UNITED STATES OF AMERICA

FOOD AND DRUG ADMINISTRATION

+++

CENTER FOR DEVICES AND RADIOLOGICAL HEALTH

+++

MEDICAL DEVICES ADVISORY COMMITTEE

+++

CHEMISTRY AND TOXICOLOGY DEVICES PANEL

+++

October 20, 2022

9:00 a.m. EST

PANEL MEMBERS:

| | |
|---|---|
| Karol E. Watson, M.D. | Chair |
| Wang, Ping, PH.D., DABCC | Voting Member |
| Jennifer Higgins Ph.D., M.B.A. | Consumer Representative |
| Wilson Compton, M.D., MPE | Temporary Voting Member |
| Adam J. Gordon, M.D., MPH, FACP, DFASAM | Temporary Voting Member |
| Timothy J. Ness, M.D., Ph.D. | Temporary Voting Member |
| Sherif Zaafran, M.D., FASA | Temporary Voting Member |
| Walter S. Dunn, M.D., Ph.D. | Temporary Voting Member |
| John T. Farrar, M.D., Ph.D. | Temporary Voting Member |
| Anne-Michelle Ruha, M.D. | Temporary Voting Member |
| Brian T. Bateman, M.D. | Temporary Voting Member |
| Laura J. Bierut, M.D. | Temporary Voting Member |
| Lawrence S.B. Goldstein, Ph.D. | Temporary Voting Member |
| Cheryl Walker, Ph.D. | Temporary Voting Member |
| Colleen Gallagher, M.D. | Temporary Voting Member |

| | |
|---|---|
| Elijah Wreh, M.S. | Industry Representative |
| Elizabeth A. Joniak-Grant, Ph.D. | Patient Representative |
| Kellie Kelm Ph.D. | Director, Division of Chemistry and Toxicology Devices |
| James Swink | Designated Federal Officer |

FDA PRESENTERS:

Peter Mr. Yang – Training on De Novo Program

Program Lead for the De Novo Program in the Division of Submission Support in the Office of Regulatory Programs in the Office of Product Evaluation and Quality in CDRH at FDA.

Ouided Rouabhi – Breakthrough Device Designation Program

Assistant Director, Policy and Operations Team 1, Office of Clinical Evidence and Analysis, Office of Product Evaluation and Quality, CDRH

K. Melody Gussow, Ph.D.
Toxicology Team Lead, Division of Chemistry and Toxicology (DCTD)

OHT7, OPEQ, CDRH

INDUSTRY PRESENTERS:

Keri Donaldson, MD, MSCE – Study Design and Results

CEO of SOLVD Health

Joseph Garbely, DO, DFASAM, FAPA—Epidemiology, Current Practice Guidelines and Unmet Need

American Society of Addiction Medicine

American Psychiatric Association

Faculty, Penn State and Drexel University College of Medicine.

Christine Brauer, PhD – Additional Analyses Performed to Address FDA Questions
SOLVD Regulatory Affairs Consultant

Chris Zacko, MD – Clinical Perspective
Penn State Health Milton S Hershey Medical Center

Chris Mullen, SAMHSA Statistician

OPEN PUBLIC HEARING SPEAKERS:

Andrew Kolodny, MD
Brandeis University
President, Physicians for Responsible Opioid Prescribing

Dr. Michael Abrams Senior Health Researcher with Public Citizen's Health Research Group

RECORDED COMMENTS:

Dr. Suzet McKinney DrPH, MPH

Dr. Brand Newland, PharmD, MBA

Dr. Kamran Hamid, MD

Dr. Eric G. Fox, DDS, MS

Dr. Guillermo Chacon, DDS, FACS

Richard Jones, MA, MBA, LCAS, CEAP, SAP

Ken Kaufman

Jodi Barber

Ken Daniels

Megan Barry

Whitney Kannaka

Chris Fox

Joe Janasek

Cal Beyer, CWP, SCTPP

Brett Large

Bradley Sorte

Jeff Horwitz

INDEX

CALL TO ORDER

Dr. Watson: Good morning. I would like to call this meeting of the Clinical Chemistry and Clinical Toxicology Devices Panels to order. I'm Dr. Karol Watson, the Chairperson for today's meeting. I note for the record that the voting members present constitute a quorum as required by 21 CFR part 14. I would also like to add that the panel members participating in today's meeting have received training in FDA device law and regulations. For today's agenda, the committee will discuss, make recommendations, and vote on clinical information related to the de novo request for the AvertD test, sponsored by SOLVD Health.

Before we begin, I would like to ask our distinguished committee members and FDA attending virtually: please turn on your video monitors if you've not already done so and unmute your phone before you speak. I will call your name. Please state your area of expertise, your position, and affiliates. Before we start, I would like to note that I am Dr. Karol Watson. I am a cardiologist specializing in preventive cardiology in women's cardiovascular health at the David Geffen School of Medicine at UCLA. Ping Wang PhD, DABCC, please introduce yourself.

Dr. Wang: Good morning, everyone. My name is Ping Wang. I'm a Professor of Pathology and Lab Medicine here at the University of Pennsylvania. My specialty is clinical chemistry and toxicology. Thank you.

Dr. Watson: Thank you. Jennifer Higgins, please introduce yourself.

Dr. Higgins: Dr. Jennifer Higgins. I am regular member of the APAC Committee, and I'm a consumer representative to today's committee.

Dr. Watson: Thank you. Elijah Wreh, please introduce yourself.

Dr. Wreh: Hello everyone. My name is Elijah Wreh, and I'm the rep for industry on the FDA Advisory Committee Panel. I work for Boston Scientific, and my role is Senior Manager for Regulatory Affairs at Boston Scientific. Thank you.

Dr. Watson: Elizabeth Joniak-Grant, please introduce yourself.

Dr. Joniak-Grant: Hi, I'm Elizabeth Joniak-Grant. I am a sociologist, but I'm here today as the patient representative. I represent chronic pain and chronic migraines, and so I'm happy to be here today. Thank you.

Dr. Watson: Wilson Compton, please introduce yourself.

Dr. Compton: I am Wilson Compton. I'm a physician scientist and the Deputy Director at the National Institute on Drug Abuse. My area of expertise would be addiction psychiatry and epidemiology.

Dr. Watson: Adam Gordon, please introduce yourself.

Dr. Gordon: Good morning. My name is Adam Gordon. I'm an internal medicine/addiction medicine physician, Professor of Medicine and Psychiatry here at the University of Utah at Salt Lake City. My area expertise is opioids and opioid addiction.

Dr. Watson: Timothy Ness, please introduce yourself.

Dr. Ness: Hi, I'm Tim Ness. I'm a practicing anesthesiologist and pain clinician from the University of Alabama at Birmingham. My expertise relevant to this is that we have studied genetic markers in relation to analgesic responsiveness and pain responsiveness, but not addiction.

Dr. Watson: Sharif Zaafran, introduce yourself.

Dr. Zaafran: Morning. Sharif Zaafran. I'm the President of the Texas Medical Board. I'm a standing member of the ADPAC Committee, former Subcommittee Chair of the HHS Pain Management Task Force.

Dr. Watson: Walter Dunn, please introduce yourself.

Dr. Dunn: Hi, good morning. Dr. Walter Dunn. I'm a psychiatrist and a Professor of Psychiatry at UCLA and the West Los Angeles VA.

Dr. Watson: John Farrar, please introduce yourself.

Dr. Farrar: I'm John Farrar. I'm a professor of neurology, epidemiology, and anesthesia at the University of Pennsylvania. I've been involved in pain clinical trials for most of my career, and I'm very interested in the epidemiology and methodology of clinical studies.

Dr. Watson: Julie Lysis, please introduce yourself.

James Swink: Julie could not attend today.

