the device's intended use will be needed.

In terms of image quality evaluation, FDA has provided general guidance on ongoing clinical use in typical scenarios. For the deep learning-enabled image acquisition technique, the question will be more on whether the existing image quality evaluation methods can be adopted directly and if not, what additional considerations we should be aware of.

Taking along is how often should the image quality check be done after release because the output image is now used for diagnosis right away. I believe checking it every time will not be feasible and in the real world, not beneficial. But a smart way to address this problem should be considered.

The risk, therefore, will be different for an output image being used with other available information to support the physician's decision versus the image is the only information available to the surgeon during the surgery. I think that a flexible framework to capture the challenges and the risk level with respect to the intended use will be welcome.

Bias detection and correction has caused some attention at this meeting and FDA has provided guidance on checking the validation dataset and results for potential bias. However, details on checking the potential bias from training dataset and intermediate result will need to be worked out.

And as a user, I have read the user manual carefully enough or am experienced enough to understand the boundary of my AI-enabled device. I may accidentally use the device on something that it has not been trained on, that would be another challenge we

should think about.

Lastly, FDA proposed a good machine learning practice last year. I would appreciate if more detail can be provided to allow us to determine how to make this conceptual idea to work in practice.

Let's take a look at some regulatory challenges. Other speakers spoke about these various steps, so I'm going to pose some questions. For the release of retrained and refrozen improvement within the created intended use, it can happen very fast once the data supply, training, and testing pipeline are set up. Imagine that if there's a manageable and cost effective way for FDA and industry to communicate about these releases, what would that look like?

And product development can also become very fast if we can formulate the problem correctly. Imagine that there is a way that can synchronize the 510(k) submission and review with the product development within the existing regulatory framework, what would that look like?

Ensuring the safety and effectiveness is always be paramount for medical device manufacturer and FDA. Having said that, the regulatory sciences development to align with the speed of technological advancement has always been a challenge because the crystal ball that can precisely predict the future may be difficult to find.

And lastly, we agree that the transparency is one of the pillars of responsible AI practice. So how to determine the right amount of information such that the user adequately informs -- no, confuses it by the overload information will be something we need to think of practically. I believe these challenges we just talk about can become some opportunities in some aspect and I also believe there's no one-size-fits-all solution.

So here are some thoughts we may consider. First, I would like to acknowledge that CDRH, especially DRH and DIDSR have done an outstanding job in setting up the foundation

for IA-enabled radiological imaging devices.

The current premarket regulatory pathway is flexible enough to accommodate advancement of IA-enabled imaging devices.

Additional work will be more on leveraging and recalibrating the existing regulatory approach and tools that make the best of them to safeguard the safety and effectiveness. To support this, I think that setting up an agile quality evaluation and monitoring system which is comprised of a combination of performance measurement, monitoring feedback, and mitigation mechanism would be favorable.

A benefit-risk assessment framework that is adaptive to a range of benefits and risk levels could also be put into the system.

Flexibility in pre- and post-market data to demonstrate the total product life cycle safety and effectiveness could also be taken into account.

And of course, the real-world performance after release should also be a core component.

I believe that FDA and some folks in industry have thought about these items multiple times, but implementing them in a cohesive way with minimal disturbance to the operation of the existing regulatory system is something we should keep in mind.

Quality and appropriate data plays a vital role in the overall success of any quality AI-enabled devices. By establishing a practical guidance for data governance to allow users, manufacturers, and the FDA to build trust on the data that we use and that we are going to use in training and testing that are representative and with minimal bias and reflecting the clinical outcome of the intended use, we essentially set the baseline expectation.

And user education is important because AI-enabled radiological devices are getting more and more sophisticated. Therefore, some approach to develop effective user training, risk and benefit awareness to the user, and channels for the user, manufacturer, and FDA to

report and announce the issue where accidents occur are important.

And finally, I would like to conclude my talk. I think that regardless to the approach we are going to take, all of us are responsible for sharing the responsibility to ensure patient safety and well-being. Thank you.

(Applause.)

DR. VAEZY: Thank you, Dr. Lam, for discussing the benefits and challenges of AI and deep learning.

And next we're going to go to a company with a unique technology, Dr. John Martin of Butterfly Networks is going to talk about -- with the title of Can Machine Learning Tools Bring Diagnostic Imaging to the Home with Safety and Efficacy? Dr. Martin.

DR. MARTIN: Thank you very much and I appreciate the FDA offering their invitation for us to participate in this. My name is John Martin, I'm the Chief Medical Officer of Butterfly Network. We produced ultrasound on a chip. But I want to give this talk from a number of different perspectives and that is the perspective of the CMO of a company, as well as a physician with 30 years of practice history, as well as a history of being an administrator in a healthcare system, and probably the most important one to me now is as a recent cancer survivor and a consumer of healthcare. It gives you a very different perspective on technology and how our healthcare system works, and I want to put that perspective into this conversation and my goal here today is to be a bit provocative.

I think the most important lesson that I've learned along the way and as Simon Sinek says, this is the "Y" in butterfly and that is because time really matters and we rarely actually talk about this in healthcare. Two-thirds of the diagnostic dilemmas that we face, we don't know exactly what's wrong with a patient when we instantly see them, can be

solved with simple imaging modalities. The problem is the second you order a test you introduce the element of time.

Now, that time may be minutes, it may be hours, it may be days, it may be weeks and we face that all the time and as physicians we often discount that. But I can tell you now, being on the opposite side and any of you that have sat in an emergency room or sat with family members waiting to know what's wrong with you, time really matters. When we can accelerate the diagnosis for people so they know what's wrong with them, we not only relieve the anxiety that's created in patients themselves and family members, and in me as a physician, because I want to know, I want to start the right treatment, so we also can improve outcomes. So the goal here, both of our device and others, as well as in artificial intelligence tools is can we accelerate that diagnosis because time matters.

Now, time is important because as we move in healthcare and address the issue of the cost of healthcare, which is staggering, we need to move to the left of the care continuum. Where can we advance the diagnosis and treatment decisions to the left where they're less expensive?

And certainly, as I started in medicine and I ordered a test for my patient, that was a journey, it was a destination. You went off to get a test and that is not the most efficient way to deliver care. What we want to do is have the test at the bedside when we see a patient so we can make a decision and start care. Now, obviously that makes a lot of sense inside hospitals, but the more we move that to the left, the more efficient it is, the less cost it probably introduces and certainly, from a patient's perspective, that's a heck of a lot better. We talk a lot in medicine about re-admission rates, but what we're thinking about in our company is doctor-free days. How do I stay out of the healthcare system altogether, because candidly, that's the most cost effective.

We can't argue the concept that home use and devices have proven value. I

certainly started my career 35 years ago with the thought of every single patient being able to measure their own glucose was something that was unfathomable. Now it's commonplace. Who would've ever thought defibrillators would be on every course, every corner and in every church and every gymnasium? No one ever thought that was possible. And that's a device that delivers an electrical shock. If it's in the wrong place it actually kills you. Now it's commonplace. Home diagnostics actually can have proven value and I don't think it's so far off to see in the future where ultrasound, what we work with, can take that same place.

Now, we understand that the core of the FDA is safety and efficacy and we certainly have that as our core as physicians. What are those thresholds for all of us and certainly for us in our company? A consumer-level interface, they've got to be able to use it. You need the exam to be relatively easy to perform and the results need to be reproducible. And for us, our standard is we want the accuracy of that test performed by a patient or a consumer to match that of the professional and ultimately, much like we have with diabetes, we need to have some guidelines so that physician notification is known and appropriate.

If we're going to go down this pathway to the home what are the key issues we need to resolve? First of all, are there clinical conditions where this actually makes sense? Can a patient actually be taught to scan themselves reliably? Can AI make this easier and more reliable? Are there actually devices that can perform ultrasound now available and they can practically achieve this and be affordable? And are there reimbursement models in place to incentivize? At the end of the day we all realize you have to follow the money. If there's no incentives in there from a financial model to make sense, it's not going to happen.

Let me start with the clinical conditions. One that you may not think of but if you're awake at night, you see all the commercials for catheters, which I often am, you realize that there are hundreds of thousands of people in the United States and across the globe that

self-catheterize themselves. They have neurogenic bladders and there's a lot of other indications, but they catheterize themselves. What do they do that on? Time. We all know we don't go to the bathroom the same amount of time. Wouldn't it be just great if they could only catheterize themselves when they need to? And the fact that they can't leads to a whole host of office visits every year, emergency visits on average every other year, and a third of those lead to hospitalizations. They may be reduced or prevented with routine ultrasound.

And how about congestive heart failure? Estimates are from 40 to a hundred billion dollars spent every year on congestive heart failure. Four million hospitalizations at the minimum. The main cost of that is $12,000 a time. That is a problem that is screaming for a better solution and ultrasound could very well be that solution if done in the home.

We certainly know and have plenty of evidence inside hospitals and urology clinics that bladder ultrasound can safely and accurately measure volume. Can we teach patients to do the same thing at home?

And B-lines are becoming more and more prevalent as a tool for physicians that use point-of-care ultrasound to measure lung volume and correlate that with congestive heart failure. Wouldn't it be great if we could actually measure those B-lines at home instead of in an emergency room or inside a hospital and correlations to this data would make a lot sense for treatment.

So if we have two good use cases where we can improve quality, improve patient satisfaction and reduce cost, then we've got to ask ourselves the question okay, we've got a use case, how hard is it actually to learn? Courses are going on across the country and I would agree with one of our previous speakers, education is key here. Learning how to do this is a critically important part of that. Some of that actually could depend on artificial intelligence tools for guidance.

But just how hard is it to learn ultrasound? We paint it with a broad brush and think about everything, but if your education has to be focused on a specific use case, can you actually do that efficiently and quickly? And our fellows up in space are actually proving that. Those are not ultrasound experts up on the space station, but they've been taught to do this and can do it repeatedly because there are very limited applications in which they're using it.

How long does it really take? There's some pretty good data out there and studies that show in a number of different instances and if you read through these use cases, some of them are pretty specific. So evaluation of left ventricle systolic function. Can a person learn this yet? Eight hours and 20 practice sessions. That's not terrible. And if you look through this whole list, the screening exam for triple A, incredible sensitivity and specificity, 50 practice exams and the time to record that 4 minutes. I would argue this data certainly suggests we could teach people if there were focused indications that we want to teach them on.

And we have studies underway, some -- Nova at Penn actually published this study of home use people actually learning how to do this and applying it in the home, so the geography actually doesn't really make sense. And now we have two studies going on right now under IRB with patients scanning themselves for B-lines in the home. And I think the data will be very compelling to show that we may actually be able to do this.

Here's my conclusion. I'm a surgeon, we're very visual people when we operate, and I like to think about this journey. It's not as easy as just walking up a hill, but it's sure not climbing a mountain. And I think if you look back and say are there education barriers? Yes. Are they surmountable? Absolutely.

What about devices? If you've been involved in the ultrasound industry, you've seen the tremendous evolution of these devices from the big carts we started with in my days as

a vascular surgeon with Gene Strandness, all the way now to the device we now have that fits in your pocket and plugs into your iPhone. So clearly the portability and versatility of devices has gotten to the point where something in the home makes sense and now the price point is getting to the point where it's actually affordable.

There are a number of different devices that are out there that actually accomplish this goal and the key features of all of these are a cost structure that actually makes sense, the portability, you can get them into your house and you don't need to have a big space for them, and competitive image quality and I want to get to that point again in the future, but competitive image quality to answer the questions we're trying to answer.

Now, what does machine learning bring to the table for us? Two important elements, if we're going to do this, and that is image guidance and image interpretation. Image interpretation work is exploding and we have a whole 2 days over this with all kinds of different algorithms suggesting that and there are certainly tools out there already to count B-lines.

And I want to pause here for a moment because image interpretation is a really interesting thing. We're holding machine learning up to a very high standard, which I think is really important, but let's not lose sight of the fact that as humans we're not perfect. My colleague from Harvard presented great work yesterday that showed the tremendous variability that's out there in detection of cancers. It went, I think, for two to six at different centers. We have certainly seen that, as well. We know humans have tremendous variability and anyone that looks at echoes knows the really appropriate way to interpret that is a range. There isn't an absolute number. It's hard to explain to a machine-learning scientist that there's not just a single number that's the ejection fraction. It doesn't make sense to a mathematician, but in clinical work we know that there is beat-to-beat variability, there's situation variability, so what we use as a gold standard for our machine-

learning tools, we have to take into the context as humans we very well may be the gold standard.

What about image guidance? This is a critically important tool and I actually commend Caption for their work because I think this is important. We need this tool to accelerate confidence for novices and by the way, for us so that we can get there quicker. I also think it will promote consistent image quality. It will reduce scan times, which is critically important because of the role imaging plays. It will expand access to other classes of healthcare professionals.

Candidly, when I was in the hospital sick, the person that used the stethoscope the most was the nurse. Boy, would I like her to have a tool where she can go from listening and guessing to looking and knowing. And is it really the path to the consumer, this image guidance tool?

The other important issue about images, though, and we have to think about this, and I've learned this going around the country, image is in the eye of the beholder. One sonographer may like this image, another may like that image. A certain physician may say this is diagnostic where another may say it's not. There is significant variability in that. And at the end of the day, we are always striving to make better and better image quality. I would challenge you, particularly when healthcare economics comes into it, we have to ask ourselves the question how good does the image need to be to answer this specific question? That's our guiding light.

Now, why is this so important? Why does it make sense, actually, to say we have use cases that make sense, we can actually have devices out there that are affordable that we can bring into the home. We can use tools with machine learning actually to guide the user to get there.

And I'll stop at this one particular point which I think is so important. One out of

every four congestive heart failure patients are re-admitted within 30 days, one out of every four. Does anyone in this room actually believe that's acceptable? And the answer to that is an obvious no. If we had the tools at home to detect volume overload occurring earlier, the patient actually diagnoses themselves in conjunction with their physician and could prevent emergency room visits or hospitalizations. We all believe that's better for our healthcare system, it's better for the patient. So that's why this work is so important.

I will end with this comment, which is a really important one. What about reimbursement? We don't have great models for this now. We certainly have reimbursement models for telemedicine, which I like to call today tele-conversation. When we move diagnostics into the home, we may really see telemedicine realize its full potential. But we're going to have to address this issue if we bring it in a home, what is an appropriate reimbursement model?

As we move to value-based healthcare, this is great. I see these devices at the bedside reducing the cost of care and in a value-based system, they will make sense. But we also have to understand that if we have a device that we want a patient to have, how can they afford it, how can we create a model where it makes sense? If you have a $13,000 re-admission, that pays for a lot of devices. We just have to figure out the model.

My final personal opinion is can machine-learning tools bring us to bring diagnostic imaging to the home with both safety and efficacy? Absolutely. Do I think we need to move quickly? Absolutely, because finally, time does matter. Thank you.

(Applause.)

DR. VAEZY: Thank you, Dr. Martin. I inadvertently gave you infinite amount of time but you graciously used a finite amount of time, I don't exactly know how much, but yeah, we're going to take care of that next.

So for the last talk of this session, we're going to go to GE Healthcare Ultrasound where folks have developed AI-based, I would say, products in the ultrasound imaging space and so the title of the talk is AI-Guided Ultrasound Image Acquisition; GE Perspective.

DR. WASHBURN: Thank you. When I think about what we're doing in ultrasound and the history of that, you know, traditionally ultrasound has been this powerful tool that enables clinicians to make critical decisions and that doesn't change moving forward. But more and more, we see it as a powerful tool that's intelligently assisting clinicians in that task.

And of course, as ultrasound has become less expensive and smaller, more portable, where that's happening has been growing, as well, in terms of the care setting and in terms of the user for sure and seen with some of the examples from today.

And if we think about an analogy here to a surgical assistant, right, the surgical assistant knows at a similar level, has knowledge at a similar level to the surgeon to be able to provide this assistance and therefore, they can anticipate what's next and that provides a level of efficiency to the surgeon as they are helpful withwhat tools are needed next and so forth. And they could also serve as another set of eyes which provides improved -- you know, the goal of that is to provide improved outcomes, right? It would avoid a mistake. And this idea is that this awareness of the surgical assistants provided assistance to the surgeon. And we can think about this in a similar way for ultrasound assistance, right, and the idea that if the ultrasound machine or this ultrasound assistant can anticipate what's next, it can provide workflow assistance or enhancement to the clinician, which is a form of efficiency. And it can also serve as another set of eyes, this idea of providing some level of decision support, which ultimately has the goal of improving outcomes. And so again, this idea that as the ultrasound machine can be more aware of what's happening in the exam

and interpreting the results of that, it can provide another level of assistance.

And in general, we can think about the Quadruple Aim of AI in medical imaging, improving outcomes, which I will talk about as effectiveness in this presentation, providing efficiency, and then ultimately the patient experience and clinical experience. I'll really focus on the first two of these as we think about this problem in terms of guidance.

If we try to break down the anatomy of an ultrasound exam, there's a lot going on and it's an iterative process, as well. Many of these steps are performed over and over, if we think about a diagnostic exam, for example, and it gets more complicated than that because all of these steps have sub-steps that need to be performed and AI certainly will play a role in terms of helping to guide many of these steps of an ultrasound exam.

But if we think about guidance and acquisition in a focus on this view identification piece, I think it's always important to think more broadly than just one step because that one step is very dependent on what happens before and what happens after, right? So we're identifying a view for some sort of analysis or interpretation that needs to happen and we're making adjustments to the scanning to try to get a good enough image to solve the clinical task at hand. And when we look at the entire spectrum, the idea of using AI in the guidance space is to make this process more effective and efficient in every way that we can.

And we can’t talk about what we're trying to do here without sort of understanding the clinical task, right? The images have a purpose, not imaging for the sake of imaging. The purpose could be to guide a procedure, it could be to support a diagnosis, it could enable the acquisition of some sort of measurement information, various types of tasks depending on whatthe clinical task is.

And there are imaging challenges to achieve these particular goals, right? Getting the right view can be a challenge and getting a good quality image can be a challenge and so

the question is how do we help in that?

So obviously we have to have a purpose on one end, but we also have to think about this problem in the context of the user, right? And we have a broad spectrum of users, there's variations in terms of users' ability to be effective and efficient, and this could be a class -- and so we know that they're going to range in some sort of distribution and this could be a class of trained -- these users could be trained sonographers and they're going to have variation. These could be nurses and they're going to have variation. These could be non-clinicians in a home environment, right, and they would have some variation, as well.

And there are things that are going to cause -- there's nonuser factors that are going to cause this entire set of users to be more challenged or more enabled. The difficulty of the scan is one of those. Not every task that you're trying to perform in ultrasound has the same level of difficulty. The patients themselves may provide additional challenges with respect to achieving the task at hand, and then the capability of the equipment may be particularly well designed for the task at hand and provide some ability to shift effectiveness and efficiency in a given direction.

But if we sort of fix that and we say okay, we have a particular exam that is a particular difficulty, we have a piece of equipment that provides a certain level of value, the question then becomes where does an individual user fall on this curve, right? Individual users are going to fall in different places. And if we think about what's dictating where an individual user falls on this spectrum, for sure training and experience come into play here, right? And so there's been some focus and talk about how do you provide that good training and, you know, it's absolutely about trying to move people's ability to be effective and efficient with the technology. And the goal of AI is really to help this individual user, wherever they fall on that spectrum, to move in the right direction, right, either in terms of effectiveness or efficiency or ideally, both, right?

And then if we think about this in a broader sense of a whole population of users and if the goal is to move them all in the right directionas we help this entire group, it should be possible to make the outcome of that entire group to improve their effectiveness in a way that draws out a higher level of consistency and reproducibility among that user group. And if we think about what we're trying to achieve with AI from a user perspective, it is this kind of individual help and to drive the population of use in a positive way and a more consistent way, which certainly has value.

We've kind of talked about the users and we've talked about the outcome but then -- so knowing those things, what are the types of ways that ultrasound can provide guidance, right?

Getting to the desired view is a key aspect. And there are many ways that the ultrasound machine can provide some level of guidance in that activity, whether it's real-time detection of pathology as the scanning is happening. We've seen examples of segmentation happening in real time, that can provide a level of guidance segmentation, potentially, with labeling, as well. And then we saw a very nice example from Caption about guiding probe movement by recognizing the anatomy relative to the targeted view, right, so that the ultrasound machine knows what the target is and can understand how to help make that translation. And, you know, that can be combined with some sort of an indicator, you know, the red-yellow-green in terms of where you're at in that process of getting to the view, of course, that kind of real-time feedback can be very important, as well.

And if we think about getting a good quality image of that view, obviously there's a lot going on to improve the core image quality of ultrasound and AI will play a role in that, as well.

Image recognition, so knowing what's being scanned to drive better imaging

settings, being able to interpret the image that's coming back and making adjustments automatically for the user. Ultrasound has a lot of controls and so forth, traditionally, that allows a skilled operator to get to a good endpoint, but the idea, of course, is to automate that process as much as possible so that, you know, a less skilled user can achieve it and a more skilled user can also achieve it perhaps more efficiently.

You can also think of scoring an image as a form of helping to guide to a good quality image, as well, right? The idea is, is not to necessarily get the perfect image, the goal is to get an image that's going to solve the problem and how do you know that you're there or how do you know that you could work harder at it?

And there's far more examples of each of these, I'm going to go through a few specific things with respect to getting to that desired view.

This is a particular clinical problem where we have a product offering today and the objective, the clinical task at hand, is to make some measurements in the fetal brain or the fetal head, more specifically, and the idea is that to do these measurements you need to identify the specific plane where it's appropriate to perform this measurement. And the approach for this AI-based solution is rather than having the person scanning try to acquire those planes, instead a volume is acquired of the entire fetal head and then the AI is used to detect those planes within that volume and then as a next step in terms of the analysis or the interpretation, it goes ahead and does the automated measurements on those planes, right? This is targeted at somebody who would normally perform this type of an exam, but this concept of maybe targeting a volume versus a plane can certainly help for a novice user, as well, or even a non-clinician operator, potentially, in terms of reducing the specificity of the activity that the user needs to perform successfully.

This is just an example of real-time segmentation and real-time segmentation can certainly provide guidance in many ways; it can help to highlight a particular structure that's

of critical importance or provides a good landmark for a user who may not be as familiar with the area that's being scanned for a given exam, but I really show this to think about another piece of this. If you put an ultrasound image in front of an operator, there's really two goals in that. Earlier, we had a very good description of the goal of how the human brain is using the information that's there to be able to guide the positioning, but the other thing that an operator is often doing is interpreting what they're seeing, right, they're looking for anomalies as they're scanning.

But if you start to put the technology into the hands of a non-clinician, well, the non-clinician is not trying to interpret the image, they're only trying to guide. So then one could question is the image really the best tool for guidance or is some cartoon version of the image or something that feels more like a video game more appropriate for the guidance task when there isn't that dual interpretation or analysis task going on by the user?

And then earlier we talked about the fact that images have a purpose, right, and if that purpose can be performed in real time in a way that the operator sees that purpose happening, that sort of real-time analysis can provide feedback to the user about whether the view activity, view portion, identification of the view portion of the exam has been done successfully.

This is an example of something that we're supporting for the ER to be able -- when a patient comes in, to be able to very quickly determine if the patient has a shock or not and the -- this is two out of three of the tools just being shown -- where the idea is that you scan to a particular view and then the AI-based tool will perform the measurements in real time. If you think about the measurement itself, because it's being performed in real time and the operator can see that, it provides a level of confirmation, at least, that the view selection is adequate to the task at hand, which is why it's important to think about it not just as an independent task, but as a complete system. We were talking before about

adjusting, identifying a view, and then the analysis piece because they can be virtuosic.

And maybe just one final thought is that in an ideal situation, somebody who's been using AI assistance would actually improve their skill level even if the AI went away, in the sense that the AI would be essentially performing a level of training and helping an experience that would help the operator to improve their skill set. This may not be appropriate or possible in all forms of guidance, but it's something certainly to keep in mind as we design things, as well.

So just kind of circling back, you know, my goal was to give you some ideas on how ultrasound can be that intelligent assistance both for clinicians who need to do the scanning and the interpretation, but also you can see pathways towards assisting non-clinicians in the acquisition and guidance of an exam, as well. Thank you.

(Applause.)

DR. VAEZY: Thank you very much, Dr. Washburn. And this brings us to the end of this session, thank you very much. Thank you for your participation and thank you to all the presenters.

At this point, we're going to have a 10-minute break and then after that, we'll have presentations on the clinical realities, if you will. Thank you.

(Off the record at 10:03 a.m.)

(On the record at 10:33 a.m.)

DR. GARRA: Good morning, everyone. Welcome to the next session. I'm Brian Garra, a radiologist with the Office of Science and Engineering Laboratories at the FDA, and this session is going to change the focus a little bit from the people producing the products and developing the products to the people who will be using the products. We have both

physicians and patients and sonographers involved in this session.

I'd like to introduce our first speaker, which is Helen Feltovich. She says it's 20 years ago and I'm too old to remember, but I was a radiologist at university when she was a fellow, so we have a long association together and she's going to be talking. She's a neonatologist/obstetrician and she's going to be talking about the use of AI from the clinician's perspective and since OB is one of the primary applications of ultrasound, this is very appropriate.

DR. FELTOVICH: Thank you, Brian. Well, thank you so much for having me here. Here is my thanks to the FDA for this workshop. There's a lot of hype and fear and crazy stuff going on with artificial intelligence to us lay people, as you well know, and so these are just a couple of quotes that I pulled out of very, very recent publications. And I just want to say thank you for having this workshop to make sure that we're all on the same page.

As I just discussed, there's a lot of hype and promise and all these kinds of things around artificial intelligence and it seems like people fall into one of two camps, either they think AI is this great big monster that's going to take over humanity or the sort of goddess/savior omnipotent part of it. So here is AI as a job-stealing monster.

At the CS show last year, this woman, Bridgette Carey was one of the reporters and she was reporting on one of the feature products and she said this is amazing, you know, there's an ultrasound equipment that can be used from anywhere with software that's smart enough to help guide someone, even a layperson, to capture the images correctly even if you're not trained to do a proper ultrasound. She happened to be pregnant at the time and she said this is awesome, here I am, let's look at my belly. And she was really disappointed because the people at the booth said actually, that's a job best left to your own doctor. So what did we obstetricians think about that? Whew, we're safe but those

poor radiologists, they're really in trouble from this job gobbling monster, right?

And this is why we think we're safe. I'm going to show you a couple clips about how much fetuses move. Those little stinkers are all over the place all the time. It's like half my life is spent just chasing fetuses around a screen.

DR. FELTOVICH:

DR. FELTOVICH: I'll go to my bottom line right away. So the reason I was so excited about this opportunity and so grateful to be invited here is that as a person on the front lines looking at patients every single day with big, big problems, we're going to get to this in a minute, but I'm going to feature two obstetrical problems that I deal with every single day that are big deals, tragic, and I really, really think that AI could help.

What I'm hoping for and my e-mail is on my very last slide, what I'm hoping is to inspire some of you very smart industry people, researchers, something like people, anybody, your 13-year-old brother, Isabella, who already does coding, you know, somebody to really get invested in this problem of fetal imaging because there's a lot going on in that space and we give so much money and time and attention to, you know, admittedly extremely important problems like cancer, heart defects, you know, things like that. But I'm telling you what, everything starts in the uterus. These are problems that happen in pregnancy. This is awesome. I get to be on my pedestal. Problems that happen in pregnancy affect all of society, it affects our bottom line, the economic health of societies and everything. Everything starts in the uterus. I'm really hoping to get people super motivated about a couple of obstetrical problems.

This is the bane of my existence. You will notice that -- I'm not faking this, the transducer isn't moving, look at the top. Those are fetuses. Look at the one on the left. That little person is laughing, like actively giggling, do you see that? I get this kind of like sarcastic stuff from fetuses all day every day.

(Laughter.)

DR. FELTOVICH: And this is exactly the problem with AI, this is hard to capture. Okay. Now, the other side of it, AI as omnipotent goddess. So more than 10 years ago now one of my very respected colleagues got all excited about AI. He said as long as enough data are available, computational analysis using the technology will extract any and all regularities, patterns within the data and formulate a solution.

Now, this was awesome because back then the proposed advantages were that it wouldn't require a priori knowledge or any sort of predefined algorithms, no preexisting biases, you could just send the AI in and golden, everything you needed would come out of it and it would save the world. It was kind of like a holy grail situation.

He said it is astounding to think in the year 2006 obstetricians are still struggling to understand the cause of labor, maybe computers can help us solve the puzzle of parturition. Okay, well not yet.

This is a very depressing article from just a few months ago in November on World Prematurity Day that shows us that preterm birth rate is worsening. Not only not getting better, but worsening. And 10% of infants are born prematurely and I just told you about the severe consequences of that. That's my love, that's the focus of mine in Tim Hall's research lab, we look at prematurity. And I see that as one application of AI, but we'll talk about some others. Anyway, we're not there yet but I think we could get there and we need to get there. So where are we today? Wow, it is stuck again.

(Laughter.)

DR. FELTOVICH: And this is not even a video. Why does this kind of -- all right. My daughters are going to love this so much because they think that I have a magnetic field around me. They're always fighting with each other about who has to be my tech support and they think I have an actual magnetic field around me that makes -- a second, one

second, okay -- that sort of jinxes everything. Which is really funny since my research is in a medical physics lab, which I love to say because it makes me sound so smart.

(Laughter.)

DR. FELTOVICH: Truth is, I could barely pass a physics test, probably. But I have tons of people, you heard from Tim Hall yesterday, we have a great lab. I'm on the board of the directors of the Perinatal Quality Foundation with Jean Spitz, as you're going to hear from in a moment. I'm surrounded by so many amazing people that I don't have to be that smart, fortunately. Okay, all right.

This is where we are today. Fetal imaging today, where have we come from? This is a very, very, very rough idea of the evolution in fetal imaging. So almost 20 years ago people started looking at biometry and basic anatomy of the fetus, like you just saw some in the previous presentation, we measure the fetal head, the abdomen, the femur, the thighbone, to get a sense of the estimated fetal weight, are fetuses the estimated gestational age we expect, are they growing the way we expect, things like that. And now actually, most of the high-end ultrasound systems have some sort of automated anatomy visualization on them like you just saw, which works variably well, I'll be honest.

People have also been starting to look at acquisition guidance and user training. There's a whole bunch of software coming out right now that signals the sonographer or the scanner to let them know hey, you're in the right plane, this is the right way to look at this fetal heart or this fetal brain, whatever the structure is, and some of them have software user packages with them for training and this is kind of growing in concordance with all the simulations, research, and applications that are going on in medicine for training of our residents.

More recently, people have been trying to look at fetal anomalies which, interestingly, or interestingly to me, anyway, is affected by some of the things that we were

talking about yesterday in terms of the generalizability. It works really well in one place, doesn't work so well in other places, you know, over -- is that over-training, under-training, I don't know, but the same problems that were discussed yesterday, we are seeing some in the obstetrical space, too.

And by the way, I do not propose to be an expert in AI, I certainly know almost nothing about regulatory processes or anything like that, but I certainly can tell you that there are some problems that right here, right now could be addressed in obstetrics. Am I getting everybody, like, excited enough about OB? Yes. I have Isabella excited. If you can excite a 15-year-old, you can excite anybody.

Okay. So the most important development, to me, has been that we've come from this place where we thought in 2006 that a priori knowledge and bias were not a thing and now we do recognize hey, there really are -- you can't just like put an ultrasound transducer down, send a beam in and expect awesome things to come out. It needs some guidance, it is certainly susceptible to bias, unconscious bias. There's a bunch of this in the fetal growth space that we won't get into, but it certainly is susceptible to bias. I think that leads us into where could we go with fetal imaging tomorrow now that we have this really much more sophisticated understanding of it. And I like to think about a culture of AI, right? I like to think in terms of what is culture and what is a cultural drive? By the way, this is a great book. Everyone should -- everything should read it, some really fascinating -- *Darwin's Unfinished Symphony* came out a couple of years ago. Point is that this author, like, sort of turns evolution around. They say, you know, it's not our large brains, intelligence or language that gave us culture, which is the prevailing thought, but rather our culture, each other, our organization, are working together that gave us the large brains, intelligence and language that we have.

Here's the question and the proposal: How can we think about cultural drive for

machines? How can we use the large brains, intelligence and language of machines to identify and rapidly implement novel approaches to make our health better? I mean, really, from my perspective in terms of the monster/goddess thing, I don't really care if it's machines or humans that are in control as long as people's lives get better. There's a communication there,

Okay, so this is, like, lesson number one when you are in residency or in sonography school or in nursing school or anywhere near any sort of imaging equipment, lesson number one is garbage in, garbage out. So AI is not that holy grail. You cannot just send ultrasound waves into a tissue and have it, you know, come back and tell you everything that you need to know. You have to have very precise data acquisition.

So again, I don't know a lot about regulations or, you know, how we have to make this work within the FDA, but I do know that everything starts with precise data acquisition. And I have a lot of opinions about applications. I'm going to give you two examples of the challenge. As I was telling you a few minutes ago, preterm birth is my love because, you know, honestly because I just sit in the clinic too often with patients and all I can do is hold them while they cry and -- because we don't have good solutions for them and in fact, there is exactly one FDA approved medication for preterm birth prevention, it's a progesterone analog and that is probably about to get pulled by the FDA. There was an FDA committee hearing a few months ago that said the confirmatory test didn't show benefit, should we take this off the market? So not only do we not really have much, the one thing that we kind of do have is probably about to get pulled right here in Bethesda.

So that's why I'm using the survey as one example of that. There's a ton going on in that cervix, there are various reproductive tissues, there's a uterus, the membranes, the cervix, all of these have their own lives, all of these reproductive tissues change throughout pregnancy, they're very, very dynamic, and we don't have good ways to see them, so that's

one of my examples.

The other one is the fetal heart. And as I tell patients, the heart is the very hardest thing that we do because it's a moving thing inside of a moving thing inside of yet another moving thing. This is difficult and this is the reason that almost 60%, six-zero percent of fetal heart defects are missed in this country antenatally. This is a difference because of timing, like somebody was -- Dr. Martin was talking about this morning, timing. It matters where a fetus with a heart defect is born, right, because if they are in the right place things can happen very differently and it can be literally lifesaving.

So, by the way, 60% aren't missed in the office of, you know, a maternal fetal medicine person, we have 3 years of extra training in it, but in the general population, this is a real problem. Can you see the AI application here? To both of these, detecting fetal heart defects and also really understanding what's going on within the pregnancy tissues.

Here's the dream. The dream is that the fetus stays spine posterior like this with the heart up, beautiful views, and you have a quick -- this is not the same fetus, you'll notice the stomach's on the other side, but -- and you can just have the AI really quickly go down and make sure that the cardiac access is pointed in the right direction with respect to the stomach. That would be beautiful. It would be beautiful if all cervices were like this, you know, nice and long because a long cervix is associated with a lesser risk of preterm birth. That's not perfect and this isn't a preterm birth talk, so I won't go into it, but just take my word that this is a beautiful cervix.

There's so much information in that cervical microstructure about tissue properties that are making it beautiful, that is making this woman at low risk for a preterm birth, which we don't have a great way to get out right now, lots of people are working on it, but at this moment we are not using any of that information clinically. And if these were the two situations, a not-moving cervix, a perfectly behaving fetus, we would be good but, in fact,

that is not the situation.

Here's the situation. Fetuses move like crazy. This one in the middle, this is the best views we could get of that fetal heart and that woman's BMI is only 40. I mean, we have lots of women fifties, sixties, so you can imagine how difficult that is. And this third movie is not working, unfortunately, because it's probably the coolest. But this picture over on the right, that's a little tiny heart and that heart is pretty much structurally normal but you know what, it's outside of the chest, right? This is a lethal defect that, theoretically, you know how yesterday people were talking about context in that, so that is a lethal defect but if you were AI, how could you make sure that the structure was where it's supposed to be, does that make sense?

This brings us to the concept of clinical significance, clinical meaningfulness versus statistical meaning and somebody was talking -- a couple people, actually, were talking about that again yesterday. The little fetus on the right, you tell the patient the good news is, you know, the fetal heart looks fine, the bad news is it's totally in the wrong place. What if a system said hey, the heart looks fine, checks it off when you have this kind of a defect. These are the kinds of things that we're looking at.

The other problem with things like this is that they're rare and my understanding of AI machine learning is that you need tons and tons and tons of data. Then we're talking about things like data augmentation or how can we work with some of these rarer conditions?

Okay. And the cervix is even a bigger nightmare. Remember I showed you that beautiful cervix that was just behaving itself and staying just nice and long and still? This does not always happen. These are three cervices, these are transvaginal images, three cervices from three different patients. Look at how they're moving, right? Look at what those tissues are doing. These are all patients in about the middle of pregnancy.

The one on the right, there's hardly any cervix there. That baby looks like it's going to come out any second, like on the screen, right? And this person has twins, which is actually an additional risk factor for preterm birth. The two patients, the one in the middle and the one on the left, those cervices are problematic, you know, they don't look like that cervix that I showed you at first, but they are -- they're not as bad as the one on the right. Can we all agree with that, even if we've never -- I mean, Isabella, you can see that. You can even see that, right? You've probably never seen a transvaginal image of a cervix before.

The problem is that -- that was a good guess. The problem is that we don't know what's going on in there, there's so much valuable information that we could get at. You know what was the difference between these -- these are all patients of mine. Oh, by the way, gave me permission to put their cervices up on here de-identified.

(Laughter.)

DR. FELTOVICH: These patients on the left, their cervices felt soft to my finger. The one on the right, her cervix felt stiff. The one on the right is at way higher risk of preterm birth. She stayed pregnant for 12 more weeks and delivered her just fine twins. The two women on the left had had babies before, they only have a single ton (ph.), they were at super low risk, their cervices don't look as bad, one of them delivered a few hours later and the other one delivered a couple of weeks later, both lost their babies. Why? I don't know, but we need to find out and there's a massive amount of data in these cervices that I know that we can mine and figure out. I'm getting lot of help from Tim Hall, we're doing quantitative ultrasound to try to figure out some of these things but it's slow, we need AI, we need more people in the preterm birth space get inspired.

Okay, low-hanging fruit. So I just told you about some really sad things, I'm going to give you some low-hanging fruit even though there's a little bit of a trick in it, but I have worked with -- going places, training people, you know, the usual model, people in the field,

these people all gave me permission, too, by the way, but these are women in Bangladesh in a rural area. Rural enough that if one of these women lined up to see the provider right there has a placenta that's over the cervix and they don't know about it before she goes into labor, she is dead, okay, so that's how rural this place is. Time matters. I have gone to places and taught the providers there, the midwives there, how to do scans, how to look for these things.

And by the way, I have access, if anyone's interested, to still images. I mean, that seems like a really easy AI problem. Still images to look at growth and things like this, my e-mail is on the last slide, let me know if you're interested in such a project.

But what's really difficult is the live things that I was talking about. Okay, so there's your low-hanging fruit, easy project, lots of hope, here's your trick. This leads us into the untrained user. This is where the money is. This is where, as you heard this morning, people need to able to do these where they are, in their homes, in these rural areas, whatever. And because of these women that I trained on the left, this was just a few years ago, half of them are gone. Half of them are gone already. And in many places it's not acceptable to look different and -- I'm distracted because I'm running out of time. But in many places it's better to fit in. If we can work with these midwives that sometimes are not even literate, right, and put into their hands the kind of things that they need to find out really simple things like where's the placenta, how many fetuses are in there, that is literally lifesaving, right, because time matters.

The kinds of things that you have to deal with, with the untrained user is where do they scan, you know, you have image optimization, forget about that, that has to be all internal. They have to know where do I scan, how do I scan, how will I know when enough data are acquired, you know, these are the important points.

In summary, how do we get real artificial intelligence? I think that tenets of a

culture that support what we want, which is identification and rapid implementation of novel approaches to make our lives better.

Principle number one, garbage in, garbage out. Principle number two, we need to think about the clinical relevance and the context, you know, not just with that heart but will all kinds of things.

There's marked biological variability, as someone said yesterday. I think it does seem impossible to train for every single condition, so what do we do about things like data augmentation and, you know, adaptability and flexibility?

Awareness of biases is important, unconscious bias, like we talked about a few minutes ago.

Generalizability of the findings that has to do with under- or over-training the algorithm, there's -- I mean, that's like two days of a conference itself to talk about that.

And then, finally, access for everyone, everyone everywhere at the point of care. And then we will think that AI is -- no, I'm just kidding. I won't go there. But I have a lot of hope in -- I have a lot of hope in humans' ability to figure these things out, to really make life better, not just for my pregnant patients but for everyone.

Thank you so much for listening and caring about this problem.

(Applause.)

DR. GARRA: Thank you, Helen.

The next speaker is Jean Spitz, she's a sonographer, she's executive director of the Perinatal Foundation, Perinatal Quality Foundation, and there are very few people that have been more instrumental in guiding sonography than Jean. She's a representative of one of the major players in current ultrasound and that's the sonographer community, so I'm happy to have her here to talk about the perspective of the sonographer. Jean.

MS. SPITZ: Good morning. Thank you for having me and thank you to Helen for bringing up some of the problems that we deal with every day. I want to talk about artificial intelligence and the art of sonography.

MS. SPITZ: I want to thank the FDA for inviting me as a sonographer. I think it's important that we're part of the conversation and I hope that I'm the first of many that will be here in this conversation. You have it there but you don't have it there. Okay.

(Pause.)

MS. SPITZ: Okay, okay. Now, I'm talking as a sonographer, I've actually scanned patients for over 40 years and I taught sonography in a baccalaureate program for 30 years. And currently, I am executive director of the Perinatal Quality Foundation. I'll be talking about all three aspects of my career and will be talking from that perspective.

In teaching sonography, we always taught that sonography was the art of drawing with sound, that's basically what the word means, and our goal was to create beautiful images with a clear message or a clear interpretation in those images. Many sonographers take this to heart, they have their own idea of what a good image looks like, they have their own eye, they prefer images with a lot of gain or images with very little gain and it's hard to change that, I've learned, as a reviewer with NPQR. You can get six or seven batches from a participant, a sonographer, and ask him to turn down no gain and maybe after seven or eight batches, which is about 40 images that you've reviewed, they might consider that but that image looks good to them because that's their eye.

In teaching sonography, we really stress the recognition of image patterns and the production of images and I have to say with apologies that we really didn't teach protocols or standardization of measurements until maybe the late '90s or early aughts. I'm a firm believer that we're getting better, but if you put six sonographers in with the same patient

and those sonographers come from a different practice spot, you're going to get six different sets of images and you're going to get six different sets of measurement slightly with a different characteristic of the images. There's a lot of variation and this sort of falls into what I call the art of sonography.

PQF, that I work for now, the Nuchal Translucency Quality Review Credentialing Program, has tried for 12 years to standardize one single measurement, which is the depth of fluid in the back of a young fetus's neck. We require education, tests, and image review, web based; we apply standardized image criteria and the participants have to submit five images and then additional batches of three to five images to pass the image review.

Our reviewers have scored about 5 to 10,000 images a year. We have about 59 or 60,000 images that were scored since 2009. We scored before that, but we changed our criteria, so we only count the images with the current criteria, which was 2009.

Once they're credentialed, we actually get their measurement values from the laboratories that are using that measurement as part of a laboratory test for risk assessment. And we send reports to our participants telling them if their measurements are consistent in distribution with the reference curve, are they measuring at a range higher, at a range low?

Now this has been very successful in making sonographers and physicians more effective at producing this standardized measurement. If you look at our current data, we have about 88% of sonographers that are in range with their measurements that are provided to the laboratories. We have about 95% of the physicians that are in range, so we feel relatively good about our credentialing and our quality monitoring.

And you can also see that as we go -- that through the years, we reduce the numbers that are measuring low according to the reference curve and really increased the numbers that are within range. So we have created a more effective measurement with our

education and quality monitoring.

The problem we have, though, is a variety of images that we get. Some images have very high dynamic range of giving us a lot of detail and texture and they had -- both of these images are -- all of these images actually are good. All of these images probably would've passed our image review but when the gain is low, you can't really see the neural landmarks for the midsagittal, which is part of our criteria.

When the gain is high, as in these images, the boundary lines are much too thick and we know that that is an under-measurement, but our criteria says they should put the calipers at the inner edge of the boundary line and so they have a very thick boundary line and we know it's an under-measurement but it meets the criteria. So the variety of images, as a reviewer, we would much prefer to see an image like this or like this consistently and we don't.

Again, you can see the problem with the thick boundaries. This is an image which was measured correctly. Keeping that same measurement, you can see that the caliper is now not where the criteria say it should be, at the inner edge, but it's sort of in the middle. I have another example of that here where you have to place a caliper in the middle of the lines or under the lines, which is against the standardized measurement in order to get the same accuracy with the measurement.

Another problem, and you can't miss it, is echoes within the space. Now, humans can sort of recognize, you can't see this very well, but they put the caliper here, they ignore the reverberations, and they put the caliper down here, they ignore that echo within the space, so they ignore the reverberations, they ignore the echoes and they correctly place the calipers.

I played with an SNFM recently, I played with some of the machine-learning measurement tools, and they just move the calipers in, you know, their

segmentation/classification of that boundary line is not accurate because if there are echoes within the space, it will close, they're not -- that and then moving the measurement in. So we get inaccurate measurements even with machine learning. You can see here that this is a machine-learning technique that was used.

So in conclusion, AI needs to help with acquisition of input images, not just with standardization and analysis.

I want to go through sonography and how we acquire images because I think some of it may be helpful and the first stage in sonography, much like the first stage in AI, is classification and you're scanning a very wide area. And it's relatively easy to pick up the heart because it's moving in a characteristic way, the fetus is relatively easy to pick up, fluid structures are relatively easy to pick up. It's not a surprise to me that that's where AI is going first because I lived to help students for 30 years and they can pick that up pretty easily. When you're watching a student trying to find the ovary or the pancreas, they'll be scanning and scanning and scanning and you'll see the pancreas and you finally just grab their hand and tell them to point and press and it will come out. Classification is going to require supervision, particularly in the soft tissues, and maybe color to help pinpoint the regions of interest.

Segmentation. Once you grab their hand and apply pressure, you've basically selected a pivot point where you can search for the correct angle and the correct pathway to the area of interest. You have to compensate for large BMIs and we typically try to just move the fat pad away. I think IA can just reduce the noise, reduce the speckle and digitally move that fat pad away.

You have to assess the quality of the echoes throughout the region of interest, you have to move the pivot point or angle differently if you are losing echoes in any space and then you have to determine the acceptability of the image for a specific algorithm.

If you're doing a measurement, you need to reduce the thickness of a line, you need to reduce your dynamic range if you're looking for texture. If you're looking for more detail, you're going to increase your dynamic range, but you have to determine the acceptability of that image for that particular function.

Then optimization and detection, again, you can magnify the region of interest, fill in the boundaries, decrease thickness of the lines and blurring of the lines, optimize the dynamic range for the intended function, optimize the gain and the strength of the echoes, apply and optimize color and then apply the measurement criteria or algorithm.

I think if AI could provide real-time help in these processes, if we had something that would tell sonographers reduce this, increase this, that would help so much with standardizing our images, which is an issue.

My conclusions are that AI is very much needed in sonography for acquisition of images as well as for the analysis, that acquisition requires assessment of image quality throughout the process of image production, and that mimicking sonography imaging methods may provide a roadmap for AI acquisition.

Thank you again for having me.

(Applause.)

DR. GARRA: Thank you, Jean.

We're going to move on to the next speakers, but before I do that I wanted to thank Tracy Gray and Michelle Tarver in the Patient Engagement Office at FDA for providing these people that are going to give us such an incisive patient viewpoint.

The next two speakers are Joshua Basile and Harsh Thakkar. Joshua is founder of the Determined to Heal and SPINALpedia and Harsh is program coordinator at MedStar National Rehab Center and president of United Spinal. They're going to give us some patient

perspective on why AI may be helpful.

MR. BASILE: All right, so Josh Basile and --

MR. THAKKAR: My name is Harsh Thakkar. I'm actually the president of United Spinal Metro DC Chapter.

MR. BASILE: And we're excited to be here today to share with you our quadriplegic and paraplegic perspectives surrounding artificial intelligence, imaging, and the community and in home. To begin, we're going on our paralysis journeys.

I had a very happy childhood filled with sports, chasing girls, having fun with friends and family, and my life was flipped upside down when I was 18 years old, I was on a family vacation at the beach. On August 1st, 2004, while in waist high water, I turned my back to the ocean and a wave just picked me up, threw me over my boogie board, slammed me on my head, and shot up my neck. That day I woke up a C4-5 quadriplegic. And since then, I went back to school, I graduated from the University of Maryland, then went to law school and now I'm a practicing trial attorney doing catastrophic injury work for mainly the medical malpractice field.

MR. THAKKAR: And once again, my name is Harsh Thakkar, my journey was a little different than Josh's. I got injured in December 2005 due to violent effects. I basically got injured due to a gunshot wound when I was 20 years old and my life turned completely upside down. I didn't realize what my potential was until I started getting into adaptive sports. Sports was always something I identified with prior to my injury, so it was one of the first things I wanted to get back to, and I went to the Edinboro University of Pennsylvania on a wheelchair basketball scholarship, got a chance to play wheelchair basketball at a collegiate level and currently, I'm employed at MedStar NRH, which is the National Rehabilitation Hospital, as a program coordinator. And I like to think the only

thing that I can't do anymore is reach the top of a cabinet somewhere.

(Laughter.)

MR. THAKKAR: So basically, we want to talk about different aspects of what's AI and what the benefits could be, but we want to start off by addressing some of the struggles and current barriers. Is this better? All right, cool. Sounds better.

All right, so basically we want to discuss some of the struggles and barriers of current healthcare aspects and so what's happened with different specialized attention that we need, having a spinal cord injury, I know for me, specifically, I went to go see a urologist that that wasn't specialized in neurogenic bowel and bladder and when I went to go see this urologist, basically I was misdiagnosed. They were saying that my bladder walls were too thick for a normal person of my age range but they didn't recognize my specific needs. And so not having specialized care can be detrimental to persons with other health complications.

MR. BASILE: I've had my own troubles with imaging over the years. Four years ago, I was at the dentist's office and they wanted to get an X-ray of my jaw and my teeth. I went to go in front of the X-ray and my feet were sticking out too far where I couldn't even get to the imaging, so I wasn't able to get the treatment that I needed that day and I had to go to another office, to another dentist, to be able to get that. Also, I've had X-rays in the past where I had -- that when they put the lead plates behind me, they ended up pulling it down my back, scratching my entire back and causing wounds. A lot of people just don't know how to treat our unique bodies, the sensitivity of our skin, and it's just -- it's scary when you have to go places and you trust that people are going to take care of you and they don't. So there are a lot of struggles and barriers within the community to get proper care.

MR. THAKKAR: So along with the struggles, there are also things that, after learning about this, that excite us greatly about AI and having some form of imaging at the home.

The first aspect of excitement for me is the fact that it would be so freeing to be able to do things, to do things at a home level to prevent some of these things that Josh and I talked about prior to this. It would not only be convenient in the aspects of not having to leave our homes, but it would take away some of the secondary challenges that we have, such as transportation, and it would make it so much more convenient. At the same time, it would also save a lot more time in our personal lives because believe it or not, but we do hold full-time jobs and we still participate in full activities, but so it would save so much time in our lives to be able to get access to things a lot quicker.

MR. BASILE: Preventative care, for me it's all about avoiding secondary health conditions. Secondary health conditions are avoidable with the right approaches and the right understanding of how not to go down that road. The trouble is once you have the secondary health condition, it changes everything. You have to fully focus on it and if you don't, it gets worse and worse and worse. It doesn't go anywhere or it stays exactly where it is. The big thing for me is if you can prevent secondary health conditions, you can improve quality of life and give more opportunities to persons with the most significant disabilities in the community. So being able to bring this in the home, I think it's game changing to be able to get access to quicker, more efficient care and to do it on your own time and schedule. So, you know, how are people going to physically use this type of technology? So as a quadriplegic, I can't do it myself but with that being said, I do have my family members, I have my friends, I have my caregivers to be able to do it for me. But in order for me to do it safely, it's best if it can be done on my bed or in my wheelchair.

MR. THAKKAR: And for me, the perspective would just be a little different where independence is key in this aspect. There are plenty of people such as myself, who would want to do these kinds of things independently. And this is where, I think, for anyone involved in this, this is where design is one of the most important aspects of it because

making sure that the equipment is not too heavy, making sure that there are not too many moving parts, all of these different design aspects can serve as either barriers or hopefully with it being lighter and less moving parts, it can assist in making this a lot more functional for at-home use independently, as well as with assistance.