Dr. Watson: Anne-Michelle Ruha, please introduce yourself.

Dr. Ruha: Hi, I am Michelle Ruha. I am a medical toxicologist and addiction medicine physician. I'm a professor at the University of Arizona College of Medicine and Chair of the Department of Medical Toxicology at Banner University Medical Center in Phoenix.

Dr. Watson: Brian Bateman, please introduce yourself.

Dr. Bateman: Good morning. Brian Bateman. I'm a professor and Chair of the Department of Anesthesiology Perioperative Pain Medicine at Stanford, Chair of the ADPAC Committee, and my research expertise is in pharmacoepidemiology.

Dr. Watson: Laura Bierut, please introduce yourself.

Dr. Bierut: Good morning. My name is Laura Bierut. I'm a physician scientist at Washington University in St. Louis. My expertise is in psychiatry, addiction, and genetics. Lawrence Goldstein, please introduce yourself.

Dr. Goldstein: Yes. Good morning. I'm Lawrence Goldstein. I'm a distinguished Professor Emeritus at UC San Diego, and I'm Scientific Director of the Sanford Consortium for Regenerative Medicine. My areas are genetic cell biology and neuroscience. I have experience in drug development, and I'm a member of the National Academy of Sciences.

Dr. Watson: Cheryl Walker, please introduce yourself.

Dr. Walker: Yes, good morning. I'm a professor and Director of the Center for Precision Environmental Health at Baylor College of Medicine. I'm a molecular biologist with expertise in epigenetics.

Dr. Watson: Colleen Gallagher, please introduce yourself.

Dr. Gallagher: Good day everyone. I'm Colleen Gallagher, and I serve as a healthcare ethicist and professor at the University of Texas, MD Anderson Cancer Center.

Dr. Watson: Kellie Kelm, please introduce yourself.

Dr. Kelm: Good morning. I'm Kellie Kelm. I'm the Division Director of the Division of Chemistry and Toxicology Devices at FDA. Morning.

Dr. Watson: Thank you all. James Swink, the Distinguished Federal Officer for the Clinical Chemistry and Clinical Toxicology Devices Panel will make some introductory remarks.

James Swink: Good morning. I will now read the Conflict of Interest Statement. The Food and Drug Administration is convening today's meeting of the Clinical Chemistry and Clinical Toxicology Devices Panel of the Medical Device Advisory Committee under the authority of the Federal Advisory Community Act of 1972. With the exception of the industry representative, all members and consultants of the panel are special government employees or regular federal employees from other agencies and are subject to federal conflict of interest laws and regulations. The following information on the status of this panel's compliance with federal ethics and conflict of interest laws covered by, but not limited to, those found at 18 USC Section 208 are being provided to participants in today's meeting and to the public.

FDA has determined that members and consultants of this panel are in compliance with federal ethics and conflict of interest laws. Under 18 USC Section 208, Congress has authorized the participation of FDA employees and regular federal employees who have financial conflicts when it is determined that the Agency's need for a particular individual’s services outweighs his or her potential financial conflict of interest. Related to the discussions of today's meetings, members and consultants of this panel who are special government employees or regular federal employees, have been screened for potential financial conflicts of interest of their own as well as those imputed to them, including those of their spouses or minor children, and, for purposes of 18 USC Section 2 0 8, their employers. These interests may include investments, consulting,

expert witness testimony, contracts, grants, credos, teaching, speaking, writing, patents and royalties, and primary employment.

For today's agenda, the Panel will discuss, make recommendations, and vote on clinical information related to the de novo request for the AvertD Test sponsored by SOLVD Health. Information from AvertD Test provides patients 18 years of age or older and healthcare providers with objective information to be used for informed decision-making prior to the first prescription of oral opioids for acute pain. The AvertD test is intended to be used in combination with clinical evaluation and assessment of the patient.

Based on the agenda for today's meeting and all financial interests reported by the panel members and consultants, no conflict of interest waivers have been issued in accordance with 18 USC Section 208. Mr. Elijah Wreh is serving as an industry representative, acting on behalf of all related industry. He's employed by Boston Scientific. We would like to remind members and consultants that if the discussions involve any other products or firms not already on the agenda for which a FDA participant has a personal or imputed financial interest, the participants need to exclude themselves from such involvement, and their exclusion will be noted for the record. FDA encourages all other participants to advise the Panel of any financial relationships they may have with any of the firms at issue. A copy of the statement will be available for review and will be included as a part of the official transcript. Thank you.

I will now read the Panel's Memo. For the duration of the Clinical Chemistry and Clinical Toxicology Devices Panel meeting on October 20th, 2022, Doctors Brian Bateman, Laura Bierut, Walter Dunn, John Farrar, Lawrence Goldstein, Adam Gordon, Timothy Ness, and Michelle Ruha, Sharif Zaafran, and Cheryl Walker have been appointed to serve as temporary

voting members. Dr. Jennifer Higgins has been appointed to serve as a contemporary non-voting consumer representative, and Dr. Elizabeth Joniak-Grant has been appointed to serve as a temporary non-voting patient representative.

For the record, Dr. Brian Bateman serves as Chair of the Anesthetic and Analgesic Drug Products Advisory Committee, or AADPAC, in the Center for Dev Drug Evaluation and Research, or CDER. Dr. Timothy Ness also serves as a voting member of the AADPAC. Dr. John Farrar serves as a consultant to the AADPAC, and Dr. Jennifer Higgins serves as a consumer representative for the Analgesic Drug Products Advisor Committee in CDER. Dr. Laura Bierut serves as a voting member for the Tobacco Products Scientific Advisory Committee in the Center for Tobacco Products, or CTP. Dr. Walter Dunn, a regular government employee, serves as a voting member of the Psychopharmacologic Drugs Advisory Committee, or PDAC, in CDER. Dr. Adam Gordon, a regular government employee, serves as a consultant to the PDAC in CDER. Dr. Lawrence Goldstein serves as a consultant to the Cellular Tissue and Gene Therapies Advisory Committee in the Center for Biologics Evaluation and Research, or CBER. Dr. Elizabeth Joniak-Grant serves as a patient representative consultant for the Arthritis Advisor Committee in CDER. Doctors Michelle Ruha and Sherif Zaafran serve as consultants to the Drug Safety and Risk Management Advisor Committee, or DSRMAC, in CDER. Dr. Cheryl Walker serves as a voting member of the Science Advisory Board to the National Center for Toxicological Research, or NCTR.

These individuals are special government employees or regular government employees who have undergone the customary conflict of interest review and have reviewed the material to

be considered at this meeting. These appointments were authorized by Russell Forney, Director Advisory Committee over Insight and Management Staff on September 21st, 2022.

I will now read the Appointment to Temporary Voting Status Memo. Pursuant to the authority granted under the Medical Device Advisory Committee Charter of the Center for Devices and Radiological Health, dated October 27th, 1990, and is amended August 18th, 2006, I appoint the following individuals as voting members of the Clinical Chemistry and Clinical Toxicology Panel for the duration of this meeting on October 20th, 2022: Dr. Wilson Compton and Dr. Colleen Gallagher. In addition, I appoint Karol Watson, MD to act as temporary chairperson for the duration of this meeting. For the record, these individuals are a special government employees or regular government employees who have undergone the customer conflict of interest review and have reviewed the material to be considered at this meeting. This has been signed by Dr. Jeffrey Sheerin on September 27th, 2022. A copy of this statement will be available for review and will be included as part of the official transcript. Thank you.

FDA encourages all other participants to advise the panel of any financial relationship that they may have with any firms at issue, and here are a few general announcements as follows. In order to help the transcriber identify who is speaking, please be sure to identify yourself each and every time that you speak. The press contact for today's meeting is Laura J. McCarthy. For the record, FDA has received one written comment. Thank you very much. Dr. Wilson.