So for AI, basically what's happened in the past for me before is something such as I went waterskiing a couple years ago, my foot came off of the strap a little weird, felt there was something wrong with it, didn't know that I had a minor fracture in my foot until 3 or 4 days later and that was just because my foot was inflamed more than it's ever been before. And so again, something to detect some of these things early and quicker, I think is the key and basically if I would've gotten that information, if it was ready on hand, I would've been able to pull that information up a lot sooner than waiting 3 days to go to an imaging center, get the images, get them read and have it take so much longer than it actually can.

MR. BASILE: So I spoke with my dad about this slide, luckily, my dad has been a urologist for the last 30 years and I talked with him, I said how would you use this in the home or how would I use this in the home for my bladder and my kidneys and he specifically said, "Josh, this would be incredible to be able to monitor the kidneys, see if there's any backup of urine, " because if you have a backup of urine in your kidneys and it doesn't get addressed properly, that can cause permanent damage. And then in the bladder, specifically you can detect abnormalities, you could monitor the size of the bladder, thickness of the bladder walls, if there is any bladder stones, and there's just many different levels of the bladder and the kidneys to be able to get quicker access to care and I think that just really, really excites me about this.

When it comes to concerns of this type of technology reaching the community, you know, a spinal cord injury is the most expensive injury you can have in your life and it's a lifetime injury, it's a long-term injury, and that creates extreme financial burden on those

with spinal cord injuries and their families. And this type of technology cannot just be made for those who can afford it, we need to be always looking at these type of communities as a long-term investment.

Getting this into the homes of people will prevent secondary health conditions. Preventing secondary health conditions will increase quality of life, which is great, but from a payer's standpoint, this is going to save billions of dollars for the healthcare industry and the healthcare payers. We need to look at how we can advocate to those that can purchase these technologies, to get it into every single home for persons with spinal cord injuries and significant disabilities.

MR. THAKKAR: I'm pretty sure everyone's been evaluating some sort of safety concerns regarding all of these things, these were some of the things that we came up with. More specifically, the questions of what happens when you use this equipment too often incorrectly, those are some of the more prominent questions. But one of the most prominent questions that came to mind was basically self-diagnosing and what happens if someone once trained, even after they're trained, now they think that they're professionals in self-diagnosing themselves and how harmful can that self-diagnosing be. Basically, I think it could be very harmful especially in terms of if you self-diagnose incorrectly and we know plenty of people that go on WebMD as it is and try to look up things that are going on with themselves and end up causing some distress, end up causing some mental health issues that could be prevented for sure if done the right way; and imagine, basically, from our perspective imagine what happens when these devices, again, are over-used, are used incorrectly, making sure that there is some kind of safeguards regarding those things, as well.

So basically, when I was first injured, what happens is that there's not a constant wave of information. Information comes in, like, very infrequent waves. So once you're

injured someone tells you four things that you need to know, as you're leaving a rehab hospital someone gives you another four things that you need to know, and even when people do try to provide all the information at once, it's information overload for somebody going through something traumatic. And so to make sure that there's a constant form of information coming and I think having these kinds of devices would kind of assist with some of that constant information that would be provided.

MR. BASILE: It's so important that we understand who's using this type of technology and to provide the proper training to make sure that they use it safely and effectively. There's so many different types of learning styles and so many different types of people that are using this in the care-giving world. Unfortunately, you know, a lot of people think, you guys think get this into people's home, we train one person, we're good. But that's not the case when it comes to care giving. So many quadriplegics with a high level of injuries that have long-term care have 10 to 20 different caregivers in a year, so how do we make sure that we can actually train people quickly enough and understand exactly how to use this safely? Understanding the community is so, so important. They cannot be overlooked.

And the other aspect that you guys already note, we got to protect this -- all the health information that comes out of this, it needs to be sent to the proper places and stored properly and not used in ways that can hurt the most vulnerable populations that are using this.

MR. THAKKAR: So for this one, basically what happens when someone starts using this in their home, I kind of just wanted to reiterate some of the things I think that went a little ahead of the previous slide, especially when it comes to how information is provided. How information is given and ensuring that everyone's receiving the information as they need it instead of receiving all that once or once whatever information they've needed has

already surpassed itself.

MR. BASILE: So both of us are really in agreement. As soon as somebody gets injured, this needs to be used in patient rehabilitation hospitals to train people how to go home with the technology. And for me, for both of us, as chronic injuries, we want to actually bring this into our kind of protocols for addressing secondary health conditions. We're not going to use it to kind of change everything, but we're going to use it to improve and enhance our current approaches. We're not here to reinvent the wheel, but to create a better wheel or create a more efficient wheel and that's how we really want to focus on bringing this into our lives. So thank you so much.

The last thing that we want to end on, really, is to speak to you guys as researchers, as technology developers, the patient perspective is so important. Having us work with you at all stages of development, it's such an important voice to have at the table to create the best end product and especially when we then bring to the community, adopting these products, having the voices and seeing people using it and really just having branded ambassadors within the disability community, always seeing doctors and, you know, fake people, like there's just so often you're putting people able bodied and manual wheelchairs that are coming from the hospital on commercials and you see it and you immediately know that it's a fake voice just trying to sell something to you. But again, real-life stories, speaking with us. Harsh and I are here to help bring this to the next level and we want you to know that this is how you can contact us so we can be a part of this with you, a part of this journey. So just thank you so much for listening to us and I guess we're going to do panel with questions.

MR. THAKKAR: Yeah, thank you guys for listening. Hopefully, we offered some sort of unique patient perspectives to think about as you guys are in the development or whatever stages that you guys are in. Thank you.

(Applause.)

DR. GARRA: Thank you, guys, that was great.

We're going to finish up with Isabella and Amy O'Brien, who are going to bring another set of patient perspectives. They're members of the International Children's Advisory Network and I'll let them explain why they can use this at home. Isabella.

MS. I. O'BRIEN: Hi, everyone. I am Isabella O'Brien, I am 15 years old and I have aortic artresia, plastic bronchitis, Graves' disease, and liver disease. I am 15 years old, so that means I'm a sophomore in high school and it's like stressful enough being in high school, but when you have a lot of chronic diseases, it's an extra thing to be worrying about. If you can continue on your day at school, if you find it's just a little thing that your body's doing or if you need to go to the hospital.

I think this device would be amazing if it was a check for when I don't feel right or if I have, like, an event. If it's a check between appointments or just a checkup to see if everything's doing okay, like there doesn't have to be a problem. I also think that it would be great if it's for post-procedure, so if I can get out of the hospital earlier so I can get back to my life.

What has to work is I need to know that the information is available to my team but also secure so that not everyone can access it but the people who I really need to have see it sees it. I also need to know the machine is working, so to see the machine can like light up or do something or have like some type of indicator that it's working so that I know I did it right and that I'm not just messing everything up. I also want to be able to talk to someone if I'm really confused because sometimes things can be going wrong and it won't tell you what's going wrong so there needs to be something that I can call or maybe someone I could talk to.

Also, this can't replace my teams because I'm with them through everything and they know me and they care about me and something that a machine might see as not normal for someone who doesn't have any chronic diseases, it's totally fine and okay for me to be living with.

For design considerations, it should be easy to transport. I need to take this to school and I need to have it at home, even someone would need to take it to work, so something that's easy to bring around because your health doesn't wait for you to be done with school or to be done with work, but something that you can take everywhere. Also, again, there should be indicators to let us know if the machine is working and if I did it correctly.

Also, this is very important, but there should be a mute button for any alarms because if something is wrong with the machine and it starts going off in class, I can't do anything about it if there's no mute button and there's just this constant beeping in class and it really kind of draws attention and that's not what I'm trying to do. But also, if I'm trying to sleep and it's going off -- this has happened to me, I don't know if you can tell, but even trying to sleep and it goes off constantly and constantly, so you just shove it into a drawer and put it away, that's kind of unsafe because if you have an event while it's happening or something happens and you actually need it, it's away. Also, adhesives should not hurt or be as gentle as possible because if you're wearing something for a while or if you're putting it on multiple times a day, it can irritate the skin and that can be really irritating. Now not only are you worrying about your chronic condition but now you have a rash because of the thing that you're trying to help it with.

MS. A. O'BRIEN: Hi, everyone. I'm Amy O'Brien, more importantly, my official title is Isabella's mom. I'm a huge supporter of using responsibility and thinking about how we can advance technology, including AI, and able to make more equipment usable and suitable in

the home.

As a parent, caring for my child is my number one, it's not always easy to do. I'm a single mother, my job is what pays all of our bills, it also is how we have wonderful insurance coverage for our child, and she has a sibling. And I have it pretty easy because I can do my job via computer, so in many cases when I'm at the hospital I can still actually perform and be able to do my job. Every time I have to go to the hospital, I have to think about the care of another child and the impact of going to the hospital has on them.

I have it pretty easy; the volume of people that I meet in the hospital who have lost their jobs or have had to quit their jobs because they physically have to be present at work is just far too high, it is incredibly stressful to see these people in these situations. So whatever we can do to make it easier for people to do some of these tests at home, I'm absolutely not advocating not having to go the hospital because that's critical and definitely keeping my daughter alive, but there are times where we have to go to the hospital for a maintenance test or just because she's unsure if it's a normal body thing for a 15-year-old or something that's tied to one of her specific issues. If there are tests that we can do at home just as a safeguard and a check, that enables us to continue to go on with our daily life.

For many people it really is the security of their family and the welfare of other children and there are cases where parents do make the decision not to go to the hospital and it's not because we don't want to, but it's because there are other pressures in our world and we're constantly weighing different risks. In some cases if we can help people who don't have access to this care because of transportation or because they have other children at home or because of even just financial security, I feel that we have an obligation to do so.

So, you know, a key here, I think also is the fact that there's a lot of kind of research

and development going on, but accuracy and trust is huge to making this work and so, you know, you may look at statistics about what happens if something fails but in some cases if something fails it's literally the life of a child. I think the responsibility, the accuracy and trust are something that we parents and caregivers really need the industry to be cognizant of and make a huge priority to us.

MS. I. O'BRIEN: Thank you for listening and investing in our care.

(Applause.)

DR. GARRA: Thank you, all. Sorry, we're running a bit behind. I'm going to get Shahram up here and we'll get the panel going.

DR. VAEZY: Thank you, Brian. Yes, please, if the panel members could come up to the stage.

(Pause.)

DR. VAEZY: Okay, folks. Thank you very much. Thank you for a wonderful session, so far it's a wonderful day. Thank you very much for all the presentations.

So we're going to have the panel discussion. Since we are running a little late, instead of the 1 hour that was allotted for this, I think I just put in 40 minutes and hopefully we'll be able to cover a lot. Without further ado I want to get started, let me just start my --

UNIDENTIFIED SPEAKER: It's already there. I programmed that in.

DR. VAEZY: Thank you. All right, folks, so the topics for this panel discussion, this is going to be really on AI-guided image acquisition, image and signal acquisition. We just heard about the home use and the importance of that, so I want to get started with that. We heard from the patients, but also we heard from the experts, developers, on home use so -- oh, by the way, before I should -- before I go on, I just remembered that -- let's go through the panel members, everybody, if you could please introduce yourselves.

DR. MARTIN: I'm John Martin, I'm the chief medical officer of Butterfly Network.

DR. KING: I'm Randy King, I'm a program officer at NIBIB, part of NIH.

DR. HONG: My name is Ha Hong, I'm from Caption Health.

MR. TRAHMS: Rob Trahms, I'm a strategic architect for AI for Philips Ultrasound.

MR. WASHBURN: Mike Washburn, Chief Engineer with GE Healthcare.

DR. GARRA: Brian Garra, FDA.

DR. FELTOVICH: Helen Feltovich, maternal/fetal medicine doctor with Intermountain Healthcare and researcher with the University of Wisconsin.

DR. FRANK: Richard Frank, Chief Medical Officer of Siemens Healthineers and a representative of the Medical Imaging Technology Alliance.

DR. SAMIR: Anthony Samir, physician, scientist, and radiologist at the Mass General Hospital.

DR. SHABESTARI: Behrouz Shabestari, Acting Division Director of Bioinformatics Division at NIBIB.

MS. SPITZ: Jean Lea Spitz, sonographer and director of Perinatal Quality Foundation.

MS. I. O'BRIEN: Isabella O'Brien, cardiology patient.

MS. A. O'BRIEN: Amy O'Brien, caregiver.

MR. THAKKAR: Harsh Thakkar.

MR. BASILE: Josh Basile, founder of Determined to Heal and SPINALpedia and trial attorney.

DR. VAEZY: Thank you very much, folks. As I mentioned, the first topic is the topic of home use of AI-guided image acquisition assistance. Okay, sorry. Thank you.

Dr. Martin, you talked so passionately about home use, so I'd like to start with you and ask about maybe addressing the first two sub-topics in this area, what type of impact on the risk-benefit analysis could be expected for such devices and what approaches are

acceptable for patient consent?

DR. MARTIN: I'll do the patient consent one first because that's the easiest and I have the most experience with because as a surgeon, you live your life doing patient consent. And so I look at this exactly the same way, whenever we either institute care, whether it's a drug, whether it's a device or whether it's at home, I think you have the same kind of conversations with the patient and their caregivers of the risk-benefits of what we're trying to achieve. I don't really see this as anything different than we typically already do as a physician in healthcare.

Now, as to the first one, the risk-benefit analysis, we set a standard for ourselves, we would like the imaging to match the quality as if someone was at a hospital and a professional is performing that. We think that that's a fair standard. I think, candidly, we've had speakers at the end of the room tell you about the benefits of this and they're overwhelming. No one really wants to engage the healthcare system unless they absolutely have to have it and so whether it's improved patient satisfaction, the efficiency in the delivery of care, the reduced cost, the improved quality, and ultimately better outcomes, I think these are all measures that go into this and so as we balance -- if we can have this similar efficacy and safety, all the benefits trump everything else by far and moving into the home, I think, is something we should do expeditiously and thank you for making the case.

DR. VAEZY: Thank you very much. Dr. Samir, your thoughts?

DR. SAMIR: I think that, very much as John had mentioned, I tend to agree, the benefits are there. There are risks that I foresee, too, primarily around overuse of technology and diagnosis in circumstances where diagnosis is unhelpful. And I think that part of the organizational and regulatory framework that will go into the effect of a responsible deployment of these devices should include mechanisms to mitigate those risks so that we're not so overcome by our own enthusiasm that we forget that technology is

often a benefit and a risk.

Having said that, I share everyone's enthusiasm, absolutely. I can't wait until these tools are in the hands of patients and I hope that all of us will be able to be part of that journey and make it happen as quickly as possible.

DR. VAEZY: Thank you very much. In the interest of time, my apologies for kind of going fast through this, but I would like to move on and address the next two topics, how should the physician be involved, should they be the trainers and prescribers? And what type of image quality standards should be applied for such systems? Dr. Feltovich, you talked about this and I would like to start with you, please.

DR. FELTOVICH: Sure, thank you. Well, I actually really liked what Isabella said about how she wanted to be able to talk to a person if necessary and as I think I discussed a few minutes ago, I'm a big believer in AI. I do think there are some risks involved with diagnosis and that, but I'm a big believer in it and I have a lot of hope for it. But I think at the end of the day, that a physician does need to be involved for Isabella, 24 hours a day, for Harsh and Josh, 24 hours a day, for all of us for our conditions so that we feel like we have a connection and a person to talk to and a person to explain things to. And this ties into the risk of technology just fundamentally, which is that you can get all kinds of information and not all of it is clinically meaningful.

DR. VAEZY: Ms. Spitz, do you have some thoughts on this? You talked about your training program.

MS. SPITZ: I would agree. I think I look forward to the time when ultrasound images can be cleaned up by AI so that the risk of misinterpretation is decreased. You know, I've seen people who are not very well trained create artifactual problems, so there is a risk but I have great hope in AI, as well.

DR. VAEZY: Thank you. Again, in the interest of time, let's move forward. How open

is the community to adding diagnosis to image acquisition and what should the patient interaction be with physicians, how should scenarios with -- where diagnostic evaluation is involved be handled? So if I may ask Dr. Frank to comment on this.

DR. FRANK: Well, to reiterate my comment yesterday that patient benefit is paramount and we've heard some concerns expressed about the context of use and mitigation of risk and so on, it is much more than just the acquisition of the image and so to add to that diagnostic quality, not just diagnostic quality, the image, but interpretation of the image certainly is a bridge beyond that. Industry is open to that, but we need to take into account things like the context of use and the circumstances, not just the training and the acquisition itself.

DR. VAEZY: Thank you. And I'd like to ask our patient representatives here, how would you see this patient interaction with the physicians?

MS. I. O'BRIEN: I think that it's really important because sometimes if you get something, you can get results and not understand what they mean and sometimes a machine can try to explain it to you but everyone, understands things differently and to have an actual person sit down with you and explain to you what you have or what the results say and what that means, I think is really important.

MS. A. O'BRIEN: I'm just going to add to that, if Isabella went to get some of these tests done at a place that didn't understand her body and her current situation, some of these tests would actually freak people out and they have. So I think that interpretation, to me, is key where you may get kind of a read that is not normal, but it may be totally normal for her. So I think that's a real tricky, tricky spot where you need to have someone who has some kind of a history and perspective, and we've talked about this, I mean, we text our doctors.

I think the doctor presence, like, you want to be mentally connected with that

doctor having this communication. I don't think they have to physically be there, they could be on the phone, video chat, even texting, we've had meaningful dialog, so it doesn't make a difference, it just means like are you paying attention to me and these results.

DR. VAEZY: Wonderful. Josh, you had something?

MR. BASILE: Hello. There's something special to be said about the experience of getting healthcare and being able to have that interaction with a doctor, physician, to be able to have a good experience and motivate you next time to do it again. If you have a bad experience and it rubs off on you the wrong way, it scares you away from it; it scares you away from getting the care you need the next time around. So being able to create the better experiences, having the history, having the dialog be open promotes better care, better outcomes, and overall just a better quality of life.

MR. THAKKAR: I'm just going to add one thing real quick. We know these changes aren't going make like a 360-degree difference, like it's not going to go from like no access to like complete access, right? So like at the same time understanding that this is going to take like stepping stones to get that avenue of having these open platforms, continuous open platforms, I think is also important to kind of think about.

DR. VAEZY: Thank you very much. Please go ahead.

DR. GARRA: Yeah, Shahram, I want to make a quick comment. We have technologies at home already and I've seen what happens. A person does the examination, let's say blood pressure, they get an abnormal result, they start getting worried, they start taking it more and more often so you end up getting this overuse and then they want to talk to somebody about it and they're told you can talk to somebody in like 3 weeks. So what happens is during that interim periodthey get more and more stressed, they overuse it more and more, and so we need to have immediate access to advice, somebody to advise them that they respect to prevent the cycle from happening.

DR. MARTIN: Yeah, just make one addition. Try as I might to have a relationship with Alexa, it just doesn't work.

(Laughter.)

DR. MARTIN: She hasn't been able to read my body language or my level of anxiety along the way, although she is very helpful at times. I think at the end of the day we've done this journey before with many, many different technologies and we can all identify horrible case scenarios that can unfold. I think if we responsibly understand that that pot of gold at the end of the rainbow is something better and take a responsible route along the way, I think we're all going to be happy. We all have to accept the fact we're going to trip sometimes and learn from those, but I think this is a journey worth taking. We got into healthcare, all of us, because we care. We don't want to take that out of this.

DR. VAEZY: I think the take-home message, for me, the key is the balance between being enabled to do some of the image acquisition, signal acquisition and then having that wise, if you will, or that consulting with a physician or a healthcare professional. The last time I was in a doctor's office, I complained about the fact that instead of looking at me, he was looking at the computer and everything that I was saying he was just entering into the computer, so that personal touch should never be forgotten.

Okay, so Topic 2, Adaptive Technologies for Guided Acquisition. What type of strategies are being envisioned for managing device performance and the associated changes as a function of time and software/hardware improvements? I'd like to start with Dr. Washburn, if you could comment on this. Do we need performance metrics for continuous evaluation of safety and effectiveness of the AI-guided acquisition?

DR. WASHBURN Yeah. So I think there's been a lot of talk about market surveillance over the course of various presentations and I think that is certainly a key component to be able to look, and there was also some talk yesterday about the fact that having a

comparative standard, you know, what is the gold standard, how is it done today and how is it measured today, and often there isn't necessarily that information and so I think, depending on the intended use and the particular problem to be solved will definitely dictate the kind of solution that's needed in this space.

DR. VAEZY: Thank you. Dr. Hong, your thoughts about future changes to a device.

DR. HONG: I think that the problem has remained the same. So Richard Feynman, one of the famous physicists, Nobel laureate, once mentioned his own way of solving problem, Feynman's problem solving strategy, it has three components. Number one, write down the problem onto paper. Number two, think about the problem really hard. Number three and solve the problem.

(Laughter.)

DR. HONG: So there's a point, actually, which is to think about the problem really hard, so we all want to solve medical problems and it's very important. Oftentimes engineers get lost in the technical things but at the end of the day it all matters, the medical, the clinical problem we got to solve. So we got to find out, we got to be very clear about the aim, which problem, which clinical problem you want to solve, and from there we got toderive clinical metrics that are causally related to the clinical problem you want to solve and then from there we want to have engineering metrics which are causally related to the clinical metrics. The link between the two is not always obvious, it can be quite lengthy, but as we close to the link we can breach between the gap between the two, that will be better. If we have now engineering metrics, then, we can now start to think about making some AI models and some computer systems based on the optimizations.

Starting from there, now comes the aspect of data. Now, these days, these new computer models are data hungry, they have millions and millions of parameters oftentimes and because of that, we want to have as much as possible data. But at the same time,

quality is often very important, the quality there has several aspects. We want to cover, as much as possible, natural variability, you know, you can encounter in clinical settings. The more coverage, the better. The more patterns you see, the better. It's like same as the human. So for example, like covering as much as possible variability from -- variability from patients, operators, specific hardware, even the hospital if you can cover as much as possible data, that will be better. So quantity and quality of the data.

(Crosstalk.)

MR. TRAHMS: Okay, yeah. Well said. I think Dr. Martin also said it, that the problems that we're talking about, how to actually, you know, develop new capabilities, test those capabilities, create the metrics for those capabilities, validate in clinical context, get enough samples, statistically, you know, variant samples to solve the problem is really not unique to AI. AI is just the latest technology that we need to apply good problem solving skills and solve the problems, as Dr. Hong said.

DR. VAEZY: All right, thank you. And may I ask our patient representatives, also, do you want to know about which version of the software is being used? Do you want to know the performance metrics for the latest version or is what the physician says that this is working is going to be satisfactory?

MR. THAKKAR: No, no. I think that's a very interesting question. It kind of goes back to no one reads the agreement before they click yes.

(Laughter.)

MR. THAKKER: So it's kind of one of those, I mean, you would want to know, and I think the information should be readily available. Whether that's presented at every case of it can be dependent on what exactly the changes are within that software or hardware.

MR. BASILE: I think open communication builds trust, so I think like Harsh said, making the information available but also being able to communicate it in a way that it can

be digestible and not just go over someone's head, that's what I would say about that.

DR. VAEZY: Okay, thank you.

MS. I. O'BRIEN: I would say that it's really important and like that it's available in some way because maybe not everyone wants to know, but if you've been using a similar device or if you use a different version that does a similar thing, you kind of want to know what's different and why this one's better than the last one. I've had different event monitors and they kind of explained well, why is this one better and why are you using this one now instead of the other one, and I think that's really beneficial.

DR. VAEZY: Great, thank you. I'm going to move on to Topic 3 now, Patient and Clinician Perspectives. What would be the perception of patients and clinicians regarding the impact of AI-guided acquisition on workflow enhancement? Is the impact welcomed? What metrics should be used to evaluate the impact? Perhaps we could start with Dr. Feltovich, please.

DR. FELTOVICH: Thank you. Well, I think that, obviously, I'm very open to it and really excited for AI. I think this could markedly change our workflow so that we have more time to talk to the patient. The story you said a few minutes ago about how the doctor was at the computer instead of looking at you, I think that's happening more and more and it's really tragic because the one thing that we bring as humans to each other is connection and as we just head from our patient representatives down here, that's critical.

But I think that- I have that slide that's an AI is monster or goddess- I think there actually is a lot of resistance in the medical community because of fear that our jobs will be replaced if we are primary imagers. I think the truth is somewhere in the middle, I think that you cannot replace the human component. I mean, actually, I don't know. Maybe humans can be -- well, I mean humans. Maybe computers can be more empathetic than humans some day, I mean, who knows what's going to happen? Who would've thought

we'd have Alexas, you know, even 20 years ago, who thought that could happen, so who knows?

But I think that, again to answer your question, there's a lot of resistance but I think there's also a lot of hope for making things easier and better for us so that we can all get what we want, which is for the provider to be able to interact and for the patients to be able to interact with us and connect.

DR. VAEZY: Thank you. Any other thoughts here?

DR. FELTOVICH: Can I say one more thing about metrics? Up there, what metrics should be used? One thing that occurred to me when Brian was saying the thing about people take their blood pressure over and over and get more and more anxious, I think one thing that we need to be aware of around this introduction of AI into our homes and clinics is sort of the mental health aspect of that, that's something we haven't talked about at all, but it just occurred to me when he said something about the amazing anxiety, you know, and if a machine tells you that something is wrong or a machine tells you that something is right, like with that heart case I showed, you could go down -- you know, as a patient you could go down an entirely wrong pathway, so maybe we need to think about some mental health services around this.

DR. VAEZY: I'll come to you. Go ahead.

MR. TRAHMS: It's been mentioned multiple times both yesterday and today, I think Dr. Feltovich is correct, it's about how the clinician is spending their energy and it was basically using the analogy of the pilot in the plane, right? You know, you can spend a lot of time with the tedious nebology (ph.) of a device or you can spend time with a patient, right, and it's really about -- I think that's the promise of AI is to allow to really allow the clinician to focus their energy more to the patient.

MS. A. O'BRIEN: Yeah, I just want to kind of add another perspective here with

regards to quality metrics. It's like we love our doctors, I mean, they're a part of our family, we've known them for ages and we always talk about kind of the quality of life for patients. I'm incredibly concerned about the quality of life for all the caregivers. The hours of work, the frustration and stress trying to get good information is outrageous with regards to the impact that it has on the people who care for us. We need their brains to be working, we need them to have wellness, we need them to be healthy in order for them to be solving more complex problems, and so I do think that we need to think about that other perspective of how do we actually help everyone in this ecosystem have quality of life in order to make it much more productive.