Dr. Watson: Thank you. I wanted to mention that in addition to the written comments submitted to the FDA, several comments were also submitted to the docket. These documents have been sent to the Panel and FDA for review. We will now proceed to the FDA's presentation on the de novo program. I would like to invite FDA to begin. Dr. Mr. Yang?

Mr. Yang: Good morning. My name is Peter Mr. Yang. I'm the Program Lead for the De Novo Program in the Division of Submission Support in the Office of Regulatory Programs in the Office of Product Evaluation and Quality in CDRH at FDA. This panel meeting this morning is for a de novo request, and so, as part of training this morning, I want to provide a brief overview of the de novo process, help you understand what we're asking for in its regulatory context, so that you understand what FDA will use your input for, as far as moving forward on this submission.

Before we get to what a de novo request is, I want to set the stage by reviewing how medical devices are classified. In general, every type of medical device is classified by federal regulation into one of three classifications, Class I, Class II, or Class III. It is a risk-based classification based on what FDA believes is needed to ensure that the device is safe and effective. I'm going to start with the lowest level of risk, which are Class I devices.

Class I devices are those for which general controls are all that's needed to provide reasonable assurance of safety and effectiveness. General controls do apply to the other devices too, but for Class I devices, that's all there is. So that includes things like registration and listing of manufacturing facilities to support inspections, good manufacturing practices, we call them quality systems, medical device reporting or adverse event reporting, and prohibitions against misbranding or adulteration. These devices are generally exempt from FDA pre-market review, which means that companies can sell these devices without prior permission from FDA.

Class II devices require a combination of general and special controls to provide reasonable assurance of safety and effectiveness. Special controls are legal requirements that can include things like specific bench testing requirements, labeling requirements, specific clinical

data, or post-market requirements. Class II devices are cleared through the 510K process and substantial equivalents. Basically, what that means is these devices send a pre-market submission to FDA and say, “Based on this data I'm providing to you, I believe that my device is substantial equivalent to an existing legally marketed predicate device.” In other words, “I am as safe and effective as this other one that's already on the market.” And small improvements over time, within each regulation, within each type of device, allow that device space to evolve.

Class III devices are generally for the highest level of risk and require a combination of general controls and pre-market approval. Pre-market approval is a specific type of application to FDA that requires demonstrating reasonable assurance of safety and effectiveness for the device from first principles. They don't say that they're as safe and effective as an existing legally marketed device. They're proving for their device that it's safe and effective from first principles. It's part of an ongoing process that includes review and scrutiny of manufacturing, changes to the product ongoing annual reporting requirements where the company is analyzing their own product for safety and effectiveness and rigorous conditions of approval. That includes post-market requirements as well. So, devices are approved through the PMA process. They're cleared through the 510K process, or they're exempt from FDA pre-market review entirely, all based on their classification.

Now, we finally get to what a de novo request is. A de novo request is a pre-marketing submission, so like a 510K or PMA, but a de novo request is actually for new types of devices that are low to moderate risk that are otherwise automatically classified into Class III. So just briefly: if a device is new, if we haven't seen it before, by default, the Federal Food Drug

Cosmetic Act classifies that device into Class III and requires a PMA for that device to be legally marketed.

But, as you could imagine, not every kind of new device has the same kind of really high risk profile that we would associate with something that's reviewed and approved through PMA. A de novo request is actually a request that FDA formally classify the device into Class I or Class II based on a determination of reasonable assurance of safety and effectiveness. A de novo request is not a substantial equivalence decision. We're not saying that this device is as safe and effective as a predicate. We are formally classifying the device into Class I or Class II, so that if we grant that de novo request, we actually write a brand new classification regulation and then that device then becomes probably Class II. Then it becomes regulated through the 510K process.

The device that you're considering this morning, today, becomes the first predicate device of its kind. And then future products that are similar to this one that you're reviewing will say, “I am as safe and effective.” They will go through the substantial equivalence process, and that's how they'll be legally marketed. They'll compare themselves to the device that you're reviewing today. So, this one becomes the first one of its kind.

Okay, so I'll just skim through some of these other slides briefly, so you understand some of the regulatory criteria that are important here. We have to make sure that it's a new device. In order to be eligible for the de novo process, first, it has to be a medical device. It can't fit into an existing classification regulation, so it can't be found substantial equivalent to other kinds of devices that we reviewed in our history. It can't fit into an existing Class III regulation. There can't be any approved PMA for the same type of device, and by type of device, I mean a

combination of the intended use and the technology. It basically can't be like anything that we've seen before. That's the high level summary of the eligibility criteria for the de novo process.

As part of any classification procedure, we have these three goals to meet before we can grant the de novo request for this device. The first one, and I think this is the most salient goal to the one that you're discussing today, is to determine whether the probable benefits of the device outweigh the probable risks when the device is used as intended. The second goal is to identify what the probable risk to health are for the device or product, and, based off of that list of risks to health, determine the level of control that is needed to mitigate those risks. And we've already talked about general and special controls here, but if we have a combination of general controls and special controls, that's Class II. And together, if we meet all three of these goals, that's what provides reasonable assurance of safety and effectiveness, and that's what ultimately allows FDA to grant the de novo request. So, your assessment of benefits and risks today is pretty crucial for that effort.

To provide a little bit more detail about our benefit risk thinking: our benefit risk assessment is based off of the totality of evidence in the de novo request, and we assess probable benefits and probable risks of the device. We also assess additional factors, including things like uncertainty: what's our inherent confidence that the data that we're reviewing in the de novo request is representative of the benefits and risks of the device that we will see in the real world? Patient perspectives: whether that's looking at patient preferences or patient reported outcomes, looking at values and endpoints that are important to patients. And then the unmet medical need: what other are other kinds of devices and products in this landscape, in this device space, and how does this device address an unmet medical need in this area?

As part of granting a de novo request, we create a new classification regulation, and that generally includes a number, a name, identification language, which describes what FDA believes to be a single device type with a shared intended use and shared technology between all of the devices that belong in that regulation. And FDA will design and write this regulation based off of the totality of evidence in our de novo requests and how we want to classify these devices for the future.

We develop a risk mitigation table, which is where we list what the risks to health are for the device on the left side there, and on the right side we have mitigation measures. These are categories of testing or other kinds of requirements which together mitigate their identified risk to health. For this example table here, this is a patient contacting device. So, there's might be risks like infection or adverse tissue reaction. And then for mitigation measures, we have things like reprocessing validation, labeling, and biocompatibility evaluation as well. And then the specific risks and mitigations for this device are going to be dependent on the intended use in technology. Generally, for in vitro diagnostics, that includes things like the risks of false positives, false negatives, misinterpretation of the results, and so on, and the effects to health that that has.

Next, we have special controls. Special controls are legal requirements for all devices in the regulation. They're written as part of the new classification regulation itself, and they can include, and are certainly not limited to, analytical validation requirements, clinical validation requirements, labeling requirements, and some post-market authorities as well. Very importantly, before we can grant a de novo request, the de novo device has to meet its own special controls.

FDA is developing a regulatory paradigm to help future devices continue to be safe and effective through the application of these special controls.

When a de novo request is granted, that device can then be legally marketed. We create the new classification regulation, and that device can now serve as a predicate for future, similar Class II devices to be cleared through the 510K process. We will publish a decision summary that provides transparency into our thinking and acts as a template for future devices to figure out what data requirements they need and how and what they need to be able to get clearance. And then, we will also publish our new regulation that we created for this device in the code of federal regulations, so that information is now available for others as well.