DR. MARTIN: I look at these devices and I want the machine and the artificial intelligence to help me make a diagnosis quicker so I can spend my time helping the patient understand, to have confidence that we know what we're doing and give them an idea of what's going to happen for them in the future and what those options are. So that really is what makes the difference. If I can get to the answer quicker, I can spend my time on doing those things which are really human in that interaction that matters and I think that's where the real potential exists.

MS. SPITZ: There are other groups that are using technology to replace hospital visits in pregnancy and I've been involved in discussions of the equity and we need to keep that in mind. We need to use the at-home to increase access and not have access to a physician a luxury as opposed to at-home care, so that equity has got to be part of our goal.

DR. SHABESTARI: This is for the panel, right? I wonder if this is really enhancing the workflow. I heard Dr. Martin said he would spend more time on the patient, but is it that is going to be true or is it going to be more patients to see, so how are we improving -- in a short range or in a short time there's going to be a lot of information to kind of communicate with the patient a lot of information to understand, so it's really not at the

short time is not improving the workflow.

DR. VAEZY: Thank you. Go ahead, Brian.

DR. GARRA: Just one quick comment. We need to incorporate this into the educational program. Physicians, when they're trained, have so many technical things to learn that they forget that their real value is being a human being and caring about their patients. And as we move to an era where the technical things can be taken over by machines, we have to move along the educational system to make sure that trained -- new trainees learn that message and learn the skills that they need to show that they care.

DR. VAEZY: Dr. King.

DR. KING: I also think we need to be prepared for massive paradigm shifts and what kind of data are we going to be taking when we go from collecting discrete data to continuous data. So specifically speaking about the point of care, are we looking at blood pressure, are we looking at blood sugar, are we looking at cardiac events? If we're looking at blood sugar, is A1c really the right thing to be monitoring? We're going to get massive amounts of data, we're going to get it very quickly and we're going to have to look at an overall picture of the whole patient. Are we going to have to change our metrics and what we're going to be monitoring? So I think that we need to be prepared for that, too, as we move forward.

DR. VAEZY: Great, thank you. Since, Dr. King, you have the podium, I mean, the mike, I want to ask you about this topic. Since you talked about the NIBIB as technology developers, I want to see what type of vision do you have in the overall guiding principles, specifically when it comes to technologies that have an impact on the diagnosis and so anyway.

DR. KING: So what's very important for us is the datasets that are teaching these neural networks and guiding the AI and the imaging data for the last 50 years, a lot of it's

going to have to be thrown out and we're starting over developing new imaging databases to train the neural networks with the correct resolution and the correct parameters we need. So part of this is the preclinical participation in large databases like the ACR database. We're developing imaging standards that you can collect over time and everyone can use the databases to train their neural networks. That's one aspect.

The second are panels like this with the patient perspective where I learned today that we're dealing with flexible transducers that can be applied to the skin, do continuous monitoring or implantable devices and we're learning here from patients what's important is the adhesive applying the transducer to the body and how sensitive is their skin and how they interact with our technology and how important that can be in the preclinical research side when you're developing the technology. The patient's interaction with that technology is just as important as developing the technology because if a patient can't use it, it won't be used.

DR. VAEZY: Thank you. Dr. Shabestari.

DR. SHABESTARI: I don't know if this is a -- my job is to look at the application and I look out usually for the killer application, I call the AI-guided applications the killer application. We talked a lot about the ultrasound here, but there's other modality that they can really take advantage of this. To get the AI well trained, you need good data. It does not matter which hospital it comes from and there's a major factor for that operator, how they take the images. The same machine, the same procedure would have different images because of the operators. So if you can get AI guided that's what's really beneficial.

DR. HONG: I'd like to add three more points. First of all, I mean, it would be great if we first have high-performing algorithms. Well, I mean, if you have high-performing algorithms, then we can both optimize for true positives and true negatives, that's the first step. But, you know, reality is often more uglier than the ideal.

So that goes to my second point, which is now FDA now has a total product life-cycle approach for software as medical device, which is really exciting. In that framework we can propose some very specified change control plans and in that -- according to plan based on some agreed-upon plan between the device manufacturer and third-party reviewer, in this case the FDA, depending on the clinical situations or depending on the sites, the user of the technology can titrate between the specificity and the sensitivity, that can be one exciting future direction. The third point is that, it's simple, postmarket surveillance will be very important.

DR. FELTOVICH: I just want to go back to one of my slides, which is garbage in, garbage out, you know, principle number one that we're taught right away in medical school because in terms of figuring out the false positives, the false negatives, the true positives, true -- everything starts with acquiring precise beautiful data, which ties back into we need algorithms to do that and we have tons of help for that, we have, you know, Jean's here with the Perinatal Quality Foundation and that's their focus, we heard a talk about QIBA yesterday and that's their focus. I think we have a lot of room to work together toward that. Dr. Frank.

DR. FRANK: To answer the question as I understand it, as it's written, the impact of image acquisition, AI, on false positives and false negatives is manifest principally in diagnostic quality of the image. I think one of the speakers yesterday referred to the desire to have world-class diagnostic quality irrespective of site, right? So there's this concept that the AI can make an expert of a novice in the acquisition and I think people should understand that that making an expert of a novice is more than just getting the diagnostic quality. It's also optimizing other things, reducing scan time, reducing the administered dose and so on and thereby other benefits, as well, such as reduced need for repeat scanning, which is a savings, as well. This notion of reducing false negatives and false

positives derives from image quality but there are other benefits, as well, to that acquisition AI.

DR. FELTOVICH: And other things need to be taken -- oh, sorry. I'm ad- libbing again. And other things need to be taken into consideration, as was pointed out yesterday and by Dr. Frank right now. For instance, what's in the person's external environment, what else is going on in their internal environment, it's not just imaging although everything starts with that good image.

DR. VAEZY: Thank you. Dr. Samir, you read my mind, I was going to come to you that okay, now come -- let's come to the reality and how would recalls -- how should recalls be handled for AI-guided image acquisition systems?

DR. SAMIR: Thanks for the question. And I'll integrate the answer with the comment I was going to make, they sort of fit together. I think one of the things that we're seeing, in the field of ultrasound in particular, is not only the availability of algorithms, but also the wide deployment of sensing. Handheld or sub-handheld devices are common, where there are multiple papers appearing for patch-based ultrasound devices, as well. Some of these devices don't even do imaging, they just do sensing of physiologic parameters and un-chipped processing of related signals. And the transition there is one in which imaging or sensing was being performed in a doctor's office or at a hospital and is now being deployed out into the community. You see this with the Apple watch which many of you probably are wearing and you're seeing this with ultrasound right now, ultrasound is with a cart, then it's going to be watch sized very soon. In fact, the latest cell phones for fingerprint recognition use ultrasound for fingerprint recognition.

And that's important because it provides a huge opportunity and a huge risk. When you look at the real big-ticket items that have changed patient outcomes, they actually haven't really been medical care, they've been public policy. You know, seatbelts save

more people than all the trauma care in the United States. The reduction in air pollution from air pollution regulations reduced cancer more than all chemotherapy for all cancers combined. And the reason is when you take a small nudge, a small change, and you deploy it widely across the whole community and you push everyone's risk a little to the left, you have an enormous impact on health outcomes.

And what we're seeing with widely deployed sensing is that sensing is moving from being a medical tool to being a policy tool, a public health tool, and then it's going to fundamentally change the way healthcare is delivered to this idea of detecting something, some signal, you're too obese, your body wall thickness is too large or you have X or you have Y and you need to see a doctor is going to fundamentally, paradigmatically change how we need to approach those signals and I think that that question is going to be one that will be resolved not so much by medical imaging companies, it's going to be resolved by public policy and by the formation of new organizations to deal with those signals and positively affect the health and outcomes of the community.

DR. VAEZY: Great, thank you. Dr. Hong.

DR. HONG: So, by recalls, did you mean patient, recalling the patient for image reacquisition?

DR. VAEZY: I'm sorry?

DR. HONG: By recalls, did you mean getting the patient again for image reacquisition?

DR. VAEZY: Getting the patient back for --

DR. HONG: Yes.

DR. VAEZY: -- additional imaging.

DR. HONG: Yes. So, recall in that sense, recall is bad but I think that the bad news can depend on the situations. So in clinical settings where getting images used to be

impossible and now AI-guided, image acquisition technology is enabling more image acquisitions that were previously impossible, I think that in that case, you know, benefits would far outweigh the risks so in that case, you know, some recalls may be okay because, you know, otherwise.

And I also can see some benefits of using AI-guided image acquisition technology helping for establish the settings. So for example, if you take a look at cardiac ultrasound image acquisition echo labs, they typically scan hundreds of echoes in some cases and it is not impossible even for highly trained professionals to forget a few views or a few measurements and these new technologies can quietly look at the images in the background and can send notifications alerting them hey, you need to scan these views and, make sure so that we can minimize the recalls, so I think that's a benefit in two different situations.

DR. MARTIN: I don't want us to lose track of the fact, we're very fortunate in this room and almost everyone listening, to have access to medical imaging. And so two-thirds of the world right now has no access to medical imaging at all, none, and so when you talk about things that can have dramatic impact across the world in maternal/fetal medicine, every 90 seconds a woman dies from complications of childbirth around the world, that can instantly change with access to medical imaging.

And so, part of the promise of this group here in thinking about developing these tools, they're not necessarily the nuance of high end -- but where that can lower the cost of care, accelerate the deployment of these devices around the world where they've been waiting for this stuff to get there. They've got cell phones, but they don't have medical imaging and so where we can take technology there today, we will have a huge impact on world health.

DR. VAEZY: Great, thank you. I need to move to the next topic, very short time left,

from diagnosis to treatment, interventional procedures, and with that, I want to go to our representatives of device developers in this space, Dr. Trahms and Dr. Washburn, please, your thoughts on your AI-guided image acquisition systems in interventional procedures.

MR. TRAHMS: Okay. Thank you. I think this has been discussed in a number of ways over this workshop yesterday and I think it -- again, I'll stress that this is not unique to AI, we incorporate new technologies all the time to allow for in terms of interventional procedures using AI-based techniques or other types of technologies in medical devices I think would be defined in the documentation covered in training and ultimate control still resides with the clinician and the physician, right? And so it's really another tool in the toolbox that allows for making that, you know, effective decision -- decision support.

DR. VAEZY: Okay.

DR. WASHBURN: Yeah, from that perspective of it being another tool in the toolbox, that's certainly true and the hope is that the original tool in the toolbox, being the physician, is still very engaged, right, and that the technology plus the physician provides some additional level of ability that wouldn't be there without the combination of the two. Over time, of course, as some of these assistive technologies get better, maybe they can take on more of the task and become more autonomous over time, but I certainly don't think that's where it begins.

DR. FRANK: And I would just supplement that with the additional comment that yes, the physician is always in control and this really is assisted technology, but there's always provision for a fallback to a manual, fully manual operation, as well. It's never predicated the reliance on the AI.

DR. VAEZY: Great, thank you. Go ahead.

DR. GARRA: I do a lot of interventional radiology and I've watched a lot of it and I've QA'd a lot of it. So, we have guidance devices and we're comfortable with the systems that

tell us information about the procedure, but oftentimes it's how to direct the needle, are you hitting the target, not that there's an organ there that you missed that you're too close to. I've seen several patients die from their eventual procedures where they did an ablation but it was too close to the bowel. Bowels, it's hard to identify sometimes when it's collapsed and a perforated bowel, the person -- people died of sepsis. So, there's a lot of room for additional information in these guidance devices to alert us when we're doing something and there's something that's wrong with it that we shouldn't be doing.

DR. VAEZY: Great, thank you. I'm afraid that I have to go through these slides quickly. Topic 6 about -- is about transparency and knowing, perhaps, what's behind the scenes for AI and I think we have talked about that, we have heard from our patient representatives that some level of information is good. If it's too much, too much clutter, perhaps that's not good.

So, let's move on to the -- the next topic is the ALARA, as low as reasonably achievable. So perhaps, you know, this is of course extremely important in fetal imaging, so Dr. Feltovich, if you could please get us started.

DR. FELTOVICH: The principle of ALARA, which is As Low As Reasonably Acceptable, of course, is critical, especially for developing tissues like the fetus. And I think that, you know, to go back to what I said about that, I think that we can rely on organizations that are already established, for instance, QIBA, PQApp, organizations like this to work with industry and work with physicians to get just enough information, not too little, not too much, but just enough to regulate that.

DR. VAEZY: Great, thank you. Dr. Samir.

DR. SAMIR: I think it's important to have a historic and physics-based perspective on these issues to some extent. When you look at existing changes like the widespread use of harmonic imaging in ultrasound, you know, that increased the positive energy by probably a

factor of around 10 and we didn't even notice, right, we just went okay and kept using the tools. So yes, I think it is important and it's important to monitor energy deposition, but the evidence, at least for ultrasound, that one will be likely to cause harm that is in any way greater than the clinical harm you can cause through over-diagnosis is presently lacking and I think probably not that likely.

DR. VAEZY: Great, thank you. And let's move on to the last topic. This is something that was brought up by Ms. Spitz in her talk, art versus science. Could that last miniscule maneuver of the probe make a big difference in the image information content without affecting the image quality? How could we manage the lack of the art in the science of scanning? Jean, if you could please share your thoughts with us.

MS. SPITZ: Well, as I say, sonographers have their own eye, their own sense of style, their own idea of what's a good image, that may come from where they practice, that may come from where they were trained, that may come from the physicians they work with, but it's hard to change that and sometimes it's disruptive in terms of the purpose of the image and so, we need more examples of good images out there so that they're not in their own little world with respect to what looks good. So many times I hear "but my doctor likes it" and it's horrible. So, we need a bigger perception of quality and I think that AI can give us that.

DR. VAEZY: Thank you. Any other thoughts before we break for lunch? Dr. Hong.

DR. HONG: So precisely because of the reason I think that it is very important to have a system and all the acquired images get reviewed by experts, professional, you know, clinicians. And I think that science is a branch of art. Just wanted to say that.

DR. VAEZY: Thank you. Go ahead.

MR. TRAHMS: I just really appreciated Dr. Feltovich's picture of the babies moving around and the challenges of actually getting to the right views, etc. I think that we've been

working to solve this problem in terms of standardization for a long time and I think protocols have been kind of helping to get the right views at the right time and making sure we don't forget certain things. The image quality side of it is, I think, where AI can help in terms of kind of enhancing protocols with an AI element that can actually provide that quality feedback to the user of whether it's good or not good enough.

DR. VAEZY: Thank you. Any other thoughts?

(No response.)

DR. VAEZY: I guess with that, I just want to thank you for being available, too, as a panel member here and sharing your thoughts. Thank the audience for this session, for being patient and I apologize for running late and after lunch we'll come back and discuss the other topics, specifically regulation. Thank you.

(Applause.)

(Whereupon, at 1:02 p.m. a lunch recess was taken.)

AFTERNOON SESSION

(1:48 p.m.)

DR. LAMB: Everybody, thank you. Thank you for coming back from lunch, those of you who are back already. We're going to go ahead and get started with our public session. First up is David Ritscher. He's a senior consultant at Cambridge Consultants. Welcome.

MR. RITSCHER: Thank you, Jessica. Today I'll be talking about an emerging field in AI called explainable AI and its impact on regulation.

Think back 5 years, how many of you would picture this conference taking place at the FDA? And I'm going to say that a lot of the reason for this conference is captured in this slide. There's an annual competition for image recognition and so what you see, each bar

represents that year's best score and look what's happened to it over the years and of course, in 2015 was the first point where the AI exceeded that of the human practitioner. And that started a new era for us.

Let me just give a word about explainable AI in case you are not familiar with it. It's a very exciting new area. DARPA has invested a $2 billion grant in furthering explainable AI and think of explainable AI, first of all, it is itself an AI and it's wrapped around a host AI. The host AI is driving the car or diagnosing the image and then the explainable AI is probing it and observing it and looking for patterns in it. How is it thinking, how is it deciding, what it's deciding? So that's what the explainable AI is doing.

So much of the value of AI is that it can outperform rule-based algorithms. Of course, it's not explainable, it's a black box. The incremental benefit of the AI is exactly the non-explainable part that surpasses the rules.

What does it mean to be explainable? It means you can describe the rules that give you the outcome and of course, you cannot do that with AI. Explainable AI, XAI, is kind of a misnomer because it does not explain. What it does do is it gives you hints about what's going on underneath, it can generate insights, it's good at probing the behaviors of the AI, and even debugging the process of the AI as you're developing it.

This is an example of an explainable AI. So, there's an image recognition program that's trying to recognize the picture of the frog at the top and the AI comes up with three hypotheses: maybe it's a frog, maybe it's a pool table, and maybe it's a balloon. And the explainable AI looks, well, what's influencing it to make that decision? In a case of the hypothesis that it's a frog, it looks at the frog's head and the eyes. So, the hypothesis that it's a pool table, it's finding round things in the original image because those might be pool balls. In the case of the balloon it finds a big circular area.

Now notice that that gives you some hints about how the AI is thinking, but it

actually doesn't explain anything, so you don't actually know how it made its decision, you don't know why it decided with 54% probability that that was a frog, but it does give you some hints. So that is the beauty of explainable AI.

Now, we've seen this figure a couple times yesterday and so I'll just read the key to the plan going forward. So how does XAI fit into this diagram? First of all, does it replace any of the boxes in the diagram? I'm going to argue no. But where does it supplement the diagram? I'm going to argue in two places, one is in model validation and the other is in model monitoring.

For the model validation, it's able to help us gain insights into how the AI is working and then find potential pitfalls in it so it can help us understand are there risks we're not anticipating and particularly, for model monitoring. So as we iterate in this whole process and we keep on coming up with new models and its learning and we're continuing to model that monitor, explainable AI can give us hints into how it's thinking about things and is it starting to shift the way it's thinking about things in a way that's potentially problematic so it can allow us to anticipate potential problems that are going to be coming. I'm going to argue particularly in model monitoring, , particularly as things become self-learning and are deployed to self-learning algorithms, this is going to be very important.

In conclusion, XAI can provide hints of what's of interest to the AI, it can provide hints about the weak points of the AI. It does not explain the AI, it does not replace the good machine-learning practices that are really key to the FDA regulation of safety and efficacy.

DARPA, as I said, invested $2 billion in this field, we're going to be using it more and more and it's going to continue to develop very rapidly, so stay tuned for what explainable AI will be providing for us in the future. Thank you.

(Applause.)

DR. LAMB: Next up is Dr. Aalpen Patel. He's the chair of Department of Radiology at Geisinger, and medical director for artificial intelligence at Geisinger. Welcome.

DR. PATEL: Thank you for the introduction. When I first started putting this talk together, it was only 5 minutes so I want to be careful what I say because it can keep on going. One of the other things that I was going to initially do is talk about data assets at Geisinger and all the cool things we're doing, but I said I'm not going to talk about that, a lot of other people are going to be talking about that. What I did want to bring to you is the urgency at which the AI is needed in medicine. AI is no longer optional in medicine.

I don't have any financial disclosures.

So under-recognized urgency. All the things I'm about to say everybody knows. When you actually put it all together, there is the urgency of the need is much higher than we could've accounted for. When you look at a workflow for radiologists, this is a busy neuroradiologist practice, so not everybody always reads this fast, but even if you were to read half as fast, a radiologist can potentially look at one image every 1.5 seconds. Let that sink a little bit, one image every 1.5 seconds. Even if it were half the pace, one image every 3 seconds, it is still a staggering pace.

When we do a lot of imaging, we also measure a lot of things. So if you were take an echocardiogram, you would make tens of measurements, but as one of the speakers mentioned yesterday, humans can digest and manipulate in their head five to seven variables at a time and trying to put that into context of hundreds of patients you've seen, health records and records of everybody else like them, it is nearly impossible to actually put it into context of the MR. So that's where machines are going to help.

Can this assist physicians by providing insight from all of these measurements and then putting it into the context of a larger database?

So if you have data overload, you have data waste because we can't possibly digest all the data that's being given to us, and on top of that you put in time pressures the physicians are under and documentation burden they're under, it's a recipe for disaster.

So that's what leads to many of the errors, it's not the only reason, but it leads to many of the errors that we incur. Medical errors, as you've seen in the media and in journals, is a big problem. In 1999 when the IOM report first came out, I questioned it. That can't be right, we're not really killing 98,000 people, are we? And then over the years my own family experienced medical errors-- medical errors where they could've lost limb or life. So that sort of brought it home that yes, medical errors do happen and we have to focus on how to prevent them.

This article came out in 2016 from Hopkins and it was published in *VMJ* and that says medical error is the third largest cause of death in the U.S. The point is, it may be the third largest cause of death in the U.S., but are we actually investing as much money as the first two causes in that -- for the third largest cause? I think the clear answer to it is no, we're not and I think we need to.

The next thing that's coming and we've all talked about this and there are occasional articles in the media that said there are physician shortages coming. We, as a field, overall, we do not have a plan. If you look at China, China has physician shortages, India has physician shortages. In this particular case they are predicting about 400,000 physician shortages. United States, if you look at the Association of American Medical Colleges' data, they publish it every year and they're forecasted, they are predicting a 122,000 physician shortage in the United States. Let that sink a little bit, 122,000.

So, the key is how are we going to take care of our patients without adversely affecting medical care? We already have a problem; we can only exacerbate the problem. This is where AI is going to come in. There was a lot of talk about replacing humans. I'm

not worried about replacing humans, I'm worried about how are the humans taking care of humans going to do that when we don't have enough people to take care of them? I think that's where AI is going to come in, it's going to augment us and make us better. And there's a lot of evidence coming out that says a machine alone is not going to do the job. Humans alone are not going to do the job. The combination of machine-human partnership is where it's going to be and I think that's where we need to focus on.

So again, because of its urgency, we need to make our providers more efficient using AI, make them more accurate, reduce errors, plan for bias elimination, we all know a lot of people -- you've seen in the media that models are biased. It's not the models that are biased, it's actually the data that are biased, generated by humans with their own biases. We're teaching the models using the data that are biased, so we need to fix the data in order to do that, in order to fix the biases in the models.

And the other thing and a challenge to all of you is that we have traditional methods of validation, but at the pace we are going, we're going to get in trouble, so we need safer, faster options for testing, validation, and implementation. I don't have a solution for it. Perhaps one of you do.

Thank you for your attention.

(Applause.)

DR. LAMB: Thank you very much. Next up we have Dr. Ananth Annapragada, Professor and Director of Basic Research in Radiology at Texas Children's Hospital and Baylor College of Medicine. Welcome.

DR. ANNAPRAGADA: I just want to thank the organizers for this opportunity. So now if you all take out your cell phones and you scan that to the bar code, an AI in your

phone will take you to a comment form that will allow you to critique what I have to say.

These are my disclosures. On the left are companies that I started, on the right are companies I wish I had started.

(Laughter.)

DR. ANNAPRAGADA: The case I'm here to make today is children are not adults, they're not like adults, they're not small adults. They need to be treated differently in the context of all of medicine, but particularly in the case of AI in radiology.

So, on the left, what I'm showing you -- these are a couple of papers from our faculty, actually. One is a second opinion interpretation by pediatric radiologists, subspecialty trained pediatric radiologists, on relatively bread-and-butter abdominal ultrasounds and what you can see is roughly 50% of the reads were different between a non-subspecialty trained radiologist and a subspecialty trained radiologist. That's basically a conflict, you could be right or you could be wrong.

On the right is a particularly troubling case differentiating abusive and non-abusive head trauma in children and the numbers are roughly similar there, about half of them you get wrong if you're not a subspecialty trained radiologist.

What does this mean? It means when we're training our AIs for pediatric applications, we really should be using data that is interpreted by a pediatric radiologist.

Why is this so important? And this leads me to the ethics of treating children, they are particularly vulnerable and I'm showing you here some work, this is from my own research group, where we've been looking at AIs to recognize non-accidental trauma, a euphemism for child abuse.

And a typical workflow that we have in our hospital is shown here. It's not different from most places. You could get referred to a social worker, you could get referred to our CAP team, our abuse team, and if they see something suspicious, they'll refer to CPS. This

happens more often than you think. Last year I think there were four million such referrals across the country and something like 700, 750,000 confirmed cases of abuse, 2,000 kids, mostly under the age of three, died because of this.

The margin for error in recognizing abuse is razor thin, there is essentially no margin for error. Why? If you have a false positive, that child is going to get taken away from their family by CPS. If you have a false negative, you're sending that child back to a dangerous environment. This is not a situation where you can do a second test and then prove your accuracy, you just don't have that margin for error, so it's particularly important that we get this right.

To come back to the question of explainability, I think a couple speakers ago they talked about a LIME, LIME and SHAP. We'll talk about those in a second.

In terms of in pediatric radiology, this is the case where you have open growth plates and these can very easily be mistaken for fractures.

Here's a false negative case where this fracture was missed, and the algorithm actually ended up focusing completely outside of the bone.

Here's false positives where it looked at a growth plate and thought it was a fracture. Okay. Again, particularly important in order to gain explainability that we use pediatric radiologist-interpreted images, not general radiologist-interpreted images.

The issues with LIME and SHAP post-hoc explanations, they can be fooled. There's a recent paper from Hima Lakkaraju's group at Harvard and you can very easily train algorithms to report the wrong relevance or the wrong explanation for the AI output. And this was demonstrated in the context of an actual justification.

So explainability, we already talked about that and there's issues with knowledge gaps in user-focused explanations and concept-based explanations.

I'll stop there and allow you to exercise your AIs once more. Thanks.

(Applause.)

DR. LAMB: Thank you very much. Next up we have Dr. Hersh Sagreiya, Assistant Professor of Radiology at the University of Pennsylvania. Welcome.

DR. SAGREIYA: Hi. I'm going to talk about applications of ultrasound to an AI. The first application I'm going to talk about is renal lesion characterization on ultrasound.

Angiomyolipoma is the most common benign solid renal neoplasm and renal cell carcinoma the most common malignant one. And although they're often further characterized by CT and MRI, can we actually use AI to better characterize them on an initial ultrasound?