I just want to briefly go through an example here. This is the GSP Neonatal Creatine Kinase MM Kit. This is a de novo submitted back in 2018. When we granted this de novo request, we created a brand-new regulation here called the Muscular Dystrophy Newborn Screening Test, and I'll let you read the identification language there to see what the device is, but you'll notice that we have pretty high specificity for this device, meaning that other devices that are in this regulation will employ similar kinds of technology. Creatine kinase levels from dried blood spot specimens on filter paper from newborns as an aid in screening newborns from muscular dystrophy.

Here's just an excerpt of the special controls here. You can see that we have a very specific clinical validations special control, and I won't read through all the language here, but you'll see that there's very specific requirements about how this clinical validation study should be conducted, including how we include specific data requirements, how we assess the effective sample collection and processing steps. We require data to calculate certain reference intervals,

including sufficient samples to calculate the 97.5 and 99.5 percentile information. So, very specific data requirements for this for all devices in this regulation.

The reason why I gave training on the de novo program this morning is to help you understand the context of the questions that we are asking today. We're asking questions about the benefits and risks of the AvertD device. FDA will use your input from this meeting today in helping decide whether or not to grant the de novo request for this device. And if we grant the de novo request, that will establish a regulatory framework for ensuring the safety and effectiveness of similar devices yet to come, similar devices that will belong in the regulation that we create for this product, and that could include clinical validation, analytical validation, labeling, and post-market study requirements. We will design a lot of this with your input. Thanks very much.

Dr. Watson: Thank you very much, Dr. Mr. Yang. We will now proceed to the FDA's presentation on the Breakthrough Device Designation Program. I'd like to invite the FDA to begin. Dr. Rouabhi?

Dr. Rouabhi: Good morning. My name is Ouided Rouabhi, and I'm the Assistant Director for Policy and Operations Team 1 in the Office of Clinical Evidence and Analysis. Our office is located within the Office of Product Evaluation and Quality in CDRH. My team oversees several CDRH programs, including the Breakthrough Devices Program, which I'll be discussing today.

Our learning objectives for this session are to provide an overview of the Breakthrough Devices Program, to review the criteria for breakthrough device designation, and, lastly, to identify program features. At a high level, the intention of the Breakthrough Devices Program is to provide patients and healthcare providers with timely access to devices that provide for more

effective treatment or diagnosis of life threatening or irreversibly debilitating diseases or conditions. I'll talk more about what qualifies as a breakthrough device and the slides to come, but the key thing to keep in mind here is that for certain devices that meet the program eligibility criteria, FDA will expedite their development assessment and review.

The program guidance outlines the key principles of the program, which help to achieve that goal of more timely patient access. I won't go into each of these, but I wanted to highlight a few key points. The first is the opportunity for interactive and timely communication. Sponsors of designated breakthrough devices are able to receive feedback from FDA more quickly and collaboratively. This allows them to move forward with device development decisions more quickly while being reassured that the data they plan to collect is consistent with FDA's expectations. Similarly, the program also aims to prioritize the review of marketing applications for designated devices as well as aims to apply efficient and flexible approaches during review, such as the enhanced opportunity for post-market data collection, all the wild preserving the statutory standards for marketing authorization.

This last bullet is really an important one. So, the Breakthrough Devices Program does not change the statutory standards for marketing authorization. This means that sponsors of designated breakthrough devices are held to the same standards as similar devices that have not received the designation. Breakthrough device designation does not imply that the marketing application will be authorized.

There are a few key things to note. First, this is statutorily mandated program under Section 515B of the Food, Drug and Cosmetic Act, which was enacted following the passing of the 21st Century Cures Act at the end of 2016. The final guidance document describing the

program's implementation was issued in December of 2018. This is a very useful document that contains all of the policy and process information that I'll be sharing with you today, and the link is listed here on the slide. Lastly, this is a voluntary program, meaning that sponsors can choose to request the designation and they're not required to do so. Sponsors can request entrance into the Breakthrough Devices Program by submitting a breakthrough device designation request. If entrance is granted, then the sponsor will have additional mechanisms for feedback and interaction with FDA during the device development process as they're working towards their marketing submission. We'll talk about these more in the slides to come. Next, I'll talk about the criteria for breakthrough device designation.

There are a few general eligibility considerations that a device must meet in order to be designated breakthrough. First, the program is only open to medical devices and device led combination products. Second, devices seeking breakthrough device designation must be subject to future marketing authorization, be it the PMA de novo or 510K pathways. Devices that are exempt from these marketing pathways would not be eligible for the program. Lastly, the device should meet the specific criteria outlined in the statute, including fully meeting breakthrough device Criterion One and one of the subparts of breakthrough device Criterion Two.

Because the designation criteria are the basis for all of our grant or denied decisions, we do take a thoughtful approach in ensuring that each criterion is met. Beginning with Criterion One, the device in its proposed indication must provide for more effective treatment or diagnosis of life threatening or irreversibly debilitating human disease or conditions. A device must fully meet this criterion in order to be eligible for the program.

Because decisions on requests for designation are made prior to marketing authorization and a complete clinical data set is not required and may not be available at the time of designation, FDA considers whether the sponsor has demonstrated a reasonable expectation that the device could provide for more effective treatment for diagnosis of a disease or condition identified in the proposed indications for use. Reviewers take into consideration whether the sponsor has demonstrated a reasonable expectation of technical success, meaning that we reasonably expect that the device can be built and function as intended, as well as clinical success, meaning that a functioning device could more effectively treat or diagnose identified disease or condition mechanisms for demonstrating a reasonable expectation of technical and clinical success could include literature or preliminary data.

The statute specifically calls out life threatening or irreversibly debilitating human diseases or conditions. So, one thing that we consider is whether the patient population or subpopulation identified in the proposed indications for the device are representative of that statutory criteria. We generally interpret life-threatening as a disease or condition for which the likelihood of death is high unless the course of the disease is interrupted. In the case of irreversibly debilitating diseases or conditions, we consider the impact on such factors as survival, day to day functioning, and the likelihood of progression to a more serious disease or condition if left untreated.

In addition to meeting Criterion one, the second criterion requires that the device in its proposed indication should meet one of the following, subparts and Criterion Two. Either that the device represents a breakthrough technology, meaning that the device represents a novel technology or novel application of an existing technology that has the potential to lead to a

clinical improvement, or no approved or cleared alternatives exist, or that the device offers significant advantages over existing, approved or cleared alternatives, or lastly, that the availability of which is in the best interest of patients. I won't go into examples for each of these today, but our guidance document does talk about considerations for these subparts in more detail.

Next, I'll talk about the program features that are available to sponsors of designated breakthrough devices. Once a device enters the program, there are a few different features outlined in the guidance that are useful for facilitating interactions with FDA. These are just a few examples of features in the program that sponsors can choose to pursue. Some of them you may have previously heard about. First, the Data Development Plan is an optional map of the development process from entering to the program until the marketing submission and including post-market activities as necessary. The DDP is a high level document that summarizes the plan, non-clinical, and clinical testing, so that everyone is on the same page about data collection expectations. This can hopefully add predictability, efficiency, and transparency to the device development process in a way that's least burdensome. Another example of a feature open to devices within the program are Sprint Discussions. These are a highly interactive process to facilitate reaching rapid agreement on a single development issue. And lastly, regular status updates can be used in between submissions, which can be useful for planning purposes.

There are a few items that I wanted to note at the marketing submission stage. First, it's important to keep in mind that breakthrough device designation must be requested prior to FDA receiving the marketing submission. This means that, as we mentioned previously, a decision on the designation request is typically made while the device is still under development and

complete data has not yet been collected. During review of the marketing submission, the program principles and benefits are applied, including those who mentioned earlier, such as expedited interactions and priority review, as well as senior management engagement, and the opportunity for pre- and post-market balance of data, when appropriate. I want to note again that statutory standard for marketing does not change. As with any marketing submission, the review team reviews the totality of clinical and non-clinical evidence to make the regulatory decision.