So what we did is we used measurements from a technology called ultrasound shear wave elastography which measures tissue stiffness and we used those inputs, those values as inputs to machine-learning algorithms and the four machine-learning algorithms we looked at included logistic regression, naive Bayes, quadratic discriminant analysis, and support vector machines and we compared that against a clinical baseline of the median tumor shear wave velocity. And what we found was that the support vector machine algorithm, in yellow, performed best and that with a receiver operating error under the curve of 0.98 and it outperformed all the other machine-learning algorithms and it also outperformed just taking the median tumor shear wave velocity, which performed pretty poorly at 0.62.

Another examination we did was using AI to standardize measurements from different vendors. So essentially, liver fibrosis or stiffness can eventually lead to cirrhosis, but different vendors use different technologies for determining clinically significant fibrosis. We wanted to see if we could use machine learning and with magnetic resonance

elastography as a gold standard to standardize the fibrosis gradings from two different vendors, Siemens and Philips, for elastography.

What we found was that using those ultrasound shear wave velocity measurements as inputs in the same four machine-learning algorithms, that the support vector machine algorithm, again, by far outperformed the median lesion -- the median shear wave velocity amongst the 10 measurements in the liver. And for Phillips, it did the same, the performance for support vector machine was much better than it was for the median shear wave velocity.

And in doing so we can -- essentially, what we're able to do is standardize these fibrosis gradings against a clinical gold standard that's noninvasive, as pathology might be, and doing so in a way that we can evaluate new algorithms in such a way and standardize them against a clinical gold standard better than using the traditional method.

Another application I want to talk about is early cancer treatment response. The idea here is that is it possible to determine whether or not a cancer is responding to chemotherapy and in essence, are we able to see if there's a change in the blood flow to the tumor perfusion before there's a change in the size of that tumor? The idea here was, we had a preclinical dataset in this case and it was the response of mice to the anti-angiogenesis agent bevacizumab. So what we used was something called curvature learning, which is an unsupervised learning algorithm, it's a little bit different than the convolutional neural networks that we've heard about a lot, but essentially what it does is that it takes this contrast-enhanced ultrasound data, which is 2-D plus time, and it compresses them essentially into this 2-D heat map that represents how suspicious different tissues are. We take those pixel classes and we essentially look at how those pixel classes differ within a given imaging session with time and also how they differ pretreatment and post-treatment and we use that information to calculate essentially a

tumor score and we compare that tumor score between the control groups that were either untreated with a chemotherapy agent or were treatment resistant and also with the treated treatment sensitive group and we found that there was a statistically significant difference.

So essentially, we were able to show that there's a change in tumor size without a change in the tumor's blood flow perfusion without a change in tumor size and this was despite the fact that the tumors were still the same size and without having to perform any segmentation or manual steps.

A final application I'm going to talk about is anomaly detection with ultrasound images. Essentially the idea is you have hundreds and hundreds of images that you get as part of an ultrasound examination, but only a few of them are really important or significant. They might have a cancer or a lesion that you're looking for.

So essentially this technique detects, you know, frames that are different from the rest of the data, it employs something called a generative adversarial network that consists of a generator that creates synthetic images and a discriminator that discriminates real versus synthetic and once you get a new sample, what it does is that it if there's a new sample that's not in what's known as the latent space of the generator, it's deemed anomalous. And we hope to apply this to thyroid ultrasound and a number of other techniques. Thank you very much.

(Applause.)

DR. LAMB: Thank you very much. Next up we have Dr. Enhao Gong from Subtle Medical.

DR. GONG: Hi, I'm Enhao Gong, Founder and CEO of Subtle Medical. I graduated from Stanford in electrical engineering with a focus of applying AI in medical imaging, and I

cofounded Subtle Medical with a practicing neuroradiologist and right now, a 2-year journey, we have received two FDA clearances of our AI technology and mostly focused on using IT to improve medical imaging quality for better image acquisition and enhancement, and today I would like highlight the technology and the research behind it and to see how AI can potentially improve the acquisition and enhancement and with three highlighting applications. The first one is using AI which can improve the quality and efficiency for PET-CT exam and second one is we can improve for MR image and a third one is we can use AI to potentially reduce the contrast dose for MRI patients.

And most of our technology is based on deep-learning CNN and with a lot of research we find, actually, a lot of times it's not only the specific AI architecture -- model that matters most. But more importantly with AI technology, we need to be trained and validated by representative unique paired clinical datasets.

So in a lot of patients the input is trained on the image that we want to enhance, either it's because we need to improve the quality plus the low dose or it's acquired with a further exam or we just want to see better resolution. And also, usually it's real clinic cases but using a longer exam and it has superior image quality or even exceeds standard of care quality. This way we have this unique pair of datasets that does not really exist in clinical standard of care and that's how we train and validate it. And the beauty of deep learning is it will keep evolving instead of like become fixed within that, like, conventional algorithm and daily add to the research we have done and also, at Subtle Medical, we are accumulating and curating more and more datasets to make sure we get the better performance. And more important is we need to make sure it's always generalizable for all the patients, with all the manufacturers' device, and with all the PACS systems and more importantly with other downstream locations either on the diagnosis side and also on the downstream quantification as well. So that's a lot of validation we are doing.

And for PET, we have published clinical research shows we can use AI to improve PET image quality. It's a lot of use cases, where typically, for example, on pediatric cases, you want to use low dose and for all other sizes they may use record protocols and then find the image quality is not as good, or just because some cases, the image quality is intrinsically lower. And here we show using AI improves the image quality and compared with conventional algorithm, it's actually better in terms of both the visual quality, it looks real because deep learning is trained with data-driven masses, not with certain fancy recommendations. And also, from a quantitative perspective, it's way more accurate.

And we show several research studies on how we can use AI to apply in MR exam. MR is a very valuable exam, but usually it takes a long time and we can use this to enable just like a faster protocol, to improve the image quality. And this is one site that used the record protocol already and the image quality is good enough maybe for certain diagnoses, but you want to get the most out of the scanners you have. So use AI, you can train with paired datasets on the input image and what the image looks like with longer exam. That's how we train and validate it as ground truth and we show we can improve the resolution and the quality using AI.

And this is another patient's -- MSK (ph.) of patients shows you can improve image quality without sacrifice the details. And there are several researchers have been evaluating using AI and CNN which can improve the image quality, SNR resolutions compared with longer exam shows equivalencies.

And last, but not least, we want to highlight research using AI to reduce contrast dose. We received NIH grant on this, we have done research using AI to reduce dose. Around a third of MR exams use contrast dose, gadolinium contrast dose, to highlight a tumor. But they recently find there are risks around it and both NIH and FDA have paid a lot of attention on patient safety. So here we show AI actually can tell the subtle difference

and enhance it to gather same quality.

In a way we show AI can basically help us to renegotiate the tradeoff clinicians always have to make between better quality and the cost and potential risk. It might be the time cost or radiation dose or contrast dose that you don't really want to inject to the patient too much. In this way AI can provide better imaging workflow for physicians and provide better care of patients as well. Thank you.

(Applause.)

DR. LAMB: Thank you very much. So Dr. Greg Zaharchuk from Stanford could not join us due to his clinical responsibilities. His colleague, David Larson, is going to give his talk. Thank you very much.

DR. LARSON: Thank you, Jessica. As she mentioned, I am not Greg Zaharchuk, I am his colleague at Stanford and a big fan of their work. This also is not my presentation, but I'm going to do my best to share with you the great work that they're doing.
Dr. Zaharchuk is cofounder, along with Enhao, of Subtle Medical, so a lot of this work is a continuation of what Enhao already has discussed.

If we step back, we think about the whole radiology value chain, a lot of what we talked about yesterday was really on the diagnosis side, the analytic side, and today it's great to see that we're talking a lot more kind of further upstream on the acquisition side. And if we think of the whole picture, 80% of the cost actually at the current time is on the acquisition and processing side. No matter how much we do on the diagnostic side, that's actually a minority of the overall cost. If we can improve that on the upstream side, then we can improve the overall cost impact. That also, of course, is where our patients are at the interface and so we can improve those patient experiences.

We're going to talk in just second about how we can improve the contrast based on information we already have with other imaging, we can put those together, and how the force of the data from the upstream can help to improve the classification downstream.

So, the first case is MRI in terms of decreasing time. This is material that's a labeled image that takes 8 minutes to acquire and if we use a sequence that takes less time, you have a lower signal-to-noise, a noisier image, we can actually use other sequences that we're going to acquire anyway on that MR and dramatically decreases that time to acquire an essentially equivalent image to the high signal-to-noise ratio, arterial thin labeling image with a significant decrease in overall time.

Another example is reducing dose, so not just using the image itself but again, if we take the information that we have from the accompanying MR image and we apply that to a low-dose PET scan, we can decrease the dose by a factor of a hundred to 1% of the dose and get a roughly equivalent image with applying AI algorithms.

And we can actually get what's been termed a zero dose FDG PET image where we take the information on the MR, including the contrast, and then we train that with an algorithm that's compared to the FDG PET and then we can predict what the PET scan looks like without even getting the PET scan. So here on the top we see two cold lesions in the frontal lobe. In the bottom we see lesions that are predicted to uptake activity on the PET scan.

This is a really exciting one from my perspective. Here in this case we can actually train models to evaluate what would happen with a given patient if you used different treatment strategies. So, this is a stroke patient, this is a perfusion study and we can apply a contingency or a scenario to each patient, on a patient-by-patient basis. We see on the top, for example, if we give tPA in this patient, we can see it would significantly reduce the signal abnormality on the predicted follow-up exam compared to if we don't give

treatment. Whereas we see on the bottom patient that if we give treatment or don't give treatment, the algorithm predicts that it will not have a significant impact on the image.

And we cannot only predict the response at the individual patient level and then select those who this is the most likely to respond to, but we can do it in a graphical way rather than just a narrative. And there is preliminary information that indicates that this can be applied not just to stroke but a broad number of applications.

In summary, there's a lot of "hidden" information in the current imaging that we can use AI techniques to harvest in ways that we really wouldn't be able to do as humans, and it allows us to change our assumptions regarding cost, speed, dose, safety, contrast. We can really use information that's already there and really change that paradigm.

Again, models can be trained then to predict future imaging based on personalized treatment that's not just the worries about what might happen, but we can see that visually, what is likely to happen to the patient depending on different imaging strategies. Thank you.

(Applause.)

DR. LAMB: Thank you very much. Next up we have Dr. Shandong Wu, Director of Intelligent Computing for Clinical Imaging Lab at the University of Pittsburgh. Welcome.

DR. WU: First, I would like to thank FDA for this opportunity. My name is Shandong Wu from the University of Pittsburgh, where I lead the Intelligent Computing for Clinical Imaging Lab and the newly launched Pittsburgh Center for AI Intervention in Medical Imaging.

So, I think by today we have a lot of evidence support that we have proved concept of imaging AI in terms of building deep-learning models for a range of different applications,

as we can see from a lot of published papers and the quick advancement of imaging AI in industry.

From the transition to what we believe imaging AI to be, we trust imaging AI, I think it poses a great opportunity and of course a lot of challenges to build trustworthy AI. This is not a comprehensive list of definitions of trustworthy AI, but I think some of the important factors that I listed here, we want the model to be, you know, generalizable, but I think we shouldn't really expect the model generalizability without any changes to the model simply because of the complexity and heterogeneity of the data and the populations.

So if we can customize a model and then we can trust this AI model because we know it is localized, it is customized to a local cohort and dataset or institution or a population, and we want the model to be robust and, under different quality of the data and also reliable and safe under intended attack. We want the model to be explainable, but we want to think about where we really wanted the model to be explainable.

For example for self-driving cars, we really don't need the explainability here. You know, I think we don't need it, we're comfortable without that because we know the physics behind the decision making of a self-driving car. It has sensors and it computes the distance and, you know, speed and direction. So, we know that decision making, we're comfortable with that. In terms of medical imaging it's different, but for segmentation, I don't think we really need explanation; we can see the segmentation, we know if this works or not.

It's actually often the scenario that for disease diagnosis and outcome prediction using deep learning, that's where we really expect some interpretability because often these are the scenarios that actually, even for human experts, is hard to explain some of the processes because it often involves a qualitative or a complicated reasoning process in this person, even for human.

Knowledge is very important. If we can incorporate domain knowledge in the curated data-driven modeling, I think this will help us build the trust because then we know we feel comfortable and we know there's something meaningful incorporating the model. On a biological level, we can also provide some support to build trust and I will give you some examples from some of the work in my lab.

We built a deep-learning model to distinguish, to identify potential false recalls in breast cancer screening and as you can see, this deep-learning model generalized from one dataset to another dataset. But we see a large variation of the model performance, right? So, if you want to really trust this model in a different institution, we need to customize that with the model being adopted into a local site or a local cohort.

We also built a deep-learning model that can deal with inaccurate labels. So we, through automatic identification and the correction with these missed labels, wrong labels in the dataset that we canmake the dataset clean and build a better model and improve the model performance. This way we also feel we can trust this model because we know it's built on clean data and it can deal with noisy data or errors in data.

We also generated adversary samples, you know, these fake images from real images and using the GAN network. So, we're using this. It turns out that this fake image actually leads you to a 67 percentage of cases of missed diagnosis that we inputted into the deep-learning model. So if you look at these fake images, for some of the images, we should look at an experienced radiologist guy to identify that, especially after an education intervention, but it still doesn't skew the number of images that we cannot identify that and the model is fooled, made a wrong diagnosis. So, this is very important, we have to make sure we can prevent this from happening because otherwise we cannot trust this model, right, we cannot trust the model's output.

We also built a deep-learning model to predict short-term breast cancer risk from

normal histo-mammography. We have several cases of work, I don't want to go through details, but as you can see here, those heat maps, we identified automatically from deep-learning modeling that it helped us actually gain, you know, at least help us understand where the models are looking at, right, so for risk prediction because I think this also helps us actually gain some insights and some clues and some sense how the model works because in reality I think, for radiologists, it's where -- we don't really know where to look at it for short-term risk predictions.

We also did a study to look at the correlation between imaging features and the protein level expression of both breast cancer patient and brain tumor patient. As you can see from this kind of correlation between imaging and biological-level data, that we can potentially identify some biological pathways so that it helps us interpret imaging features we identified from radiomics or from deep learning. And I think this is another way to gain trust from physicians because we know, by training a physician they tend to think clinically and biologically, right? So, they tend to find evidence from a biological perspective, not a computational perspective.

So, you know, as a summary, these are several, I think, important factors I expect that we want to really address in order to build trustworthy AI, and this is also one of the important topics in our Pittsburgh Center for AI Intervention in Medical Imaging, where we really want to leverage the computational expertise and clinical resources and expertise in University of Pittsburgh, in University of Pittsburgh Medical Center and -- University, so that we can build innovative and trustworthy imaging AI.

I want to thank my group and the funding sources. Thank you for your attention.

(Applause.)

DR. LAMB: Next up, we have Mr. Armen Vidian from DCVC. Thank you.

MR. VIDIAN: Thank you, Jessica, and thank you to the FDA for inviting me to speak here today, very flattered. Armen Vidian, I'm a partner with DCVC or Data Collective, which is a venture capital group based in Palo Alto and San Francisco, California.

But first, here's some legal stuff. I'll wait while you read all of it, every word. Just kidding. Don't tell our attorney, he worked really hard on that.

Anyway, so some background on DCVC and what we are and why we're here today, perhaps. So DCVC is the leading firm investing in artificial intelligence and data science startups for the past 9 years running now. Our thesis is really where artificial intelligence meets propriety data, either public or private data sources or in some cases exclusive arrangements where that intersects with deep computing and in some cases robotics or capital equipment. No, we are not afraid of capital equipment-related startups, where that can be fundamentally disruptive to CapEx and OpEx for major industries, and solve problems in industries such as healthcare.

In fact, about 35, 40% of our investments from the past two, three funds have been in the healthcare space. I focus exclusively in healthcare and the life sciences for our firm. We invest early. We have five funds spanning about $1.75 billion under management. Average amount of first money in is maybe 2 to $5 million. That's increased with that with our latest fund, which is a $725 million vehicle, but we still do like to go early. And we are very proud investors in Caption Health and Subtle Medical, major investors in both, and we are very proud and grateful for their working relationship with the FDA and of the entrepreneurs at those companies.

So why is this topic important? I was having an interview with someone out here of some importance recently who said you know, a lot of investors say they should -- entrepreneurs should avoid working with the FDA or figure out how it is that entrepreneurs in healthcare can circumvent it or -- and we find it odd that you would say that they should

have a close working relationship with the FDA. I said that's very strange because we encourage the opposite and I said maybe this is an opportunity here today to either set the record straight or say what the majority of us really think.

There's a lot of great technical entrepreneurs in this space. We are grateful for them, they have a lot to add, but this can often be kind of intimidating. So, we encourage entrepreneurs and other investors to work directly with the FDA very early and often. This is about your credibility, this is about knowing early on what your message is to the users of your product and potentially to patients and understanding why that is going to be important to them. It's part of not just an obligation, but a strategy that you articulate to the market and so it confers advantage to you, not just something with which you have to comply.

So, understand the regulations early, understand what it means for your business, understand likely what kind of staffing needs and financial needs you're going to have as you move along with developing your business.

And this is important. Perhaps, if there's one thing I could leave you with today, it's that strong product management and regulatory collaboration early on is key because the regulations may affect the type of product design you need to have and vice versa. What you don't want to have happen is explain to your investors later why it is your product needs to be iterated and why you need more money or the other way around, why you're trying to force the regulators to believe that what you have is okay and that's not a good look. So, you have to have that collaboration early on.

Be transparent. Instill that value early on and understand, have your investors understand what is going to be required in order to get an FDA authorization, and make sure that you understand why that's a critical part of the strategy, right? What claim are you trying to make and why?

So finally, I'll just leave you with working to integrate the regulatory strategy into your fundraising strategy and your message to customers is important so that everyone can be in alignment as to why you are pursuing certain claims, why that's important to customers and ultimately, how that's going to make you a really big business. Incorporated is part of your strategy.

That's all I have to say. Thank you for having me here today.

(Applause.)

DR. LAMB: Thank you. Next up is Mr. Brian Scarpelli from Connected Health Initiative. Welcome.

MR. SCARPELLI: Okay, all right. Thank you. Hi, everyone, a pleasure to be here. My name is, as Jessica just said, Brian Scarpelli. I'm with a not-for-profit multi-stakeholder policy advocacy organization called the Connected Health Initiative, and I'm pleased to just share a few thoughts with you here. I'm very honored to be here and thanks to the FDA and all of you for being here.

I won't read all of this, but what may be more helpful would be to, you know, kind of describe, you know, like the widget we're making essentially is we're trying to advance policy changes, largely at the U.S. federal level, to allow for the responsible use and deployment of digital health innovations widely and we work on a wide range of topics across a whole bunch of different venues. What maybe isn't listed here is that we have kind of like an advisory board, a steering committee, and we've gone to all of these great lengths to incorporate a bunch of different viewpoints. In other words, not just another industry group, but we're happy tohave participation and membership of the steering committee from, for example, the physician community as well as industry, large and small, some

academic medical centers, providers, some just what people might still call tech companies, whether or not they like to be called that, and the insurer/payer community, etc. And the whole idea is anything that we're doing reflects a thread of consensus across all those different communities.

So, you know, this is just an image about our association there. The reason that I am happy to kind of share what we're up to in this space, and really I am seeking in-kind collaboration at, you know, super granular, high level, anything is gold to us, but you know, in an effort, you know, to be proactive rather than be reactive and wait for, a request for information from -- from a government, , a regulator or a government agency or Congress or something like that to come out and react to that, we wanted to be proactive.

So about 1 year ago, a year ago and a month, we formed up a task force that again is open to any interested stakeholder. It's not like money ask or something like, we just want to collaborate. In an effort to be proactive, we came up with a set of policy principles for AI in healthcare. The idea is for regulator X or policymaker X to consider these different angles and facets as they think about a regulator's role and how AI fits into healthcare to advance the Quadruple Aim. I know that these slides are going to be made available or something, I hope. Great, okay, because there's a link to those policy principles and we're dying for anybody's input on that.

And this lists the major areas, but it was an attempt, at a high level, to be comprehensive and it ranges from some core issues, quality assurance oversight, major themes here, thoughtful design, payment, affordability, and as well as, you know, workforce issues, even educate -- by education, we mean like curriculum and medical schools, for example, and things like that. So that's just kind of, you know, something that we've done looking back.

Moving forward, we've been socializing these policy principles with a number of

regulators, including the FDA, who gave us some really cool feedback, but also Congress, the European Commission. We recently presented on these to a United Nations World Health Organization conference in Atlanta that was pretty great. So really we're trying to pull on threads here and find interesting things.

Kind of based on conversations, really inspired by the FDA's great work in being proactive through putting on conferences like this, our thought was that, you know, what we'd like to do is to kind of survey the landscape and look at key use cases like this one, like what brings us all here, and try and pull out some good machine-learning practices, as people like to call them, that build on all of this good work that's already occurred, and then to socialize that with you, the community, and hopefully that will move the ball forward. There's a whole bunch of things on the policy level that probably need to happen before it would be implemented, you know, before it would go into effect, like through an FDA regulation and stuff and we're working on that, too.

But these bullets here try to capture what our good machine-learning practices consensus document is going to reflect and we're soon going to be publicly releasing that and again, I would love to work with any and all people to take in their input on that.

I should recognize too, I think some very good points were made yesterday and probably today about locked versus unlocked algorithms, I guess, and realizing that's something we've beenworking on with this document. The attempt is to squarely address locked, like today, but also hopefully to provide something that can kind of be picked up and dropped in the unlocked context, too, and all of the -- you know, we plan to start these consultations with all of you in the community before the end of the first quarter. So, I'm very excited to share that and I look forward to working with you all.

That's about all I really had to say. Thank you very much.

(Applause.)

DR. LAMB: Thank you very much. And coming up, Dr. Houston Baker from the National Cancer Institute.

DR. BAKER: How's my signal, am I coming through right? Back of the room? Okay, all right. Houston Baker and George Redmond presenting. I'm responsible for -- mainly for Slide 1 and 2 and George is responsible for Slide 3. He was detained so he couldn't be here today, but I think he's online, wherever the camera is. Hi, George.

Anyway, we wanted to make note of the fact that, with the big changes coming in artificial intelligence, we saw a lot of presentations today and one of the common themes that kept coming through was "needs research," okay? We're into research funding. We wanted to present some material to you on early, mid-stage, advanced stage, and translational stage imaging development funding opportunities which includes artificial intelligence.

So NCI, the first slide is to give you some general information because I notice a number of people in the audience are not familiar with the grant application process at NIH. And NCI publishes funding opportunity announcements and these are notices of availability of funds.

All right, in our case for cancer research, in 2020 we expect 1,300 new applications and continue to pay 4,000 continuing grants. These grants support basic and applied research and all things cancer and the most common variety of funding opportunity announcement is program announcement, which we abbreviate as PA and they focus on an NCI interest. We also have PAR, which is program announcements with special "R" activities, namely receipt dates or referral to limited sponsors or special review criteria.

We also have another fairly common funding opportunity and that's the request for

applications, also known as RFA. These tend to have a narrower focus than a PA or a PAR and most RFAs use set-aside funds, smaller amounts of money for a smaller number of grants, and they use cooperative agreements where an NCI scientist is appointed to participate.

And so that gives you a little bit of background on the more common funding opportunity announcements.

What I want to bring to your attention is the bioengineering research grant announcement series and these accept artificial intelligence imaging developments. This series covers the life cycle of technology development from the feasibility stage to early preliminary data, prototyping, more advanced stages and all the way to translation to end users, which usually means a product.

The Exploratory Bioengineering Research Grant is bimodal because it has exploratory and developmental features, exploratory to test feasibility, developmental to get early stage preliminary data. Most study section reviewers don't recognize that dichotomy and so these don't get the best review we could hope for, but they work.

The Bioengineering Research Grants are special review characteristic applications and they have two aisles, clinical research not allowed or clinical research required.

The Bioengineering Research Partnerships is the late-stage advanced development activity. It's not so popular at NCI as it was originally, because we've come up with this Academic-Industrial Partnerships for Translation of Technologies for Diagnosis and Treatment and that one is an advanced stage with our early experience. For first ones funded that are finished, we have over 70% showing products on the market at the end of their final report, usually five or five and a half years after they finish.

George's contribution is biomedical informatics and information technology RFAs for informatics technologies. These have set-aside funds and so the number of applications is

limited, but they have an R21 grant series for exploratory and/or early stage developmental grants. They have next-stage prototyping and hardening type of activities, which is a U01, it's related to the R01, which is the basic NIH grant. The U01 is a cooperative agreement. Advanced development is a U24 resource-related cooperative agreement. We also have two announcements there, one for sustained -- sustaining algorithm development, the other one for revisions of existing activities that want to do an upgrade or explore a new idea. Thank you.

(Applause.)

DR. LAMB: Thank you very much. Thank you once again to all of our public presenters, we really appreciate your contributions to this workshop. And my colleague, Dr. Garrett Astary, is going to lead into the next session. Thanks.

DR. ASTARY: Thank you, Jessica. So my name is Garrett Astary, I'm a lead reviewer in the Division of Radiological Health at FDA. I'll start the next session, which is regulation of imaging devices containing AI software.

It's my pleasure to announce our first speaker, Dr. Shahram Vaezy. He's a biomedical engineer in the Division of Radiological Health at FDA. Shahram.

DR. VAEZY: Thank you very much for the introduction. Okay. This is the last session and perhaps maybe the last stage or the last step in a technology, an idea coming to the bedside for patient care.

So Regulatory Considerations of AI-guided Image Acquisition and Optimization, that's the title of my talk. For the outline, I'm going to talk about, of course, AI-guided image acquisition and optimization. We'll describe the current regulatory status of this

technology. Clearly, any regulatory consideration is done in the context of the benefit-risk analysis. I'll describe a little bit about general and special controls that go into the current regulation that we have for this technology and cover a bit the area of device modifications, future device modifications that a product, a device of this type can have. I'll briefly touch on future directions for this technology, and key elements of regulatory success.

At this point, probably, it's not really needed to talk about an example or an application of AI-guided image acquisition, but this goes back to my area of research before, bleeding detection that could be done using ultrasound imaging. Clearly, there is a specific area or there are specific areas that one could look for bleeding detection. In this particular example, I'm showing an area between the liver and the kidney, Morrison's Pouch, where blood would accumulate in the case of a blunt abdominal trauma. You can see that it would be extremely important for this type of examination to be done at the scene, in a remote setting.

But clearly, there needs to be some guidance to novice users, perhaps the buddy soldier who is going to do this and/or maybe the medics who do not necessarily have that training. But this is a very clear example of how AI-guided image acquisition could be helpful because if there is a suspicion of blunt abdominal trauma and bleeding internally, the staff at the hospital could be ready for that patient to come in.

So as far as this technology is concerned, AI-guided acquisition and optimization, FDA has authorized the first type of this technology and this is for the Caption Guidance system. The authorization was done on February 7th of this year. This was via granting a de novo request for classification as radiological acquisition and/or optimization guidance system. This establishes a Class II and for this regulation to be generated, there needs to be special controls, there needed to be special controls and they were developed and we're going to talk about that. Based on this Class II regulation classification, clearly a 510(k)

regulatory pathway is established for following devices of this type.

The regulation itself, radiological acquisition and/or optimization guidance system is a device that is intended to aid in the acquisition and/or optimization of images and/or diagnostic signals. So, you can see that it's not just the images, but also signals. The device interfaces with the acquisition system, analyzes the output, and provides guidance and/or feedback to the operator for improving image and/or signal quality.

As I mentioned, a Class II regulatory pathway is established. These devices would be non-exempt. Basically, they do need a 510(k) submission to the Agency. Devices of this type would be considered prescription use devices and not for the over-the-counter use, and images and signals have to be interpreted or read by a qualified healthcare professional.

As I mentioned, this establishment of a regulation is done in the context of benefit-risk assessment. For this benefit-risk assessment, a summary of benefits has to be outlined for this particular device type. Clearly, as we have discussed today and yesterday, there's improved access to clinical care and in terms of location of delivery of care and the time, which is an absolute, I guess, paramount importance in this. Improved performance of these devices for experts who are using this technology to standardize the image quality can clearly be seen. And as we heard this morning, there is definitely a benefit of patient convenience, again, in terms of location and time for getting a scan done, basically.

Clearly, there are risks too. A summary of risks would be low image quality, low signal quality, and delayed clinical care if there are low-quality images, images that will not have diagnostic quality.

Other factors also go into benefit-risk analysis. If there is a gap in the clinical care, we certainly consider that gap, how that can be filled with the establishment with this regulation, with devices that actually fall under this regulation and can be authorized for

clinical care.

Again, by now it must have -- it must be very obvious that the patient perspective is extremely important in this context. We need to know how the device impacts the balance of risk and benefit for clinical care, and what level of uncertainty could be expected, anticipated, and perhaps accepted for a device that clearly has certain benefits and risks.

For this particular regulation, this was basically hinted at this morning, that the risks of the device have to be considered with the associated mitigation methods. For these types of devices, two types of risks were identified. One is a device error where there's perhaps a problem with the algorithm, it leads to failure to provide guidance or acquiring diagnostic-quality images or signals. That would clearly lead to delay or prolonged examination or additional unnecessary procedures, as I mentioned, perhaps due to algorithm failure or hardware failure.

There's also a user error, risk of a user error, where the operator would not necessarily follow the instructions provided by the system and then the low-quality images would be acquired, images that do not provide sufficient information for the diagnosis to be performed. The mitigation measures would be established in the context of the special controls that I mentioned. And we're going to talk about that.

So as sort of an introduction, this was presented yesterday. General and special controls are applied to all medical devices unless they're exempted by regulation. Examples of general controls would be electrical safety. If a device is an electrical device, clearly, its electrical safety has to be demonstrated. If it's coming in contact with the patient, the biocompatibility of patient contact material has to be established.

The special controls are those that apply to the specific type of device. For this particular device, of course, this is AI-guided image acquisition. This is, of course, done in the cases where general controls are not sufficient to provide a reasonable assurance of

safety and effectiveness, and there is sufficient information available to establish special controls for risk mitigation. Examples for this type are performance testing relevant to devices in the regulation and labeling. They're specific, again, to this type of device.

Let's take a look at the special controls that were developed or established for this type of device. One is design verification and validation. This has to be -- the components of this design, the device design and verification and validation include a detailed device description, a report on nonclinical performance testing, and a report clearly on the clinical performance testing of the device.

You can imagine that for the clinical performance testing, the protocol has to be described, you have to have quantitative evaluation of the diagnostic quality or diagnostic utility of the images that are acquired. And I'm not going to go through all of them, but clearly the patient population has to be considered, the user population has to be considered in that performance testing. These devices are going to enable users that are not necessarily experts in acquiring images, so clearly an effective training program has to be included. And in case there are future changes, a detailed protocol for what these changes are and are anticipated and how they could be tested, their performance could be tested, should be included.

The second component of the special controls for this type of device is labeling. Again, a description of the device, the patient population, the intended user population, very clear and detailed instructions for use. For this particular device we thought it would be very important to have basically the physician or a qualified professional interpret the images for these type of -- so these devices that would come with a 510(k) to the FDA basically, they would be prescription use and the images would be interpreted by a qualified medical professional.

The reports, performance reports, have to be included in the labeling for the users

to know and be familiar with the performance they're going to get with the device. And since these are ultrasound -- in fact, I should say not just ultrasound, there are other imaging modalities, ALARA has to be also upheld and the labeling should specifically discuss that.

As I said, there could be possible device modifications in the future. This is a new paradigm at the FDA done -- what has been basically been provided to the public is a white, a proposed regulatory framework for modifications to artificial intelligence/machine learning-based software as a medical device. Other folks have presented on this specific paper and its content. For this regulation we included special controls, as I mentioned before, that does consider future modifications to the device and as I mentioned, it has to look at the changes and how they could significantly affect the safety and effectiveness of the device and the risks proposed or posed by these changes.

So future considerations. At this point you look at this and, you know, with ultrasound being such an operator-dependent imaging modality, one might think that, for the most part, this regulation would apply to ultrasound imaging systems. But we could envision that, and as we have heard, you could envision that this technology -- this regulation could also cover other imaging modalities as well.

As far as we are concerned and what would be acceptable under this regulation, it could include other indications for use and clinical applications. It could include other users, perhaps patients in a home use environment, and it could include other technological features as well, like measurements and analyses. But performance testing clearly has to be very specific to the device type that's coming in a premarket notification, a 510(k).

Just perhaps the key message that I'm sure you have heard from other folks as well, is that early interactions with FDA are highly encouraged. We have seen a very definite increase in the quality of the submissions and that, I would say, is in large part attributed to

the early interactions with the FDA. And there's a specific mechanism called Q-submission program that developers can utilize to interact with the FDA. In this context, I would highly recommend that you would have well-defined device characteristics, well-defined intended use, and well-defined performance testing plan and to come to FDA and share with us your plans for device development and we'll be happy to discuss with you and engage with you and help you in your regulatory pathway. Thank you.

(Applause.)

DR. ASTARY: Thank you, Shahram. I would like to invite up Scott Paulson. He's the senior director of regulatory affairs and quality systems at Fujifilm SonoSite.

MR. PAULSON: All right. It's a pleasure to be here today. I thank the FDA for the opportunity to present. As I sat here and watched all the presentations this morning, there were a lot of messages and themes that have already been spoken about and presented on even yesterday, that you will see commonalities with what's in my slides. But my message just takes a little different angle to it, so it ties a few things together.

My presentation is partly on, you know, how an AI-guided ultrasound system or product can help providers in clinical settings achieve the Quadruple Aim, and then also how the current ultrasound premarket guidance aids in getting those products to market.

Let's first look at the key points and questions that I'm going to look at today. What will be the impact of the incorporation of AI on existing models of ultrasound and new technology? Who will benefit from the incorporation of AI into ultrasound? Is there a regulatory framework in place to support innovators? I think this has been covered, I'll take my own spin on this, and a few conclusions and takeaways.

Generally speaking, ultrasound is good for public health for a number of reasons.

The imaging modality has been around for decades and we've seen the evolution of the technology, you know, in the last 20 years. As mentioned earlier today, non-ionizing radiation, this is a benefit, especially when you're talking about a wider population of patients, even young pediatrics. Real-time image capture and display. Being able to capture an image in real time with the physician and patient there together, it really enhances that interaction between those two parties. And then safety. Besides diagnostics associated with gathering an image or acquiring an image, there's also guidance that can be provided by incorporating ultrasound in certain procedures. Maybe those procedures were done blindly or using a landmark method. And then just general expanded use of the ultrasound technology itself, hopefully AI will lend itself to that.

Okay, ultrasound has come a long way. Then and now, really you had cart systems that were really, really large and it was difficult to move them around. You usually had to bring the patient to the ultrasound. Ultrasound now, it's so small and portable, you have -- you have large, highly complex, powerful ultrasound systems and you have point-of-care systems you can hold in your hand. Obviously, they serve different purposes, but really today ultrasound systems grow legs, they go to the patient.

The ability to use ultrasound to make decisions is a powerful tool today that can be utilized hopefully with the incorporation of AI by a wider population of providers, increasing the impact that the technology could have on the general population.

All right, so AI and ultrasound, you know, what is it? Does it augment human intelligence? Does it replace it? Personally, I think due to the provider-patient relationship, AI really, really, really benefits that by enhancing that human intelligence element and then benefiting both the patient and the caregiver.

Let's talk a little bit about the Quadruple Aim. I'm sure it's been brought up a few times and for those who aren't familiar with it, so you have four elements, it used to be the

Triple Aim, but you have quality of care, population health, cost reduction, and then finally, the well-being of the care team. So when people talk about burnout and additional work that comes with caring for patients, this also factors now into this focus. So how are hospitals measured? It's more on quality now than it is quantity, which is why these four factors are important.

All right, how do we get there? We'll take a closer look, briefly, at the current regulatory framework for diagnostic ultrasound systems and how AI fits into that framework. Also take a quick look at medical device development methodology in general and touch on some design control aspects which fall under quality system regulations with the FDA and use of standards such ISO, particularly 13485.

So where does AI fit into all this? Let's take a look at that. And, Dr. Vaezy, it was a nice lead-in with your presentation and talking about general controls and special controls, so I'll stay clear of getting into that, but as you mentioned the de novo pathway and then the non-de novo 510(k) pathway, you know, those are available for moderate class devices, so Class II.

I think it's important to note that -- excuse me. What is not depicted here is the Q-sub, as Dr. Vaezy mentioned. The Q-sub is a mechanism that is in place to make sure that industry engages ahead of time and is proactive with regulators and having that in place enables a more collaborative approach to, I guess, a regulation, a process to getting product to market while maintaining efficacy and safety.

What it also does is it is effective because it helps manufacturers or innovators as they're designing AI and planning the incorporation into whatever product it may be incorporated in. It helps define the direction for the company to go to try to save time on the Agency's side and the innovator. Plus the additional interaction and communication between regulators and innovators. As FDA has kind of stressed over the last couple of

years, the more collaborative approach is the more beneficial approach for both sides and it can foster a greater feeling of trust for both.

All right, let's quickly talk about algorithms in ultrasound machines today, and we're going to be focusing on locked AI algorithms for the purpose of this discussion. We can ask how different or similar is the locked algorithm from kind of a conventional algorithm or an automated measurement, perhaps, that has been in ultrasound systems for the last decade.

The pathway for regulatory clearance is really no different fundamentally. There is obviously the more complex nature of AI algorithms. Kind of it's its own animal, but fundamentally the regulatory pathway doesn't really change. The submission elements are consistent. As Dr. Vaezy mentioned, the special controls associated -- that have been established most recently with Caption -- kind of can come together with special controls that have already been in place for ultrasound.

Of course, you need to talk about claims, intended use, labeling. These drive the performance characteristics and the qualification of the design, ultimately. And then today modern datasets are used to fine tune the AI algorithms. And using edge detection in an ultrasound image, as an example, image processing with conventional or AI algorithms can ultimately perform the same function. Now, obviously the newer AI algorithms are more elaborate and can do it in a different way. But if we follow the path through to market, the use of, I mentioned verification and validation here, it's generally consistent.

Let's go through this spaghetti chart to kind of explain, you know, what this -- what does this mean? So, the V model in design controls, so in the quality system regulations -- from a high level, on the left-hand side, so it's everything in blue, you have basically your requirements. What are your user needs, what do they need in order to use the product successfully, safely and effectively? And then the on the right-hand side you have the measures, I'll call them measures, where you basically check when you're done designing

your product, how do you make sure that you've made what you said you were going to make? You know, how do you make sure it meets all the requirements?

So over here is sort of a standalone depiction of kind of the start of data collection. The important thing to call out here is the discrete nature of the training set from -- here we go -- ultimately, the validation set or the testing set, depending onthe term that you use. And then finally, there's been a lot of discussion around postmarket surveillance.

So, what I wanted to stress here is that even AI, without AI, medical device manufacturers are obligated to ensure that their products that they've sold into the market continue to behave and perform the way they're intended to. Whether or not there's AI, the obligation is still on the manufacturers to ensure that the products maintain their originally intended performance.

This is a quick depiction of -- this is ultrasound guidance and the numbers up above are sections or clauses within the guidance. I just want to draw attention to some important areas and within this guidance, if we were to release an ultrasound system without AI, we're still obligated to show who's qualified to use this, who are the users? You know, what associated labeling, warnings, will go in place? What's in the instruction for use? This is all driven by user needs and those user needs are established by understanding who your customer is. So really, the introduction of AI doesn't really change this approach.

Okay, finally, key takeaways. One thing I wanted to stress is the existing marketing authorization pathways for medical device manufacturers are basically clear, they've been proven effective and they're predictable. It's been stressed numerous times that there are particular steps to move forward through the FDA to get a product to market. FDA does their very best to provide guidance to understandwhat steps need to be taken in order to achieve your goal, which ismaking it to market and having products used by our customers and impact their patients.

The medical imaging industry has already started to collaborate and in the context of AI, to better understand how can we put a common framework together, analyze existing frameworks and figure out whether there are any gaps in terms of a regulatory framework that needs to be put in place to manage AI development and demonstrating safety and effectiveness.

Development of consensus standards, so this is an important piece. While I stress that the existing pathways are clear and effective, the flexibility may really be with standards that may be adopted as consensus standards by the FDA. It's already been proven effective for DICOM, ISO standards and IEC standards. Why wouldn't this work for AI?

And then finally, continued development and access to AI algorithms by clinicians ultimately help them even more easily or further meet the Quadruple Aim, which will benefit our healthcare system, the patients, and everyone involved. Sorry for the lack of a better term there. And that's all. Thank you.

(Applause.)

DR. ASTARY: Thank you, Scott, that was great. Next up we have Berkman Sahiner, a physicist in the Division of Imaging Diagnostics and Software Reliability.

DR. SAHINER: Thank you. Over the course of the past 2 days you heard a lot about the potential of AI in medical imaging and especially today about AI-guided image acquisition. In the presentation, I would like to go a little bit deeper into methods for performance testing for these devices.

Here are some broad definitions about nonclinical and clinical performance testing in the context of AI-guided image acquisition. In nonclinical testing, one is trying to

characterize the technical performance of the system for image acquisition, and in clinical performance testing, the goal is to evaluate the diagnostic utility of the device when representative users use it on a representative patient population.

So why do we need both, both nonclinical and clinical testing, because the answer is because they complement each other so well. Nonclinical testing, because it's less resource intensive, can be performed on larger datasets and may be able to answer questions that require good precision, it may be able to answer some questions about generalizability. And also, it may be helpful about future iterations of the device, as was discussed extensively earlier. And of course, clinical testing helps us get a quantitative analysis of the expected utility of the device in the hands of users.

I'll go a little bit into more detail about each component, nonclinical and clinical performance testing. Nonclinical performance testing is the technical performance of the AI system by itself. In this part, an interaction between the AI system and the user in the intended manner is not necessary, although it might still exist and it can be component by component.

What do I mean by component by component? Your AI system may be such that there are multiple components stacked together, so you might have one component that estimates what the image quality is and as we heard, that's really the start of a good AI-guided image acquisition system. It might have another component that tells the user, for example, in the context of ultrasound imaging, how to move the probe or angle it to obtain a better quality image. So, these are the different components that you may be able test by themselves.

Some of the potential nonclinical testing questions for an ultrasound image guidance device might be, is the AI system able to assess or determine the image quality, closeness of the probe to the desired location to acquire a particular ultrasound view, desired

manipulation of the probe that will improve the image quality, and others.

Here are some basic steps. I hope that there are no surprises here, and this is somewhat similar to the quote that Dr. Hong gave abouthow do you solve a problem? You, of course, first define the nonclinical testing questions that need to be addressed. Then you define what the assessment methodology is going to be and what the metrics are going to be. The third step is maybe hardest for some AI systems, is to determine what the reference standard or comparator or ideally, what the ground truth is. And of course, you then go and acquire the data that you need to acquire for nonclinical testing and then provide relevant statistical data on the assessment results, subgroup analysis, etc.

So, this is quite generic. In the next few slides, I will go through them with an example that's more focused on AI-guided image acquisition. I'd like to emphasize that what I'm going to give is just an example. It's not necessarily what CDRH would ask you to do specifically in a submission, but just to illustrate some of the points that are listed in this slide.

Let's say that your nonclinical testing question is the algorithm performance in assessing image quality. As discussed earlier, this is really a basic, very first step. And let's say that for the assessment method you choose to use the comparison or the quality ranking for pairs of images between the algorithm and the experts. So, there's a pair of images, what does your algorithm give for which image is higher quality and what do the experts say? And let's say that your assessment metric, just to be specific, is concordance measured by Kendall's tau.

What's your reference standard or comparator? In this case it's sort of easy because that was in the definition in the previous slide, the comparator is the side-by-side comparison of two ultrasound images by experts. Let's say that for these two ultrasound images, your computer, your algorithm came up with the result that Image 1 has better

image quality than Image 2 and then you have, let's say, four experts and they also tell you which image they think has higher image quality. And then depending on the direction of the comparison sign, some experts may be concordant with the computer and some may be discordant. And just for completeness, I have the definition of Kendall's tau at the bottom. You look at the number of concordances and number of discordances to define what your Kendall's tau is.

Now, the dataset, needless to say, has to be independent from the dataset used for algorithm training. Ideally, it's a dataset that the algorithm has never seen before and you need to consider, of course, when you're collecting or choosing your dataset, the relevance and representativeness of the dataset for the component to be tested and some considerations for sample size.

Once you collect the data, then, of course, you look at your performance metric, but not only a point value but, you know, confidence intervals for the metric are very important. And then what are some of the other relevant comparisons? For example, if you collected image quality rankings from multiple experts, as I've shown before, then what is the concordance between them and how is the concordance between the computer and the experts compared to the concordance between them? And what are some relevant subgroup analyses? For example, if BMI is an important factor in your image analysis, what are your results with respect to the different BMI subgroups?

Now I move on to clinical performance testing and I'd like to say that when we say clinical performance, it does not automatically mean that all of the steps obtained are obtained in the clinic. For example, you might have a good virtual system that closely mimics some of the components in clinical testing and if you justify their use, you should be able to substitute them in clinical performance testing.

So clinical performance testing looks at the clinical safety and effectiveness of the

device used for its intended use when used by the intended user. And as I mentioned, it might be truly tested in the clinic, but it might also include accepted virtual or physical systems designed to capture the clinical variability, it might involve comparison to a closely related device with established clinical performance, and other sources.

In clinical performance testing, you want to get a quantitative evaluation of the diagnostic utility and quality of the images acquired or optimized and the performance in representative user and patient populations under anticipated conditions for use. Again, this might be a little bit generic, so I'll try to illustrate with an example and again, just in the case of nonclinical performance example that I gave, this is just for illustration purposes.

Let's say that your clinical performance testing is the comparison of diagnostic performance of experts when they interpret images acquired with standard of care or images acquired by a representative user population with the guidance from the AI algorithm.

To be more specific, let's say you have a patient with or without condition X. The upper row is the standard of care imaging, so you do the standard of care to acquire the image, then an expert interprets the image and provides the diagnosis for condition X, whether it's there or not. And at the bottom you have the new device, the image is now acquired with device guidance and then again, there is an expert interpreting the image and provide a diagnosis and at the end, you compare the diagnosis between the standard of care and the new device. And as I previously described, some of the blocks here can also be virtual if there is good justification for it.

The appropriateness of the clinical testing method will depend on device benefits and risks. So what I just gave was just an example about diagnosing a condition, but the comparison might be an image quality assessment comparison with the standard of care and the new device or it can be a comparison to a closely related device with established

clinical performance or can be meeting a benchmark, and this would be a good topic of discussion at the Q-sub meeting with the FDA, as discussed earlier in the session.

I'd like to now spend a few minutes on other important considerations of performance testing. One of these is the statistical analysis plan and here, pre-specification is the key. Some elements in a statistical analysis plan that should be pre-specified are the endpoints, statistical analysis methods, the process for defining the reference standard or the ground truth, statistical and clinical justification of sample size. If you're going to do multiple hypotheses testing, if you have multiple hypotheses, then a plan for that and a plan for handling missing data because in all the studies I have seen, although at the outset your goal is, of course, to have no missing data, there is some missing data and the goal is again, to pre-specify, before you see those missing cases, what you will do with them.

Also, statistical analysis for subgroups is important, where a sufficient number of cases from important subgroups and sometimes powering for subgroups may not be necessary depending on the clinical application and of course, unless specific subgroup performance claims are included.

Technological characteristics are important to provide with your submission. They are needed for understanding the scope of change in a modification if you're going to modify your system in the future, or comparing two devices, and they may reduce performance testing requirements.

As discussed earlier in this meeting, algorithm modifications, FDA's open to discuss plans for modifications during the initial premarket review and as outlined in our discussion paper last year, maybe a way to do this is to come to the FDA with a predetermined change control plan that shows the ability to manage and control risks of modifications at the time of your submission of the device.

Generalizability is an important consideration and it has heightened importance

because of the data-driven nature of most AI algorithms. So, it's important in all kinds of device submissions and we typically look at, for example, generalizability to different patient and user populations, but also in the case of AI, important considerations are how does this generalize to data acquisition equipment or conditions and data acquisition sites, etc.?

And I think there is a huge rule to be played by medical physicists in really bridging the data signs and clinical views and about the generalization to conditions that are or are not present in the training and test datasets. I think a good understanding of the physics of the problem and a good understanding of what are some of the sources of variability are, they're important to make arguments or to look at data for generalizability.

I'll stop here and hopefully we'll be able to expand some of these in the panel discussion. Thank you.

(Applause.)

DR. ASTARY: Thank you, Dr. Sahiner. Next up, Dr. Xin Feng, Human Factors and Reliability Engineering Team at FDA.

DR. FENG: Hello, everyone. Good afternoon, this is Xin Feng from FDA/CDRH's Human Factors Team. So I'm going to talk a lot here today about the human factor considerations for medical AI applications.

First is the disclaimer page that this only represents my personal position.

So hopefully what I'm going to talk about today is giving a simple introduction to human factors and them come up with an overview for FDA/CDRH human factors role and considerations in our premarket review of medical devices, and then present a few considerations for user interface design for the AI medical applications, and end up with a

hypothetical example of using AI-guided image acquisition to illustrate the human factors medical device process.

What is human factors? I'm sure many of us have experienced when we had difficulty trying to figure out how to use something like a microwave oven, figuring out how to use this water pot or how to use a presenter's remote, and we scratch our heads and we say to ourselves, this just doesn't make sense, right? So human factors is trying to figure out how to design a product that makes sense not only to engineers and developers, but also to the users who will use it. If you're inputting a message, users should understand it. If you're inputting a workflow, users should be able to follow it. If you put a controller, users need to figure out how to turn it on and off. And that's human factors.

I appreciate that Scott mentioned the QSR and user needs in his presentation not long ago. Really, the FDA/CDRH has a long history looking into human factors in medical device regulation. Back in 1996-1997 when QSR and design control were published, the Agency specifically stated that design input shall address the need of user and patient. Design validation shall ensure that device comfort to defined user needs and the intended uses.

And to further demystify the misconception that human factors is just afterthought and testing after the device development, FDA specifically mentioned in the QSR preamble that human factors studies, analyses, and tests should start from the early stage of design process until the point where the user interface is finalized. In other words, it should be a process integral with your device development.

So since then, FDA published several human factors guidance to share our current thinking on how to apply the human factors in the premarket review for medical devices. The current human factors guidance from CDRH was published in February 2016 and in that guidance we explained that FDA is primarily concerned about the safety effectiveness

perspective of human factors. We also explained that for the medical devices where the user can cause serious harm to the patient or user, we believe that the manufacturer should include the human factors data as a part of premarket submission. And we further explained that the harm here is to define include compromised medical care.

I'm sure this is very intuitive that use error on many medical devices can cause serious harm to patients directly. Now, when an AED doesn't deliver the intended shock to the patient or auto-inject or infusion pump doesn't deliver the medication as is expected to the patient, we all understand that can cause serious harm to the patient. Now, for people in this room, I guess the question is can the use error in a radiology imaging device lead to serious harm to the patients, right, because this is a device where you only make diagnostics instead of delivering the treatment directly to the patient. And the answer is it depends, right?

When the radiology imaging devices generate a wrong treatment decision or delayed treatment decision due to the use error, but that wrong treatment decision or delayed treatment decision lead to serious harm to the patient, then you should submit human factor validation as a part of your premarket submission data, and your risk management should consider both direct harm as well as indirect harm.

So here is an example to make that point more clear. This is where a use error which may lead to wrong treatment decision and will lead to serious harm to the patient. This is a CGM, continuous glucose monitor device, where the patients put a sensor on their body and it will read the continuous glucose value on the smart app or a dedicated reader. If it works as intended, this will reduce the effort of a patient manually tracking blood glucose value by using the BGM, right?

If you look at this picture on the screen, what you see is if you read the number, you read the trend, you read the color, what do you get? You get the patient was a single, the

number is in the right range, we don't need to do anything, our blood glucose value is good, right? But then if you turn to the labeling and instructions for use material, which explain all the user interface elements on that screen and you will see a little bit different story here and the difference there is an icon on the screen and if you read the IFU, it will further explain that at this time you shouldn't make a treatment decision based on the CGM value. What you need to do is you need to manually measure your blood glucose value by a blood glucose meter. So, this is an example where user interface design matters and if a user made the use error of not understanding the icon, he may make the wrong treatment decision, which may lead to serious harm to the patient.