In summary, the Breakthrough Devices program is intended to provide patients and healthcare providers with timely access to breakthrough devices. Devices are designated by meeting the statutory criteria, and designated devices can benefit from program features intended to expedite the development, assessment, and review of these devices, both during device development as well as throughout the regulatory submission process. And with that, I'll be happy to take any questions.

Dr. Watson: Thank you so much for those very informative FDA presentations. Panel members, I'd like to open it up. Does anyone have any brief clarifying questions for the FDA? If so, could you please unmute, identify yourself, ask a question? I see Dr. Goldstein has a question.

Dr. Goldstein: Thank you. This is Lawrence Goldstein. My question is: are the data standards for quality the same as those which would be used in drug development?

Dr. Watson: Dr. Rouabhi?

Mr. Mr. Yang: Excuse me, this is Peter Mr. Yang. I'm not sure she's on the line right now. I will say, just from my perspective, the regulatory standard for review is a little bit different than what it is for drugs. So, I don't actually know the drug standard well enough to do a direct

comparison between the two. I think I would consider it simply from your clinical experience and your perspective. and I think FDA can sort of work through the differences between regulatory standards, if any, that apply here today.

Dr. Goldstein: Thank you.

Dr. Watson: Dr. Farrar?

Dr. Farrar: Yes. With regards to the new device standard that you're talking about, has the FDA approved chip technology or a combined genetic testing of some sort previous to this particular product? And if so, why doesn't this one fit into that category?

Mr. Mr. Yang: Sure. I can't speak to whether similar technologies have been approved. I don't have a good memory for the thousands of devices that FDA reviews every year. I will just say, from the perspective of how we consider something to be new, we will consider it based on a combination of the intended use and the technology. So even if the technology has been used elsewhere for different application, because of the risks of this particular device in terms of what it's purporting to provide information about and the consequences of false positives, false negatives, and that sort of thing, we will consider the combination of the intended use in technology. And so, because it's being used in a different application space, that might be one of the reasons why it's different from other devices.

Dr. Farrar: And if I might just follow up on that with regards to genetic testing, are there specifications? I'm afraid I didn't have chance to try and find this for sensitivity and specificity for genetic testing in general that we might use in judging the results of the studies done for this product.

Mr. Yang: I believe Dr. Kelm might be able to answer that more specifically. I'm not aware that we have sort of an agency-wide or a center-wide policy for the sensitivity and specificity requirements. I think we would generally consider those on a case-by-case basis because, when we factor in our benefit risk analysis, we'll usually factor in things like the un the inherent uncertainty in the data that we do have, but also a combination of things like the unmet medical need. So, we might be willing to accept a different benefit risk profile if there is an unmet medical need in the device space. It might cause us to say, “Hey, we think this device is worth it to be on the market.” Setting a sort of center-wide policy for the metrics and such can limit us from doing that in some cases. But I'll let Dr. Kelm speak, if she has additional comments.

Dr. Kelm: Sure. Good morning. This is Kellie Kelm. We do have information from the firm on their analytical accuracy and other assessments. Our division has cleared many genotyping assays and other types of genetic assays. We don’t necessarily have a requirement, or a standard if you will. Again, most genetic tests tend to analytically perform at a very high level of accuracy and precision. We are not going to be talking about that today at this meeting. We're separately working with the sponsor on that, but we have not necessarily any concerns to bring that to you today.

So, the question about sensitivity and specificity though, in terms of this device for this intended use, that obviously is a question that we love to get input from on the Panel. It is for the assessment of risk for developing OUD, and obviously, that's not just how the genetic test performs, but there's a lot there in terms of, also, the study, and more. So, that is obviously the

purpose of today's study, is to get your assessment of all those things and provide us some feedback on that.

Dr. Watson: Thank you all. Any further questions? All right. Seeing none.

Dr. Goldstein: Well, now, there was one. There's one by Laura Bierut.

Dr. Watson: Sorry. I'm sorry. Go ahead.

Dr. Bierut: Hi, my name is Laura Bierut and I have a general question which I think we will get to later today, which is just: however the FDA is kind of thinking of genetic tests that use multiple variants and developing it to give a probability of risk — the guidelines, how that's been applied in other diseases — just will be helpful to have me think about the risks specifically for OUD. So, I don't need an answer now, but just that's something that I'm sure that the FDA has been doing already in understanding what the risks and benefits are in general will be helpful.

Dr. Watson: Peter Mr. Yang or Kelly Kim, would you like to answer or respond to that? Okay. All right, then I guess we should be moving on. We can proceed to the sponsor's presentation. I'd like to invite the sponsor to begin. I remind the public observers at this meeting that, while the meeting is open for public observation, public attendees may not participate except at the specific request of the Panel Chair. The sponsor will have 60 minutes to present. Sponsor, please begin your presentation.

Dr. Donaldson: Good morning. My name is Keri Donaldson. I'm the CEO of SOLVD Health and a board certified clinical pathologist specializing in pharmacoepidemiology and genetics. I'm pleased to present the AvertD genetic risk assessment test and the data to support its

utility. AvertD is designed to be used in conjunction with clinical evaluation to facilitate informed decision making regarding the prescription of oral opioids to treat acute pain.

We are here today because the opioid epidemic is a public health emergency showing no signs of slowing down. We know addiction is a chronic disease best addressed with prevention and treatment, rather than judgment and punishment. However, despite efforts to reduce prescription opioid use, people are still becoming addicted, indicating that additional measures are needed to develop safer prescribing practices, including the use of more effective risk assessments. Although risk assessments prior to prescribing opioids are the cornerstone of clinical practice today, the current tools have many limitations, and there are no FDA approved tools to assess genetic risk of developing Opioid Use Disorder, or OUD.

AvertD is an innovative genetic risk assessment test designed to fill the gap in current risk assessment tools. AvertD uses genetic polymorphisms involved in the brain reward pathways to detect, identify, and analyze genetic risk of developing Opioid Use Disorder after taking prescription oral opioids for acute pain. It was specifically designed and trained for determining the genetic component of risk of Opioid Use Disorder by comparing single nucleotide polymorphisms, commonly referred to as SNPs, known to be associated with differences in the brain reward pathways among individuals who develop OUD versus individuals who did not develop OUD.

Importantly, the machine learning algorithm, using these SNPs, followed the best practices for machine learning test development, and was specifically trained to classify individuals with OUD from individuals without OUD, not to classify individuals with or without other substance use disorders or other mental health comorbidities. In short, a verdict can help

assess a person's genetic risk of developing OUD and thereby enable clinicians and patients to make a more informed decision making on prescribing oral opioids for acute pain.

AvertD is designed for use with a buckle sample collection kit. The collection kits consist of two single use sterile flock swabs, each with two vials containing 550 microliters of DNA stabilizing solution. The samples are stored at room temperature and then transported for analysis. AvertD follows a common workflow for molecular PCR-based assay and includes the amplification mix, the reagent module, the micro array, and the proprietary software to analyze the results.

AvertD testing is simple. After the buckle sample is collected, it is mailed to a high complexity CLIA-certified laboratory. Next, DNA is extracted from the buckle sample using standard laboratory methods and then undergoes amplification. The DNA extraction and PCR amplification are performed manually via standard laboratory equipment. Following amplification, the sample is loaded onto a sample plate. The remaining processes are automated on the instrument. The reagent contains a fluorescent nucleotide that is incorporated into the primary extended PCR product during the extension step of the reaction. Afterwards, the hybridization buffer is added to the sample, which is then automatically applied to the microarray. The microarray chip has spots for the assay-specific capture probes. The chip is automatically washed to remove unbound material and then dried. The microarray is scanned by the analyzer, and a specific algorithm is used to determine the genotype for each gene, and then a risk classification for OUD.