The next thing is what is the user interface design considerations for this panel, what we are talking about here, the AI-guided image acquisition devices? And how are they, AI-guided image acquisition devices, different from other types of medical devices such as medical software applications? What are the commonalities in the user interface design, and what are the unique considerations and concerns?

For one thing, if we look at the Hollywood movies, right, I think whether we're thinking of Star Wars, Matrix, Ex Machina and so on, we'll see a lot of concerns on the AI and those concerns are, if I summarize briefly, that AI may take over the world and get us into trouble. And I would argue, as a human factors professional, we're truly not worrying about how powerful the AI may become one day. What we're truly worrying about is how AI work with humankind and how AI will communicate and interact with humankind and that, I think, is a human factors issue, is how we design the user interface and let the AI collaborate and work together with human more effectively, right?

So, some examples here. In talking about the AI and the human collaboration, what I fully agree with what one presenter presented in Day 1 is that even to the degree where there is a fully autonomous AI radiology system, if we put in the whole healthcare delivery

ecosystem, it still needs to collaborate with human operators and users, right, because the results need to be presented, interpreted to the patients by humans. Some humans need to formulate a treatment plan and some humans need to deliver that treatment to the patient. In the end, the AI need to collaborate with the human, with the users. And one example I can think of is I used to work in the trauma ICU and the most difficult time and most dangerous time for the ICU patient is the hands-off time and that's from one shift to another shift, information transfer from one nurse to another nurse, and that miscommunication happened, adverse events occurred, and that patient got injured, harmed, or even death occurred. So similar here, if we're developing an AI system we predict it will collaborate with human operators, we need to define clearly and present that clearly to the human, the users, what is AI's role and what is user's role, right?

And the second thing is that I often wonder how does user tell the status and context of an AI application? And I don't know how many of you routinely use the GPS app whether it's Google map or Apple map on your phone, when I'm on a roadmap, you know, after a few hours I always wonder where is this thing going to lead me, right? Is it going to lead me to Orange County, New York, which is 20 miles from my home, or going to lead me to Orange County, California, which is 2,000 miles from my home, right, because the app just doesn't show up where is the destination anymore. And I think this is truly applying to our discussion in this room, is even though AI is working, AI is getting users through some task, we need to give the user the clear state status and the clear context of the current task that they are working on, right?

To echo the concern by the Hollywood movies, I think it's also important, when we design those AI applications, we need to ensure that even when AI is working, the user always has the option to regain the control all times. And if the user wants to take over the control from AI even when AI is working, there needs to be a way for a user to do that.

And finally, the question is how does the user verify the result? And we think this is particularly important for lay users, novice users, home users, because expert users may develop some workaround and they know how to do the tricks bypassing the application. But for home users and novice users, it is important for them if they have a second dot on the result. The developers should let them know how do they verify the result. An example would be the previous example on the continuous glucose monitor and blood glucose meter, right? In most of those CGM manufacturers' labeling, they will tell users if you have suspicions regarding the result, you should use the BGM, the blood glucose meter, to verify it. So that's a way for the lay user to verify the result and the question is how do we provide such a pathway to verify the results for what we're discussing here?

So you know, what I've said through the 2-day workshop and I heard my NIH colleague just mention this, what we're discussing here, the AI in radiology imaging applications are still needing a lot of research, the field is new and at the edge as to prove the first de novo in AI-guided image acquisition. But other AI applications have a long research history, right? The human-computer interaction community has been developing the research, design guidelines, considerations for how AI interact with human for a long time, 20, 30 years. So maybe a starting point for us here in designing the AI-guided image acquisition app is to look into the example design guidelines for the general human-computer interaction and human interaction with AI and look at what we can learn from here, right?

What I'm showing here is example by Microsoft research and they put together the design guidelines for human-AI interaction, they summarize the research from the past 20, 30 years. Not every one of them will apply to our field, but I think it's a good starting point to look at what may apply.

So here is a hypothetical example that we're using AI-guided image acquisition

devices for early screening of deep vein thrombosis, right? This application story, this is hypothetical but many elements are from real submissions. This is for family members, caregivers, lay users use and use in the home environment. And the company did the use-related risk analysis and identified that one potential use error which may lead to serious harm is the user fails to follow the instruction and the labeling or they fail to understand this is a poor-quality image, they need to take further action, so this missed opportunity of early thrombosis diagnosis, which could potentially cause serious harm to the patient.

Now, the company did the right thing, they properly identified the use-related risk and then they run a human factor validation study, included the representative lay users in the study. And while the task is to show a picture with the labeling to the test -- ask them what this means, what they need to do, right?

When engineers develop those software applications, they work with their consultants and medical physicists, radiologists and they say what does this mean, and everyone seems to understand, low res meaning low resolution. Now they bring in the representative home patients, they ask what does it mean, and a few home patients don't understand what low res mean and actually one of them say low res meaning lay low and get some rest, right?

(Laughter.)

DR. FENG: The question we need to ask ourselves is, does the message make sense to the lay users and home users? And the company did analysis and figured out what's the reason to cause the confusion and they modified the labeling message, changed it to regard to low-quality image, please try the test again.

Now by no means this means the Agency endorses this design, but I think we can get two messages here. One is if you want to present a message to your intended user, you need to figure out what is their experience, knowledge, and comfort level to those

messages, they need to understand those messages, right? And the other thing is a need to present them with the clearly next step action that they should take. Those are things that we learned from the human factors study and user research and the company was to screen a labeling message and they run the human factors testing through the validation study again and they show that now the patient get the message, they understand the message, they understand what is the necessary next step action.

And that example hopefully works with medical device human factors engineering process which is presented in our guidance, which is need to -- when you develop a medical device, you need to run the risk management, understand the use-related risk, identify the critical task and if there is a serious harm caused by the use error, then you need to submit human factor validation data to the Agency to demonstrate that the risk mitigation matter is effective and the use-related risks are low or reduced at the acceptable level, right? So, the process map is included in our guidance.

So the take-home message is that human factors is really not an afterthought or a testing methodology after design. Human factors is a process which you should include it in your design process from Day 1. And human factors is about understanding the user needs and leveraging the design guideline best practice in designing the medical applications. If you find some issues in your design and you identify some use-related issues, repeating the testing won't solve the problem, right? Design modifications will address the issue. Re-doing the testing, simply re-doing the testing one more time doesn't make those issues go away.

And finally, for our human factors testing at FDA for medical devices, we're using a risk-based approach and the goal here is to ensure the device-user interface has been optimized to support safe and effective use. And we always welcome early discussions via the pre-submission program with FDA on your human factors protocol, on your risk analysis,

so we can have an early understanding of the risk, the human factors design strategy, the human factors test strategy, human use.

And thank you, that's all.

(Applause.)

DR. ASTARY: Thank you, Dr. Feng. Okay, I'll introduce our last speaker for this session, it's Andy Milkowski, Senior Director of Ultrasound R&D at Siemens Healthineers.

MR. MILKOWSKI: We'll check off the human factors, yes, it reaches the guidance talk. For those that know me, probably find it ironic and perhaps quite humorous that I am the only non-regulatory person speaking in the regulatory part of today's presentations.

So, a couple of items. As everyone probably knows from today's presentations, everybody believes that AI can/will raise the standard of care, it will improve the components of Quadruple Aim, and personally, I really applaud the FDA's initiatives to create a more streamlined and effective regulatory oversight.

However, being a product development guy, I think there's a few opportunities that I'll highlight in this presentation and I'll perhaps start by illustrating some of the things that Siemens has done, well, three examples and then I'll go through three opportunities. Not that that's a comprehensive list, but everybody can remember three.

And perhaps in the start of this, I think today we're still at a level of immaturity. If I were to look at an AI system hierarchy, quite arbitrary, but we put it together, you'll notice that most of the things that we talk about are still at those bottom rungs. And we can meet here years from now and have another excellent meeting and discussion and really talk about more sophisticated applications.

A number of you are probably familiar that Siemens, a few years ago, released the

first kind of camera-based patient-positioning system powered by AI. At the top of that image you probably see a nice camera and that will automatically locate and segment the patient out and determine if he's centered, determine which way they're pointed and it leads to all kinds of benefits, right? It modulates the dose, it knows where the patient's head is, so it can obviously go to the right organs and scan. The benefits, obvious, it's faster, more consistent and in this case, less radiation, the quality is improved and it works well. Experts, novices, everything is just better, it's one of these fantastic cases of AI. And our current 510(k) mechanisms were adequate for releasing this product on the market.

A second example, magnetic resonance, and a similar story, kind of AI-based positioning based on patient biometric information can do things such as automatic alignment. It could do bolus timing and therefore help with improved follow-up. And in fact, one of the studies that Siemens has done is to determine now we've had enough of these AI tools on the MR side and about 90% of the ordered exams all have an AI component to them. And again, all the benefits are similar to the prior CT example, right, reduced setup time, improved variation, improved follow-up. Yeah, this picture should've been in here. And again, in this case the current 510(k) mechanisms were appropriate.

And since we're in the ultrasound day, I thought I'd throw in an ultrasound example and this has been available, again, for years and in this case, image recognition based, similar to the other stuff that we've looked at today. Automated, one-button kind of stuff, automated measurements of spectral Doppler, LA volume, LV measurements, all done by AI. And for those that are familiar, LV measurements, for those that try to do it visually, large discordance between users or those that attempt to do their segmentation manually, large eventual repetitive stressing degree, right? So again, similar story here, much faster, more consistent, both for experts or novices and in a recurring theme, the existing 510(k) mechanisms are adequate for releasing this type of product.

So that was the quick kind of overview. Now I want to switch over to what I call opportunities. And earlier today, Drs. Martin and Feltovich really talked about the applications in the underdeveloped world and from a risk-benefit perspective, let's say there's fundamentally no risk, all benefit. However, in this context we're talking about the FDA, we're all here in the developed world and how about, let's say, we're all here in a highly litigious world and so therefore our risk-benefit analysis has to be slightly more nuanced, perhaps.

I'll pick on three topics. First is the adaptive algorithm piece and for the purpose of this slide, I'm going to group both the locked and discrete adaptive algorithms together and the reason I'm grouping them together is only for the purpose that the improvements are applied through a new software version. So even if you have a discrete adaptive algorithm, if you even have like a federated model in which you're doing kind of a privacy enhanced learning at a location, but you are not actually releasing that or the machine's not changing until there's a conscious decision by the manufacturer to update the software, I'm throwing that into the same category as a locked algorithm because there's a conscious decision to release something after some testing.

That being distinctly different than an adaptive algorithm where software modifies in the field on that device without interaction by the user and it's based on some preprogrammed reward algorithm. And you can already start thinking well, was the reward algorithm properly setup for that intended purpose, right? And so those changes are done without regulatory oversight.

Now, I'm also further going to exclude things like workflow enhancement things, right? If a user has a tendency to go through certain steps, I'm not even going to cover those because anything that has, let's say, a non-diagnostic implication I can leave out. Let's think about the serious implications where there could be a diagnostic case and a lot

of previous presenters pointed some out.

What could some of those from my slide be challenges with adaptive algorithms? The second list on there is validation, maintainability, traceability. I'll characterize those as manufacturers' challenges and, you know, for the purpose of today, those are manufacturers' problems. If there is an objective here and if we need to go down that route, manufacturers will somehow figure that out and there's been excellent work done, at least that I know about, by the Air Force in terms of their adaptive algorithms for controlling airplanes and I'm sure the medical industry can figure out the same thing.

On the flip side, though, think about the liability implications of repeatability and reliability and you guys can imagine scenarios better than I can probably characterize things such as same patient gets measured on the same day by the same machine, perhaps at two different site locations and gets how many different diagnoses. Explain that one as a manufacturer, right, especially if you get called on it for whatever reason.

Or the other one I have listed is unintended consequences. Now, Dr. Lehman yesterday talked about at MGH, just the number of staff that they have and the randomized controlled clinical trials that they do and probably they have the resources to think about what would be the implications of adaptive algorithms in their setting in their population. However, we also know that every community hospital does not have those resources and you're likely introducing the same type of, let's say, bias that the algorithms -- that the AI is supposed to help you take out in the very beginning.

Perhaps I could summarize this by saying that I don't know if -- and I'll say it more strongly. We're not sophisticated enough today as an industry, as a whole, as a community, to say we know and can predict and can say that we have benefit over risk in the area of adaptive algorithms and furthermore, I believe we can achieve all of the benefits with adaptive, with discrete adaptive algorithms. And I think the best way probably to move

forward in this area is probably work with the FDA and industry and figure out what's the best plan forward. This is true today. Will it be true years from now? Probably not.

Second point, non-expert user. A lot of discussion was held over the past couple days over autonomous users. I'm going to redefine that for the purpose of this discussion. Once you put non-expert users in control of an ultrasound system, I want to forward the argument that that is now an autonomous system, meaning somebody of knowledge, an expert user, a radiologist, somebody else that has the appropriate knowledge and background -- we could talk about the human factors kind of impact -- is no longer able to intervene in a timely fashion. So, then it puts the non-expert user into the role of confirming the device is operating properly and that they're operating the device properly. Perhaps we can work through those items.

A few more things such as, you know, today's medical imaging -- if the devices are being used in a hospital, they're in a lockdown environment, there's security, there's all kinds of provisions in a hospital. Now the devices are going to be, well, let's say in home use and the images are getting sent hither and to all over the place, including the manufacturer and such and perhaps the privacy issue is outside of the data -- outside of the FDA's purview.

However, I would argue that the FDA could have a safety arm in that from the direction of ensuring that all the products, if there was a privacy issue with the product, through the vehicle of, let's say, making sure that the devices can be updated in case of some kind of issue, we already know that these issues exist, right, in hospitals themselves. Lots of private personal information has been exposed on the web, right? We're opening up another conduit for that, for that problem to occur.

Lastly, the missing actionable information liability. Today we're basically talking about cardiac images with the Caption Health release and which that's the fairly

constrained problem, but you know that these are the first days in this. We're going to cover many other applications. Think about the prior presentation and the DVT example. Now you have a location problem. How do you know you have the appropriate images being sent and diagnosed, right, because you no longer -- you have an autonomous system, you have a non-expert in use in that game.

And perhaps I saved what I consider the most important one is the foreseeable misuse and I hesitate almost to say this because, you know, in this space where you have expert users, you have medically trained folks, labeling is an entirely inadequate mitigation for a lot of risks because there is training associated with that. Even in the case where there's nurses versus sonographers being involved, it's understandable. Once you move beyond the basic medical training, one needs to consider whether labeling is adequate. We again live in the society that we do.

I'll just throw out a couple of things. Ocular scanning anyone on a device designed to scan for cardiac issues, or how about of a mother who's already lost a pregnancy before, nervously checking her unborn how many times a day, how many hours a day for how many weeks on end. I know there's no bio-effects associated with ultrasound. I wouldn't want to defend that one in case anything happened, whether it had to do anything with ultrasound or not.

So just questioning again the sufficiency of labeling mitigation and I think the proper approach is probably working, you know, the FDA and industry working together to figure out how do we best get ahead of that before it introduces any significant issues?

My last message is the regulatory framework. Again, as I said in the beginning, I really support and we support the FDA's initiatives in creating a more streamlined and efficient regulatory oversight, even if it is purely self-serving from handling a workload perspective, right? Whatever the motivation, it's fantastic.

The concern is, however, that you may introducing much preferential approval pathways as kind of the -- as the flowchart kind of illustrates, once an initial 510(k) is acquired, if you have a good approved SPS and ACP, you could easily see that you've just given yourself a broad way of just doing letters to file and expanding indications for use, if you structure everything properly.

We have a lot of traditionally approved devices that could also benefit from these improved pathways. Perhaps there is an opportunity to leverage -- and I'll say that there's been a lot of opportunities with Q-submission suggestions, that perhaps that is going to be the vehicle -- that we can utilize the existing approval pathways, yeah, take advantage of some of the opportunities that the FDA has kind of brought up through these white papers.

So, kind of in summary, industry has a lot of success in implementing cool AI applications. Three challenges that I brought up is just that most adaptive algorithms should probably today be managed through a discrete adaptive method. Non-expert users probably introduce some new risk profiles that we should question whether labeling is probably good enough as a mitigation. An existing framework should be updated probably in line with some of the recent proposals.

As I close now in the last seconds, I threw up the Garner hype cycle for a reason. For all of those involved with AI, they've known that artificial intelligence has gone through this hype cycle, arguably more than once, and that is from that peak of inflated expectations, or for those who know Alan Greenspan's comment of irrational exuberance, down to this trough of disillusionment. We have an exciting time in front of us. You see all of the neat, new developments, you see all of the applications that are already out there. You probably should be conscious of some of the risks we may be introducing in our irrational exuberance and make sure we take the proper precautions because this is litigious society and we are a medical industry. Thank you.

(Applause.)

DR. ASTARY: Thank you, Andy, that was great.

At this time, I'd like to invite all of the panel members for the afternoon session to come up. We'll give you about 5 minutes to get up to the stage and take your seats. Thank you.

(Pause.)

DR. ASTARY: Okay, we will go ahead and get started with the afternoon panel discussion. I realize we're getting kind of close to the end of the day, people need to leave to catch flights, so some of our panel members may need to excuse themselves in order to make their travel arrangements and we understand that, certainly.

I'd like to start this by having the panel members introduce themselves, starting at your side, Berkman, and then just kind of working your way around to me.

DR. SAHINER: Hi, I'm Berkman Sahiner. I'm a research scientist with the Division of Imaging Diagnostics and Software Reliability at Office of Science and Engineering Labs at CDRH.

DR. VAEZY: Hi, my name is Shahram Vaezy. I'm a biomedical engineer in the Division of Radiological Health in OHT 7, part of OPEQ in CDRH.

MR. SURETTE: Hi, I'm Sam Surette. I'm the head of regulatory affairs and quality assurance at Caption Health.

MR. RODER: Hello, my name is Tony Roder. I lead the regulatory affairs team for the imaging division at GE Healthcare, and I'm also the vice chair of the MITA X-ray section.

MR. PAULSON: I'm Scott Paulson, Senior Director of Regulatory, Quality, and Compliance at Fujifilm SonoSite.

DR. MILKOWSKI: Andy Milkowski, and I make stuff at Siemens Ultrasound.

DR. LAM: Hi, I'm Benny Lam. I'm with Philips Ultrasound, I'm responsible for the AI and digital health strategy.

DR. GARRA: I'm Brian Garra, I'm a practicing radiologist and a researcher at the FDA in the Office of Science and Engineering Laboratories.

DR. FENG: Xin Feng, human factor reviewer, CDRH/FDA.

DR. ASTARY: Thank you. Thank you for the introductions and thank you for participating in this discussion, we appreciate it.

I wanted to start off with a question for Dr. Garra and Sam. Dr. Garra, you bring a physician's perspective on AI-guided imaging acquisition to the panel and, Sam, your team recognized the opportunity to improve healthcare by developing a device utilizing this technology. I'd like to get your perspectives on how this technology could benefit patients and practitioners and just kind of expand on some of the things you touched on during discussions earlier today.

DR. GARRA: Yeah, of course, at this stage of the meeting we've covered a lot of the things, so some of them will be repetitious, but I think as Dr. Martin said, we get -- the big benefit is we have more access to high-quality imaging and we have faster access and I think that's the main benefit. Faster is helpful in an emergency situation, for example. More access is helpful in a rural environment, for example, where people have no imaging capability whatsoever. Even in the U.S., in Appalachia, for instance, people have no access to imaging and there's a lot of potential for us to fix the third-world part of the United States by having more access and more skilled access in places where there's not much skill.

There's another factor here. Have you noticed that when you have these AI applications, people change the paradigm a little to make it easier for the AI application to function? Caption Health, they didn't try everything an echocardiographer would do, they kind of limited the scope and ferreted out the key elements that you need to do to make a

diagnosis, and a lot of our medical imaging systems have never done that.

For instance, in particular, ultrasound. The paradigm for ultrasound was established in the mid-1950s and the key there in that area was, it was very hard to save images and very hard to review them, so you had a very skilled person just do a few images. They would look at the screen and say we need that image, we need that image. These days the paradigm has shifted with PACS, where people can review hundreds or thousands of images and decide for themselves what images are important. So, it's time that we relooked at a lot of those imaging paradigms and I'm hoping that the AI, advent of AI gives us a chance and an opportunity to modernize those.

Just one other thing. In our review of Caption Health and other AI systems, you can see there's an improvement in quality even for the expert who's using that system. I do all forms of ultrasound except echo, but I think with Caption Health I get to expand my capability to doing echo, as well. But I know many times where I get a good image, I never know if it's the best one I could've gotten and I'd like to have somebody that's non-threatening tell me whether I could do a little better or not. Thank you.

DR. ASTARY: Thank you. Sam, do you have anything to add?

MR. SURETTE: Yeah. I mean, I completely agree with Dr. Garra. I think there's been a lot of focus today on the nurse user that was user in our pivotal study, but we see it as enabling nurses but also physicians and other medical professionals. Even in a hospital setting that sonographers are working on a full workup, you know, an emergency room, ICU, hospitalists, all of them want to use point-of-care ultrasound, but the barrier to entry is a little too high without AI guidance.

And Dr. Garra also reminded me of something really fun that happened last year, is we took our machine to FDA and to demonstrate it live, just so that all of the features could be understandable to the reviewers and Dr. Garra and others actually scanned on a patient

model using it. So, Dr. Garra has some personal experience with AI-guided ultrasound.

DR. ASTARY: Thank you, Sam. I'm going to move on to our next question since we're a little tight on time. Scott and Andy, during your presentations you both kind of touched on how AI-guided image acquisition can contribute to improving, you know, the Quadruple Aim or contribute to trying to achieve the Quadruple Aim in healthcare. So, I just wanted to see if you could expand on aspects of the Quadruple Aim that may be impacted by this type of technology.

MR. PAULSON: Sure. I'll focus on maybe a couple elements. I guess the general quality of care, there's been discussion about -- I mean, a point being like creating a level of excellence or, you know, leveling the playing field, basically, having consistent performance regardless of the user. If you're able to achieve that, then the quality of care across the board goes up, naturally. The patients ultimately benefit from that.

Another aspect really is it's kind of new way to look at this, but with the care team well-being, if you, as a manufacturer, can come out with a product that helps reduce stress, eases workflow, the workflow that ultimately falls on the caregiver, so you probably can increase, naturally increase efficiencies at the point of care, but what if you reduced the stress level for the caregiver so that they actually do a better job of delivering their care, related or unrelated to your specific device, you know, that's maybe another aspect to that.

DR. ASTARY: Thank you.

DR. MILKOWSKI: Yeah. AI is probably the fanciest, newest mathematical tool we have, so the use of this is going to be absolutely ubiquitous and the examples that everybody bring up are perfect examples of whether it was just the third world and Appalachia kind of expanding access, improving healthcare that way, or the prior discussion about improved results from scanning even by expert users. So, it will be everywhere, those are two examples where.

DR. ASTARY: Thank you both. I'll move on to our next question. Sam, congratulations to your team on your recently granted de novo. I just wanted to kind of get your thoughts on how you prepared for your submission and how you looked at areas where you needed to provide testing data, areas where you might need to kind evaluate the performance of your device, what you saw as the critical areas to evaluate this new type of device.

MR. SURETTE: Sure. So in addition to being the head of RA/QA at Caption Health, before joining Caption Health I was a lead reviewer at FDA in the Division of Cardiovascular Devices and so one of my missions when I left the Agency was to bring, you know, the values and the animating principles of the FDA to Silicon Valley and to Caption Health, specifically. And a lot of times I'll get asked by people, both, you know, new people in Caption Health but also just casually, what does FDA want to see and I would pause at that. It's up to the company to do their homework and figure out what's the clinical problem they're trying to solve, how can they show that with evidence and, you know, bringing that benefit-risk principle that FDA talks about often to Caption Health.

And sowe looked at kind of three things. The first, of course, was the clinical study that we would put together to support the de novo, that's really the core articulation of the effectiveness of the device, all of the different features of the product in the hands of a nurse who had never scanned before, to show that they're effective at acquiring a limited TTE exam. That was the bedrock of our submission.

But we also had two other components that I think were very important. One is human factors. Our device is a usability device. Its effectiveness is making ultrasound usable for users who would otherwise not be able to use it effectively or at all. And we really took the human factors guidance from FDA to heart, we looked through it and we worked with our user experience designers to develop the usability documentation and

validation, which I'm very proud of.

And the third thing is the algorithm testing and nonclinical testing. You know, there's two aims for us at Caption Health. One, of course, is to support the functionality of the device for the de novo and then the other is to be able to articulate iterative improvements to our product over time. We obviously can't run a pivotal study with 240 patients every time we optimize the product. And so, showing the effectiveness of different aspects of it in a nonclinical way is very important to us for iterative development.

And then I think the last thing that I would remark upon just as a former FDA reviewer and a current employee at Caption Health, is that a de novo is really setting a precedent for future innovation and so it's innovation. Just like our company's promoting innovation with our products, you know, we're working with FDA on regulatory innovation and I think that's very important, and something that we thought about a lot internally at Caption Health and in our discussions with FDA, is how can the evidence that we provide in this form a model for the future? And I think I'm very proud of the special controls that came out of it and I think it should facilitate innovation there as well.

But yeah, I think it's really about thinking about the benefits and risks and I always tell folks inside Caption Health, prove to me that this works and then it will be easy to convince the FDA.

And the one recommendation I have for other companies, particularly in Silicon Valley, is to get clinical expertise. We benefit from expert sonographers. Sonography is absolutely an art and what we do every day is to try to emulate their expertise with AI and bring it to places where sonographers aren't available or places where you want that expertise to go. And we also work with incredible cardiologists as well. So clinical expertise is very important, especially because we're trying to use AI to emulate what they bring to image acquisition.

DR. ASTARY: Thanks, Sam, that was great. Does anyone else have anything to add to that great response?

(No response.)

DR. ASTARY: All right, I'll move on to the next question here, so kind of on a related topic. So, Shahram, you talked during your presentation about the special controls that were developed as part of this de novo granting, creation of Regulation 892.2100, but I just wanted to get your thoughts on, you know, what factors kind of contributed to developing the special controls and how you were keeping in mind potentially other imaging modalities this regulation may be extended to. So just kind of how those special controls were crafted in this case.