The AvertD results report indicate whether the patient has a high or low genetic risk for Opioid Use Disorder. The results, along with the clinical evaluation, should be discussed by a

clinician and patient to inform an individually tailored plan for pain management. The data that we will present supports our proposed indication that AvertD be authorized as a genotyping test used to detect and identify 15 clinically relevant SNPs. As I have outlined, these SNPs are involved in the brain reward pathways and associated with Opioid Use Disorder and can help identify patients who may be at increased risk for OUD. Information from AvertD provides patients 18 years of age or older and their healthcare providers with objective information to be used for shared, informed decision making prior to taking oral opioids for acute pain. Again, this information should be used in combination with a clinical evaluation.

For context, next, I'd like to present a high level summary of the interactions we've had with the FDA. In early 2018, FDA granted AvertD breakthrough device designation, and the clinical study was conducted in 2019. In 2020, we submitted a de novo request to classify AvertD as a Class II medical device with proposed special controls. This was based upon well-accepted methods common to many genotyping tests and modeled after several Class II genetic risk assessment tests, as well as other in vitro diagnostic tests. The de novo classification provides a regulatory pathway for market authorization and classification for novel, low to moderate risk medical devices for which general controls or general and special controls provide a reasonable assurance of safety and effectiveness but for which there is no legally marketed predicate device. Our request was declined in 2021. However, we continued to collaborate with the FDA. This included providing additional data and analyses to address remaining questions regarding the uncertainty in the study population and the applicability of the results the intended use population. We resubmitted the de novo request in June of 2022.

The clinical study, including the additional data and analyses, demonstrate that AvertD is an effective risk assessment test to help identify patients at increased genetic risk of developing OUD. AvertD achieves this pre-specified performance goals for specificity and sensitivity and is 80% accurate in identifying this genetic risk. Patients identified as high genetic risk have 18 times the odds of developing OUD after taking an oral opioid for between four and 30 days, compared to those identified as low genetic risk. These data, using conjunction with a complete clinical evaluation, will enhance shared, informed decision making regarding the use of prescription oral opioids for acute pain.

With that background in mind, let me review the rest of our presentation. Dr. Joseph Garbely will present key epidemiologic data on opioid use and OUD, along with current practice guidelines and the tools currently used in assessing risk, highlighting the unmet meet for a test like AvertD. I'll present the efficacy data from our study, and then Dr. Christine Brower will present the additional analysis performed to address the FDA's questions. Finally, Dr. Chris Zacko will give his clinical perspective on the data and AvertD utility. Dr. Brower will then return to moderate the question and answer portion of our allocated time. To ensure that we can answer all your questions, we have also invited Chris Mullen, who is a statistician with NAMHSA. Thank you, and I'll now turn the presentation over to Dr. Garbely.

Dr. Garbely: Thank you and good morning. It's an honor and privilege to be here. My name is Joe Garbely, and I'm a Distinguished Fellow of the American Society of Addiction Medicine and a fellow of the American Psychiatric Association. I'm also a faculty member at Penn State and Drexel University. I've served on the American Society of Addiction Medicine Board of Directors since 2018. I'm also the National Director of the ASAM Physician in

Training Committee, inculcating addiction medicine training at medical schools and primary care residencies across the US. We have created core competencies for both medical schools and primary care residencies, as well as core standards in addiction medicine fellowships. I've spent a great portion of my career treating substance use disorders, including developing specialized treatment and monitoring programs. Having treated many patients with OUD and having prescribed opioids to many patients as an internist, I understand firsthand the severity of the opioid epidemic, and so I'm here today unpaid. This is a testament to my commitment to help patients and be part of a solution.

Opioid use disorder, or OUD, is a condition characterized by a desire to obtain and take opioids despite social and professional consequences, an overpowering desire to use opioids, increased opioid tolerance, and withdrawal syndrome when opioids are discontinued. It is a chronic brain disease. The prevalence of OUD in the US general population is approximately 1%. It is important to understand that the opioid epidemic began with, and continues to be fueled by, prescription oral opioids. They also play a large role in drug overdose mortality. In fact, one commonly cited study noted that approximately 80% of heroin users reported that they began with prescription opioids. According to SAMHSA's annual survey in 2020, 13.4 million people self-reported that they had personally misused prescription opioids during the previous year. Overdose deaths involving prescription opioids increased nearly five times from 1999 to 2020. Also, during this time period, drug overdose became the leading cause of death for Americans under 50. I see these trends every day in patients. I. Their addiction begins with non-illicit prescription, oral opioids, and quickly escalates the use of illicit drugs like fentanyl.

Current approaches to prevent opioid addiction include following prescribing guidelines and using educational materials. Prescribing guidelines such as the CDC recommendations recommend using non-opioid alternatives, when appropriate, to treat acute pain. When opioids are needed for appropriate pain management, the aim is to reduce the duration and dosage used. Importantly, as mentioned by Dr. Donaldson, assessing addiction risk before prescribing opioids has become routine in clinical practice. Physicians have multiple strategies for assessing risk, including patient interviews, reviewing past medical records, and deploying risk questionnaires. However, it's widely accepted that current risk assessments have limitations.

The CDC has stated that currently available risk stratification tools are not sufficiently accurate for classifying patients at low or high risk for abuse or misuse. Similarly, a recent systemic review said these tools are of limited value despite being widely used. None of the current risk assessment tools assess genetic risks for developing OUD.

And genes play an important role in the risk of addiction. The genes that people are born with account for approximately 50% of a person's risk for addiction. Gender, ethnicity, and the presence of other mental disorders may also influence risk for drug use and addiction. A person's environment includes many different influences, from family and friends to economic status and general quality of life. In addition, genetic and environmental factors interact with critical developmental stages in a person's life to affect addiction risk.

Focusing now on genetics, numerous genes associated with OUD risk have been identified across multiple study types, including both candidate gene and genome-wide association, or GWAS, approaches. These include the opioid receptors, Mu and Kappa, the delta opioid receptor, or OPRD1, as well as genes seen across substance use disorders, such as the

dopamine receptors. Genetic predisposition for Opioid Use Disorder occurs in the mesolimbic system. It is there that one experiences feelings of wellbeing triggered by dopamine from natural rewards and drug rewards. Chemical messages release dopamine in this area, and genes that control these messages and subsequent release are referred to as the brain reward pathways. They begin at the ventral tegmental area and end in the nucleus accumbens, a dopamine enriched pathway. It is well established that genetic mutations affecting the brain reward pathways can result in substance seeking behavior, eventually resulting in a fully developed substance use disorder.

To summarize my presentation, since the opioid epidemic began with and continues to be fueled by prescription oral opioids, it is vital to sufficiently assess for risks prior to prescribing oral opiates. More specifically, we need to better understand a patient's genetic risk for developing OUD, because, as I mentioned, genetics contribute significantly to an individual's risk of developing Opioid Use Disorder. And with current approaches, we are limited, at least partly, because they do not take into consideration the role of genetics. We need an accurate genetic risk assessment test to provide a more complete understanding of an individual's risk and to better inform decision making around the safe use of opioids for acute pain management. Thank you. I'll now turn the presentation back to Dr. Donaldson.

Dr. Donaldson: Thank you, Dr. Garbely. I'll now review the study design and efficacy data. We ran a blinded, multicenter study designed to evaluate the ability of AvertD to differentiate those at high genetic risk for developing OUD from those at low genetic risk. This was an all-comer study, and enrollment included all participants who met the inclusion criteria. Unlike many clinical studies that have active recruitment outreach, in this study, patients at all

clinical sites were asked, as part of their regular scheduled office visit, if they wanted to participate in this study. No additional recruitment was performed. Participants were approached during their normal clinical care at ten geographically diverse private practice sites in the US. We specifically selected private practice sites because the majority of oral opioids are prescribed outside of academic institutions. There were six general practice sites, one research only site, and three sites that provided medication-assisted OUD treatment. Given the prevalence of OUD is 1%, sites providing OUD treatment were included to increase the likelihood of enrolling enough OUD-positive participants.

This study was a prospective study with one retrospective data element. Opioid use disorder may take a long time to manifest. Therefore, a retrospective design element was needed to complete the study in a timely manner. The retrospective element was the participant's index exposure to prescription oral opioid, specifically site personnel interviewed adults 18 years of age or older to determine if they had been prescribed and taken oral opioids for a minimum of one year prior to the interview. The design allowed for an adequate amount of time for OUD to develop and shortened the total amount of time to conduct the study by not requiring many years of observation time between the exposure and the assessment of OUD. In addition, the index, or initial exposure to prescription oral opioids, was required to be between 4 and 30 consecutive days. This duration has been shown to proceed persistent opioid use and is consistent with clinical prescribing patterns in the US for acute pain. We chose to use self-reported index exposure, as studies have shown that many patients prescribed oral opioids may not fill their prescription, or fill them, but may not take them. Furthermore, it was determined that self-reported exposure would be more valid and accurate than solely relying on medical records.

If participants met the enrollment criteria and provided informed consent, they were enrolled and clinically evaluated for OUD, and a buckle sample was collected for AvertD. Clinical evaluation involved assessment for OUD using the Diagnostic and Statistics Manual of Mental Disorders, Fifth Edition, or DSM-5, criteria. Participants for whom a DSM-5 OUD diagnosis was established were assigned an outcome of OUD positive for the study. These results were used as a valid clinical comparator to establish the performance of a verdict.

The study had two co-primary endpoints, chosen to ensure adequate performance. Sensitivity is defined as a proportion of participants with OUD who are correctly identified by AvertD as high genetic risk. Specificity is defined as the proportion of participants without OUD who are correctly identified by AvertD as low genetic risk. This information informs the decision making process, helping healthcare providers and patients evaluate whether the benefits of short term oral opioid pain relief outweigh the risk of harmful side effects.

The study included likelihood ratios as the secondary endpoint. In evidence-based medicine, likelihood ratios are used to assess the value of performing a test, particularly a risk assessment test. A pre-specified random representative sampling using strata was employed to ensure the study population mirrored the intended use population of adults in the US who are prescribed oral opioids for acute pain. This process minimizes bias and the impact of unmeasured confounders while ensuring a sufficient number of OUD positive participants.

Post enrollment, a blinded independent statistician randomly selected participants who met predefined categories for sex, age, time from opioid index exposure to enrollment and likelihood of OUD. This was defined as a presence or absence of any substance use disorder, which was determined by SOLVD Chief Medical Officer. This allowed the random sampling of the strata

while ensuring that the study population had enough OUD positive participants. The statistician was blinded to the test results as well as OUD status during this process, and the Chief Medical Officer was blinded to the AvertD test result.

A total of 812 participants were enrolled. The blinded statistician reviewed the demographic composition of the enrollee and determined that 689 participants were sufficient to meet the stratification criteria. Stratification sampling plan was followed, and 385 participants were selected at random to fill the strata to ensure we met the pre-specified power, and the study population represented the US population that takes prescription oral opioids, and to ensure adequate representation of patients with OUD. Shown here are the study population demographics. 58% of the participants were male. Majority of participants were white and of non-Hispanic ethnicity, and 45% were OUD positive per clinical evaluation with DSM-5. Displayed here are the results of the random representative sampling via stratification. The sex and age stratification is consistent with patients who are prescribed oral opioids in the US. There were more individuals whose time from oral index exposure was at least four years prior to enrollment. The accuracy of this information was established at enrollment and verified by SOLVD at the request of the agency. Overall, the random representative sampling achieved the goal of allowing appropriate representation of OUD positive and negative patients across all strata.

Next, I'll turn to our primary efficacy results and subgroup analysis. AvertD met the pre-specified efficacy endpoints for performance goals. The lower bound of the 95% confidence interval for sensitivity is greater than the performance goal of 59.5%, and the lower bound for specificity is greater than the performance goal of 55.5%. Therefore, both co-primary endpoints

were successfully met. Fisher's exact test of the difference in the observed proportion and the performance goal proportion for both sensitivity and specificity resulted in a significant P-value. Four participants did not have an AvertD test result. A sensitivity analysis was conducted to evaluate the worst case imputation. In this analysis, one participant was an OUD positive case that was imputed with a low risk test result, assuming a false negative test, and three participants were non-OUD cases imputed as high risk test results. Even in this worst case scenario, the results yield a lower confidence interval, well above the corresponding performance goal, as demonstrated by the statistically significant P values.

Furthermore, subgroup analysis demonstrates the robustness of the test displayed. Here are the sensitivity data based on sex and age, all of which show consistent results and no evidence of a difference by subgroup. Displayed here are the sensitivity data for time from opioid index exposure to enrollment, race, and ethnicity. Again, we see highly consistent results across all subgroups as compared to the overall population. For completeness, we have included the subgroup analysis for specificity, and again, we see consistent results. And finally, here are the specificity data for time from index exposure, race, and ethnicity. Again, we see highly consistent results across all subgroups as compared to the overall population. All participants were at least 18 years old at the time of enrollment. However, their age at the time of initial exposure could have been less than 18. Subgroup analyses were also performed for participants younger than 18 and those 18 or older at the time of index exposure. Sensitivity and specificity in each age subgroup were consistent with the overall population and exceeded the performance goal.

The FDA requested we also analyze sensitivity by severity of OUD. It's important to note that the prevalence of OUD by severity in the intended use population of patients taking prescription opioids for acute pain is not well-characterized in the literature. Severity of cases on this slide and included in the study are indicative of the intended population for this test. We compared mild to the moderate and severe groups, and the mild and moderate groups to severe, and found no statistically significant differences.

Turning to our secondary endpoint results, two co-secondary endpoints were the positive and negative likelihood ratios. These are common measures of diagnostic test performance. For the positive likelihood ratio, higher values, in particular, values greater than one, indicate an increased probability of having the disease given a positive result. For the negative likelihood ratio, lower values, in particular, values less than one, indicate a decreased probability of having the disease given a negative result. We see that both the positive and negative likelihood ratios support the primary results. Positive likelihood ratio of 3.98 supports a strong increase in the post-test probability of having OUD, with an AvertD result indicating high genetic risk. And the negative likelihood ratio, 0.22, showed a strong decrease in the post-test probability of having OUD, with an inverted result indicating low genetic risk. Subgroup analyses for these ratios were consistent with overall results.

The diagnostic odds ratio is a measure of overall diagnostic performance calculated from the ratio of positive and negative likelihood ratios. It is also related to both sensitivity and specificity and is a simple way to combine diagnostic metrics into one number to communicate performance independent of prevalence. It is technically defined as a ratio of the odds of the test being positive if the subject has the disease to the odds of the test being positive if the subject

does not have the disease. For AvertD, the diagnostic odds ratio is 18.1. In other words, a patient with OUD has 18 times the odds of receiving an AvertD test indicating high genetic risk compared to a patient without OUD receiving an AvertD test indicating low genetic risk. These data demonstrate the clinical value of the test results to both patients and physicians. This concludes the overview of our key data.

Beyond the data, we realize it's important to help physicians and patients understand how to interpret test results and how genetics can be integrated into the shared decision making around oral opioid prescribing. As has been described, genetics are important to understanding the risk of developing OUD. However, they do not tell the whole picture. To avoid over-reliance on the result and to ensure proper interpretation of results, labeling and educational materials will clearly communicate that AvertD has limitations. Our materials will also make it clear that AvertD is a genetic risk assessment test for use in informing decision making, not a test to diagnose OUD. It is also important to note that AvertD will require prescription and will be administered by a healthcare professional to ensure both patients and physicians receive the additional materials on how to interpret the results and to facilitate shared, informed decision making around opioid prescribing.