DR. VAEZY: Continuing on what Sam was describing about their work on device development, I would like to say the special controls really is -- should be done, as we discussed, as I presented, as has been discussed in the context of the risk mitigation, benefit-risk analysis and in reviewing the Caption Health submission and developing the special guidance, of course, we wanted to develop -- the special controls, sorry. We wanted to make sure that these are the requirements that Caption Guidance has met in showing reasonable safety and effectiveness of their device, and also try to cover as many devices as possible under this regulation, devices of the same type, because then, you know, by having this regulation, we established the regulatory pathway basically using -- now manufacturers can come in with a predicate, basically, using the 510(k) submission mechanism. So that's really the goal, that was the goal for developing those special controls.

And you mentioned other imaging modalities. Indeed, as you look at the regulation, there is no specific mention of ultrasound imaging, so we hope that it would also capture other types of devices, you know, using different modalities.

Also I could just comment, perhaps it will come up later, too, that devices that are --

as I mentioned, that would be for home use, for example, could be also covered under this regulation. So, it's really beneficial when we try to develop special controls that are beneficial to the community in general, for making a pathway, regulatory pathway that makes it faster and easier for manufacturers to develop devices and bring it to patients.

DR. ASTARY: Thank you. And, Brian, do you have anything to add or --

(Off microphone response.)

DR. ASTARY: Sure, yeah.

DR. GARRA: I like the fact that they're general and because the more specific you get, the harder it is to generalize it. I just wanted to -- I was thinking about the other modalities and that would be a little hard if we were -- and there's been a lot of discussion about moving to post-approval monitoring and you can see there might be some problems, but in the radiology community, it was mentioned that we have ways of checking people's performance over time and we do it by spot-checking. So, we might do every -- require somebody to do every tenth case. Not only can you have AI help with that process because it can select cases and check them maybe itself without an operator doing it, but also that's a good way to check AI.

But then if you have a system like Caption Health which uses ultrasound, you could say well, periodically you need to double scan traditionally and with the Caption Health product and see how well it stacks up. Now, that wouldn't generalize well to CT because then you're applying a double dose of radiation to a bunch of your patients, so you have to think carefully about how to do that.

On the other hand, CT doesn't have quite so much user variability as ultrasound. It's basically what order to you press buttons and things like that. I know I'm denigrating CT technologists, but -- but there are some things that they AI could do to help CT and MR. For instance, one of the big problems for radiologists is they never use the same field of view

size.

If want to compare images and I want to see whether their tumor is getting bigger, I'd like to have the same scale so that I don't have to measure every single lesion to know whether they're getting bigger or smaller. But for some reason they just won't do that and I think it's because the system doesn't remember what it did before and the person is too busy to bother to look that up. An AI system that checks for obvious common errors like that would be very, very beneficial. It shortens our workflow tremendously and doesn't impact negatively the workflow of the person performing the exam.

Another common one, and I'll stop after that, is MR. MR takes a lot of time to do an acquisition, so you often have blurring due to patient motion and time pressures being what they are in a lot of private imaging centers, they don't bother if there's motion blurring, to take another one. They hope the radiologist will say it's too much work to get them to go back and do that, so we'll just read through it. And I actually do that myself. I'm not happy about it, but -- but the system could detect that blurring and flag the person and say go back, you need to repeat this and this is why it happened. Don't do that again or reset the system. Perhaps it's a sequence that's too slow for that patient and it could reset it automatically. So, there's some opportunities there.

DR. ASTARY: Thank you. I'm going to move on to our next question. There's a desire to foster predictability in the regulation of AI-guided image acquisition systems. Predictability benefits industry by providing clear expectations in terms of submission content and benefits the FDA by promoting quality submissions. This question is for Tony and Berkman and we'll start with you, Tony, but I want to know how might industry and FDA work together to foster predictability in the regulation of this expanding area of technology?

MR. RODER: Yeah. I mean, I think the first thing I'd say is FDA is obviously focused

on this, right, and the attempt with the white paper and with communicating the various pathways, so that's, you know, all in the right direction. I think we face a couple challenges of predictability in this space. One is the rapid pace. With predictability or in order to foster predictability, transparency is super important, how you make decisions, when you make decisions, you know, what's good, what's not good. When you introduce a rapid pace of development on top of our historical practices for communicating transparency, like guidances, special controls, standards, those tend to lag, right, those tend to not evolve at that same pace. So that's one thing. You know, I mean, sometimes it comes down -- I think the quote is coming back again, you just write it down. Sometimes you just have to write it down, right? We struggle with where to write it down, can we write it down, are we allowed to write it down, frankly, right? So that's something, as we're working together in this, we can -- we have to think creatively around the various mechanisms that we can do that.

You know, what can we work on to increase predictability? We've talked about it. Like if there's certain expectations on performance or the types of things to describe or attributes that you need to evaluate for various clinical tasks, I mean, that's very technical but those are things you can just start grabbing that problem statement and working on. I think having a broad enough stakeholder community when you're evaluating these things, too, right? One group kind of running with their plan is never a good thing because that introduces maybe issues later on, right, if something's used for unintended consequences.

And then the last piece. Love the comment, over-specificity here can really cause us (1) to slow down and cause us lots of problems. So you have to give enough guidance so that we actually have something to go on, but if we get over-specific around exactly how to do something, then what we'll end up doing is having to go and do, you know, special controls for absolutely every single device that comes out and that just is untenable.

So I think there's a lot there, but I think we've got a good history in developing standards and working together on various things and that, I think, is going to be an important pathway as we hit clear problem statements and then use our venues to address them.

DR. ASTARY: Thank you.

DR. SAHINER: I'd just like to add a few points to what you said, Tony. I think one of the most important things is actually to speak the same language, the terminology, because there's a lot of new terminology introduced by new concepts in AI. So, what do I mean by continuous learning or locked or, you know, many of the other terminologies, you know, validation, training, even testing. I think it would be a great start to have, you know, some defined glossary, you know, by input from everybody involved, industry, academia, you know, other agencies and the FDA, just to standardize what we're talking about.

And I think other things could include development of standard practices, again, with collaboration from all stakeholders. And likewise maybe sometimes, you know, just getting together in workshops or, you know, other venues like this, because I think these are great steps to increase predictability. And I think we can also try to make sure that what we are developing or what we are working on together is consistent with other industries and first of all, other healthcare sectors but also other industries, so that again, we're all on the same page with the others.

DR. ASTARY: Great, thank you. I want to switch gears a little bit and talk about usability again. Xin, so this question's for you. You know, the AI-guided image acquisition systems may have various instructions for users, it might be verbal, graphical, or written. I want to get your thoughts on how might the FDA work with industry and the public to collaborate on how to best design these instructions to make them most effective and make the annotations to the devices more seamless.

DR. FENG: Sure. I will start to give some opinions from FDA perspective and, you know, company folks on the panel, feel free to chime in. I think from the FDA regulatory perspective, it's pretty clear, you know, in our human factors guidance we consider labeling as a part of user interface element, right? So when this user interface element that should follow the human factors design process, as we discussed in the presentation a few moments ago, which is you identify what is the intended user use environment, what are the use-related risks, what are the critical paths, how do you mitigate that risk and going through an iterative design process and try to figure out user interface design and risk mitigation strategy, right?

But if I take off my FDA regulatory hat and put on my human-computer interaction researcher hat for a moment and I would say, just looking generally at what we're doing here, which is sharing knowledge and knowledge can largely be broken into two categories. One category is declarative knowledge, which is the medical device is intended for use by patient who is 18 years old and above. This is a poor-quality image, you need to retake it again. So that declarative knowledge we can learn quickly by reading text, right?

And what we're discussing here, by using the AI-guided image acquisition, for example, to guide the non-experienced user, home user, to do something, I think this falls into a later category, which is the procedural knowledge. So, examples like learning tennis, learning to play golf, learning how to using the sensor to collect an ultrasound image, and that's a procedure. And like my medical officer friend often told me at FDA, you learn those medical procedures by see one, do one, teach one, right? You learn it by someone demonstrating it to you, giving you feedback, encouragement and correction, you learn it by watching.

And this makes me think it's a good topic, that the format of the labeling is very important here, which is the text may not do the work to deliver that procedural

knowledge. We need to think about, you know, this is a case where a picture is worth thousand words and -- may be worth thousand pictures, right? Give the users a directional demonstration through multimedia labeling material and demonstrate to them how to perform those procedural tasks. And then I think, again, as an HCI researcher, I think an ultimate aspiration goal might be levering the sensor WE-R (ph.), AR, AI technology and not only just deliver that multimedia knowledge in steps, it's nice to say this in total has 19 steps and Step 5 you should do this and that, but to give to the user in context help, right? You have just pressed a little bit too deep and that is what you need to correct, and that is really help to the user in that guidance and throughout your labeling material.

DR. ASTARY: Thank you. Other comments or I'll move on to the next question.

(No response.)

DR. ASTARY: AI-based algorithms can be adapted to potentially improve performance by training on expanded datasets that capture a wider range of imaging additions, patient populations, and user populations and so this question is for both Andy and Benny. Regarding adaptive algorithms, can you just kind of discuss some of the risks and benefits of these algorithms?

DR. MILKOWSKI: I'll refer to my prior comments.

(Laughter.)

DR. MILKOWSKI: I think we all know that any of these systems will benefit from additional data and will get better and I think there's an unlimited, perhaps, potential from workflow usability opportunities that AI can provide. And I think in the prior presentation I tried to make the point where if there is a diagnosis available, my belief just from surveying the best of academic literature is we're probably not sophisticated enough to unleash adaptive algorithms, just as the society's practices.

And from a risk-benefit perspective, we probably have alternatives such as discrete

adaptive where there is a conscious decision to release those and the message is kind of we do not need to get out ahead of ourselves, to use a Wile E. Coyote kind of analogy, we don't need to get ahead of ourselves and perhaps create a risk where we could get all kinds of unintended regulations to come upon us if we mess up in that space. But that's just a risk-benefit kind of calculation.

DR. LAM: So, let me deal a little bit for looking. I'm going to give some hypothetical example of a -- some benefit of using adaptive AI in radiology. Let's say we have a PACS system that can adaptively adjust the display setting, for example enhancing certain part of the image based on the radiologist reading habits on Friday. So basically, if I was a radiologist, I read it and look at different places and throughout my reading the AI algorithm, I understand that these are places I am interested in, these are the places I would pay attention to. For other time, you can show that you may direct my attention to the place that would be interesting to me. In that case I would see a potential to increase the productivity of my work.

Another hypothetical example would be like an ultrasound that can effectively control the transmission elements, like beamforming, etc., on Friday to adapt to each clinician's scanning habit to obtain the image of a sufficient diagnostic quality. In that case it can potentially reduce the clinician's variability in scanning because it kind of standardized the way -- intrinsically standardized the way that the -- how the clinician or sonographer do the scanning.

So having said that, you know, we have talked about -- a lot of talk about the risks of, you know, this kind of AI algorithm, but then I would say we have to do a biased approach, you know, in many -- not only on AI but also -- I mean only applying AI to medical radiology imaging but also in many other industry like auto driving and all this kind of stuff, if we can, you know, sensibly have a realistic risk and benefit balance and we -- I can envision that

eventually we will get to something that really can benefit everyone. Thanks.

DR. ASTARY: Thanks, Benny. Go ahead, Sam.

MR. SURETTE: Yeah, I just wanted to maybe take a more positive perspective on the future of adaptive algorithms. You know, I think an analogy that I find useful personally is manufacturing. In the medical device industry, we are very familiar with how to manufacture products, how to do process validation, how to control for variance in our devices. So, when you manufacture a pacemaker or a pacemaker lead, not every pacemaker lead that comes down the manufacturing line is going to be identical, right? And we've spent decades developing the right way to monitor, control, develop good manufacturing practices to control the manufacturing process of a device that's already been validated on a design validation perspective.

I think for machine learning, good machine-learning practices are very analogous to good manufacturing practices in the sense that you could foresee a system that you develop as a company where you the algorithm is allowed to change or allowed to operate within certain bounds, learn on new data and then automatically check itself against a help-out dataset and if it passes that check and it's somewhat analogous to a process validation, then it can be allowed to evolve or develop in real time.

I don't think it's this far-off thing that we have no experience with. I think manufacturing is a very good analogy for what this could be in the near future.

DR. ASTARY: Thank you, Sam. Anyone else have comments here?

(No response.)

DR. ASTARY: All right, so I'll move on to the next question. The question's for you, Tony. The performance of AI-based algorithms is driven in part by the data used to train the algorithms, so I wanted to get your thoughts on what are the best practices for providing transparency of the data used to train the algorithms and also to evaluate these

algorithms?

MR. RODER: Yeah, so I mean, again, this kind of comes back to the transparency piece and really ties into predictability again. The first thing is, you know, we talked a little bit about the training versus testing or validation datasets, an exact perfect example of, you know, terminology and what's what. But when you think about the training dataset, there's little that can be determined from that, from describing that. There's information that can be inferred or, you know, maybe understood a little bit about maybe how the device might -- what you think it might do, but there's not a lot that can be determined. So, you shouldn't spend all your energy on going in great detail there, right?

Where instead we should focus on more -- you know, and that should be more of a high level, just understanding, you know, what was the methodology, you know, what kind of variables did you cover, really focusing more on, you know, that testing and validation set that's really defining your performance or what you're going to say is your performance in what you consider to be the real world or the situations that the device is going to encounter in the real world.

And when you look at that, you're really trying to figure out how you're identifying your performance at the operating points, right, again providing context to the user, and what you're looking at is really reporting that performance in those variables that drive different performance of it.

The things that we really have to look at are patient factors, morphology, age, gender, etc. It depends, very much depends on the clinical task at hand, right? So again, it's going to be specific to the clinical task, looking at hardware variables and actually breaking down those subsets and being clear and transparent in the performance that you had in those so that at the end, the user actually has an ability to understand what type of performance am I going to expect in this population. And that's probably more beneficial

than really the general performance level of, you know, area under the curve, overall, for the dataset that you looked at, because your dataset is invariably going to be different than anything that that device encounters in the real world at that one clinical institution.

I think those are the main things. I think the other thing, again, you know, that understanding of that performance is not going to be equivalent in all the groups and again, the performance that you really look at in your testing dataset isn't necessarily going to match performance that's experienced when it hits a normal caseload.

DR. ASTARY: Thank you, Tony. Okay, I'll move on to the next question. So foreseeable implementation of AI-guided image acquisition, as we heard about in several talks earlier today, is enabling non-experts to acquire these images. And so, Shahram, I'm going to get your thoughts on what are the regulatory pathway or what is the regulatory pathway for home use of these devices?

DR. VAEZY: Yeah, thank you for the opportunity to say that it's going to be very hard to have very general guidelines of what regulatory pathway would be applicable to home use devices. The reason is that it really depends on the device, its indications for use, intended use, the user population, the patient population, the environment of use. So, all of those factors come into that determination.

The important thing is that I hope it's been demonstrated in this workshop that FDA is open to such use, if you will, open -- and I mean, the entire day has been devoted to AI-guided image acquisition, signal acquisition and we do see, especially for ultrasound with point-of-care devices, that this is on the horizon and if you really want to take advantage of the technology, this type of use, home use, should be allowed and now we have a mechanism with this regulation to foster and facilitate such use.

So as has been said, I guess, 50 times, the best approach is to engage with the FDA on device development efforts as early as possible, early interaction is very good and look

at FDA regulators also as advocates of the technology, as has been, I hope, demonstrated through establishment of this regulation.

DR. ASTARY: Thank you, Shahram. Anybody else have anything to add? Go ahead, Scott?

MR. PAULSON: Yeah, to add to that, I guess as an employee of a medical device manufacturer and a regulatory person at that, I can't -- I have thought about this a lot over the last year, you know, where is it going, where is point-of-care ultrasound going it as it gets smaller and smaller and more portable? And then there's the question of home use and define home use, is it use in the home by whatever prescription from your provider or is it over the counter? You know, two different animals, but they're both departures from what we have today.

And it's been mentioned a few times that besides the ability to gather or properly acquire the image, you know, there's other safety considerations as well. So you know, there's the physics and bio-effects associated with excessive use and though ultrasound is generally considered a relatively safe modality, the negative implications of overuse isn't really defined. So, I think, you know, I know that the FDA takes the position of maybe being a little conservative with that because we don't know. But yeah, there are a lot of considerations there. If we get over that mountain, it's a mountain of work, yeah.

DR. ASTARY: Thank you, Scott. I'll move on to the next question. Kind of on a related note, how do human factors considerations change with home use and is labeling sufficient to mitigate these risks? Andy, you kind of touched on this in your presentation, or kind of alluded to this, so I'll get your thoughts and maybe follow with Xin, your thoughts from a human factors perspective.

DR. MILKOWSKI: That's a neat question. I think the key message that I had is that if you remove an expert user, however we end up defining an expert user, the question is that

for all of us that work in companies, is labeling going to be sufficient if you can foresee misuse? I gave a couple of examples. I'm sure if we went around we can all think of other examples. I believe in many of the things even that I mentioned, there are probably reasonable technical solutions, as well, right, and if you are not getting the expected image, for example, in the cardiac images, if you are repeatedly not getting cardiac images, shut the machine down for a while, you know, or limit its ability to be used in the course of the day.

I think labeling is the last kind of resort and if there are reasonable technical solutions, we should implement them because in the big picture this is a brave new day, it's fantastic, we're going to improve population health. This device is excellent, you can't say enough. And, you know, personally my heart goes out the patients this morning, Isabella and Harsh and just if we can create devices to help, we have to do that. Let not our exuberance get in the way of doing the right things.

DR. ASTARY: Thank you.

DR. GARRA: Yeah, go ahead. I have something to say, but you can go ahead.

DR. FENG: All right, well -- which, you know, if you look at the medical device risk management standard 14971, which categorizes all the risk mitigation measures in three different categories, right -- design, protect and measure risk management by training and labeling. And labeling and training are Category 3, risk management -- which means is least effective and least effective for a number of reasons that labeling is not always available to the users and we all understand that. And research shows that users may or may not refer to the labeling material during the actual use, it's their own choice.

Now, for the question, is labeling sufficient to meet risks that may be introduced, I think it is another "it depends" question, right? So, when FDA look at those risks, we look at the risk level, the severity level of harm, the probability, and I think we also need to look

into very specific design questions here and there are scenarios where there are practically not a lot of choices to further mitigate the risk.

It's hard to give a universal yes or no question, again, that when we are doing the premarket review of human factors, we are taking a risk-based approach. So, we have to evaluate the risk level and consider the specific design questions and answer the question whether labeling is sufficient or not. But back to the bottom line where I started, that risk is mitigated by labeling and training is always considered the least effective option.

DR. GARRA: That was great. I want to give an example of an unintended consequence that labeling alone would not protect against. We know how to run a global outreach program and there's some countries we can't go into because they can take an ultrasound system and misuse it in certain ways. One of them is determining the sex of a baby and aborting the little girls and you say well, that's in some countries, it doesn't happen in the U.S., but a friend of mine, Greg Lynn, that was a regulatory guy for Diashonix (ph.), did a study in Southern California, we're seeing a shift towards more males and fewer females.

This can happen anywhere and we have to have some deliberateness use protection and that's one of the things. There's inadvertent misuse, there can be deliberate misuse and that's something that labeling alone will not protect against. Andy alluded to that, that you have to monitor the use of the machine and shut it off or -- and also maybe have, you know, like a PID card to turn the machine on or something, but of course, that can be given away. So, we have to think about all these unintended uses and protect against as many as possible.

DR. ASTARY: Thank you, Brian. We're nearing the end of our time here, so I just want to take a second to thank all our presenters again for taking your time to produce these presentations and share your thoughts about some things. Thanks to all the panel

members for the insightful discussions, very much appreciated. And thanks to the audience for sticking with us through this last session, so thank you.

(Applause.)

DR. LAMB: Hi, everyone. So, thanks for sticking with us to the end of this second day. Jen introduced me yesterday, but in case you've forgotten, I'm Jessica Lamb, I'm the acting assistant director for Mammography, Ultrasound, and Imaging Software. That's the MUIS up there.

I just kind of wanted to go over some highlights for today. So, you know, some of the things that we heard about this morning included kind of ways to standardize and improvement measurements and the image acquisition and then diagnoses that result from those things.

We heard about specific applications in urology, cardiology, maternal and fetal medicine. I don't know about you, I particularly appreciated Dr. Feltovich's illustrations of how hard it might be to capture measurements on a very moving fetus.

And one of the things that I think we heard about over and over this morning and throughout the day is accessibility and the potential benefits of accessibility to patients in the areas of battlefield or world areas where there may be fewer experts, but also in the home and the images and it'd be good for patients to have access in the home.

I really appreciated also the patient perspectives that came a little later in the morning and thank Joshua Basile, Harsh Thakker, and Isabella and Amy O'Brien for joining us and sharing their points of view. I know I heard from a lot of people at lunch how much they really valued getting that perspective and, you know, I think that hearing how there are indeed today situations causing to think differently about how they engage with healthcare than those of us who may not have quite as regular interactions with doctors

and imaging, at least from a patient's perspective. And it really continued the theme of accessibility from this morning, the potential game changer that could be in some lives.

One of the final notes from this morning that kind of came out of the panel discussion is also the potential paradigm shift from sort of discrete sort of monitoring paradigm to something that's sort of more continuous and the possible advantages and disadvantages that that kind of paradigm could have.

This afternoon we got into the public talks. One of the themes that a lot of the public talks talked upon was explainability and both how important it can be to, when these AI algorithms are implemented, provide some basis when possible to illustrate both to users and the doctors to interpret the information.

And we also appreciated hearing about the public and private funding and policy effort to these in this rapidly evolving field.

The regulatory considerations, I think, were of interest to a lot of you in the audience. One thing that came out of those slides was the importance of clear messaging and it kind of echoed some of the earlier comments about, you know, I think everybody appreciates the need for transparency in any aspect of AI implementation and any device implementation, but too much information or the wrong information at the wrong time or relying overly on labeling might not be sufficient to mitigate risks. So those are just interesting comments from today.

So going back to kind of yesterday, this is a slide that Robert Ochs had and I kind of liked this because it has this big open box that's going into the future and the things that we're going to need to talk about as we're going into the future, which is why we're all here today.

The thing that I think was the most common theme that I heard throughout this workshop was the belief that there's a need for continuous postmarket monitoring of these

devices in some way or another. And some of the comments that people made yesterday that highlighted this for me as a potential problem was a statement in one of the public talks, that we often don't know when AI is wrong, and an even earlier statement in one of the panel discussions, that we sometimes don't know what performance practitioners have in many areas of practice, even where there isn't AI and I think this kind of illustrates the kind of data infrastructure that we can work on together in order to bring these devices to market in a way that benefits patients and providers and everyone else in the healthcare ecosystem.

So the hopes for AI, you know, I think we all appreciate the hopes for earlier diagnosis, better treatment, more standardized imaging, just sort of the enabling -- and I think somebody put world-class healthcare anywhere, kind of regardless of the expertise that might be in that physical location.

And another thing that I think is really important is that it will allow for better quality doctor-patient engagement and experience by helping to automate some of the tasks that might currently keep a doctor from engaging with their patient and just integrate seamlessly into a system and allowing them to spend more time with patients. Of course, that's the hope and the concern is on the other side, maybe it will be just be that much more complicated.

So, this conversation isn't over. As we mentioned yesterday, the public docket is still open, so if you have further comments, I know there are a lot of people that wanted to contribute that didn't get the opportunity. Please do visit the public docket page.

As several people have already mentioned today, if you are developing a device, you can submit to our pre-submission program and I put some Google search terms in for those of you who aren't -- look up the slide and find that link later. When I made this slide I was thinking about the need to talk about data separation, that's a pretty common question and

one that I kind of wish people would ask more because it's a pretty big hang-up sometimes.

But following on some other themes, in terms of the AI white paper that people have mentioned, if you have thoughts about how you might to implement that in a way that will work for your AI algorithm, even if it's not continuously learning but you are interested in releasing a number of kind of discrete locked algorithms and you think some of those principles could apply, we'd be very interested to talk to you about that, and also any available data collection strategies you have that you think can support further development of your devices, you know, I think we'd be really interested to speak to anybody who has, you know, kind of developed strategies that will support device development.

Also, if you're making something that's not a medical device but it's intended to support the development of medical devices, we have the Medical Device Development Tool program, so this is another way for you to engage and perhaps make something that could be useful for the development of these products.

FDA, we're out there and we participate in standards development and professional societies and working groups, scientific meetings and publications, so we look forward to speaking with you folks in those capacities, as well.

Jen showed these slides yesterday, but I just want to highlight again that we're the Division of Radiological Health in CDRH and we collaborate heavily with the Office of Science and Engineering Laboratories to do research in this area. And there are four branches within the division and all of us review AI-based applications. If you have anything that you want to talk to us about, here are all of our names. And as you've perhaps already guessed, and you can see now from Kyle Myers, the algorithm to figure our e-mail addresses -- so it's not very hard.

So, another couple points of contact, the Rad Health mailbox is another way to

contact the Division of Radiological Health. Kyle Myers is the director for the Division of Imaging Diagnostics and Software Reliability. You heard from a number of members from that division, they do a lot of research, regulatory science, and if you are not looking in at imaging algorithm but some other kind of software-based app, I'd encourage you to reach out to the Digital Health Program.

I want to thank Jennifer Segui, who's the chair of this workshop and has really tackled an enormous logistical problem and scientific program, making sure that we have a fantastic program, balanced content. We got, you know, representatives from patients, from industry, from the public. This meeting would not have come together without her constant oversight and so we really appreciate her.

(Applause.)

DR. LAMB: I'd also like to thank the organizing committee, the executive committee and the many other folks who worked to make this event come together. Thank you also to the presenters, the panelists, the moderators, all of you, the public contributors who presented the content that we've been listening to for the past 2 days, and also NIH for the use of their space and their resources to -- including the AV resources, so we have a venue. And that's it. So thank you all again.

(Applause.)

(Whereupon, at 5:12 p.m., the meeting was adjourned.)

C E R T I F I C A T E

This is to certify that the attached proceedings in the matter of:

PUBLIC WORKSHOP - EVOLVING ROLE OF ARTIFICIAL INTELLIGENCE IN RADIOLOGICAL IMAGING

February 26, 2020

Bethesda, Maryland

were held as herein appears, and that this is the original transcription thereof for the files of the Food and Drug Administration, Center for Devices and Radiological Health, Medical Devices Advisory Committee.

_____

TOM BOWMAN

Official Reporter

Free State Reporting, Inc.
1378 Cape Saint Claire Road
Annapolis, MD 21409
(410) 974-0947