Our materials will emphasize that it is extremely important that test results are used in conjunction with a complete clinical evaluation to determine their appropriateness of opioids in pain management for an individual patient. They will also emphasize the importance of following opioid prescribing guidelines, even for patients with AvertD results indicating low genetic risk. In addition, AvertD will be purposely launched using Centers of Excellence. Tests will be performed in centralized CLIA-certified laboratories. We will use feedback from the

Centers of Excellence to further refine the educational process and materials. Only then will we expand to additional Centers of Excellence.

To summarize, the data we collected and analyzed support the use of AvertD test results to assist in the detection of individuals who may be at higher genetic risk for developing OUD. The co-primary endpoint and subgroup analyses demonstrate that sensitivity and specificity results are not impacted by age, sex, time from index exposure, race, or ethnicity, and the diagnostic likelihood ratios support the primary endpoint results. Furthermore, these results support the use of AvertD as a valuable test to help identify patients at high genetic risk for developing OUD, an overall prevention strategy to help address the opioid epidemic. Thank you, and I'll now turn the presentation over to Dr. Brauer.

Dr. Brauer: Thank you. Good morning. My name is Christine Brower. I'm a regulatory affairs consultant, and I've been working with the SOLVD Health Team for the last three years on AvertD. I'll now review the additional analyses we perform to address the questions raised by FDA.

The FDA questions primarily pertain to three issues: one, the potential impact of using multiple versions of the case report forms; two, uncertainty around the use of self-reporting to capture oral opioid exposure; and three, uncertainty around applicability of results from the study population to the intended use population, focusing on any difference in the study site results and prevalence and impact of any mental health comorbidity.

First, let's review the multiple case report forms used in the clinical study and how we addressed any potential impact. Prior to enrolling any participants, sites were trained using the study protocol, which included the inclusion and exclusion criteria. The inclusion exclusion

criteria from the study protocol, not the case report forms, were used to enroll participants. In a few moments, the FDA will present the details of the changes to the CRS. It is common in clinical studies for case report forms to change. Most importantly, the changes had no impact on enrollment or study outcomes.

By way of two brief examples, one change was based on an FDA request to document where the participants came from. So, we added the state of residents to the form. One other change was adding the minimum and maximum days of oral opioid exposure to the case report form. We went back and reviewed the data and found there was no impact on study results based on this change, as all participants' exposure to oral opioids felt within the range specified in the protocol. To address case report form versions used during the study, and to ensure data collection consistency after study completion, study sites documented and ensured all eligibility criteria f or each participant. They did this using a single, new case report for. Using the new case report form data and instruction specified, all 385 participants were again confirmed to meet the inclusion/exclusion criteria. In conclusion, changes to the case report forms used had no impact on enrollment or study outcomes.

Now let's look at what we did to confirm the accuracy of the self-reported index exposure used in the clinical study. As a reminder, we chose to use self-reporting because published literature has shown that prescription records for oral opioids used to treat acute pain are often inaccurate, as many patients don't fill their prescriptions, or they may fill them, but they don't actually take the. To provide cooperative documentation of the index exposure, SOLVD collected new information about self-reported index exposure. Study site personnel examined medical records one year before and after the self-reported index exposure. Participants were

then assigned to tiers based on the extent of information available to corroborate the self-reported index exposure. One example of such information would be a documented procedure, like a surgery or a dental procedure, or an accident that may result in oral opioid prescriptions. Sites also looked for documentation or reference to an oral opioid prescription being made, as well as the presence of a physical prescription itself. These tiers were defined in collaboration with the FDA. The purpose, again, was to evaluate the level of collaborating evidence of opioid exposure to address concerns around the self reported index exposure.

The tiers are not mutually exclusive and were categorized as follow. Tier 1 consisted of all participants who met the inclusion and exclusion criteria. Tier 2 included participants where medical records, including documentation of a procedure such as I have described, where oral opioids may be prescribed for acute pain. Tier 3 included those participants who had documentation that an oral opioid prescription for acute pain was issued within a calendar year of the self-reported index exposure. For example, the medical records documented that the patient was prescribed seven days of hydrocodone for knee surgery. Lastly, Tier 4 included those in Tier 3 whose records included the actual prescription, either a physical copy, electronic copy scan, or photograph of the actual prescription.

It's important to note it is uncommon for sites that did not prescribe the oral opioids to have physical copies of the prescriptions from other prescribers in their records. Therefore, certain sites in the study would not have this type of information available for the index exposure. Also, most of our participants had their index exposure prior to 2015, when it was less common for medical records to have physical copies of actual opioid prescriptions. Furthermore, it was expected that fewer participants would be included in Tie 4, which is what we observe.

We did, however, anticipate that the majority of participants would have documentation or notation in the medical records of opioid prescriptions or a procedure or event. In fact, that is what we found. 95% of participants had a procedure or an event, and 83% had notation of a prescription. These data corroborate the participants' self-reported index exposure.

Now let's look at the results. All of these analyses support the sensitivity and specificity of AvertD. Tier 3 includes all participants who met the enrollment criteria and where there was documentation in the medical records that an opioid was indeed prescribed to the participant. Tier 3, which includes 83% of the participants, also meets the performance goals as specified in the addended statistical analysis. Tier 4 had the smallest sample size for reasons previously described. Again, at the time of the opioid exposure, it was not common for medical records to have physical copies of actual prescriptions, nor was it common for sites that did not prescribe the oral opioids to have physical copies of prescriptions from other prescribers. These issues make it harder to derive conclusions about Tier 4. Overall, these data confirm the accuracy of the self-reported index exposure. They resolve uncertainty relating to the self-reported index exposure for qualifying study participants, and the applicability of the study results to the intended use population.

At the request of the FDA, we also analyzed data by site specialization. There were three sites that provided treatment for OUD. This was defined as healthcare providers prescribing medical assisted therapy medication with SAMHSA Drug Addiction Treatment Act of 2000 waiver certifications for providing such therapy during the study enrollment period. Two of the sites were mental healthcare practices that provided OUD treatment, and the third was a general practice site that provided OUD treatment. These three sites were grouped together in sub-

analyses, as all three offered OUD treatment. The remaining seven sites were general practice sites that participated in research studies, but no healthcare providers at these sites provided OUD treatment. Of note, most of the OUD positive participants were recruited at the sites where OUD treatment was available.

To evaluate whether test performance differed by type of study site, SOLVD compared AvertD performance in participants enrolled from sites that provided treatment for OUD versus sites that did not. As you can see, there were no statistically significant differences in sensitivity or specificity between OUD-specialized and non-specialized sites, providing confidence in the study results and the applicability to the intended use patient population.

FDA also requested that we review the data to evaluate whether the study population was enriched for mental health comorbidities and non-opioid substance use disorders at the time of index exposure, and also to examine potential confounding factors, and to better understand the applicability to the intended use population. Study sites collected data on mental health and non-opioid substance use disorder comorbidities documented in the medical record. Shown here is the prevalence of mental health and non-opioid substance use comorbidities seen in the study population at the time of the index exposure. The most common comorbidities at the time of the index exposure were depression, anxiety, and alcohol use disorder. Importantly, the prevalence of the comorbidities between OUD negative and OUD positive participants were generally similar at the time of index exposure. For less common mental health comorbidities, such as bipolar disorder, the sample sizes are small, and the total number of affected participants is low.

We also looked at various literature sources to determine if the rates seen in the study were consistent with prevalence rates in the United States. Our literature review included reports

from CDC, NIH, the American Psychological Association, SAMHSA, and Harvard University. The data reviewed support that the study population is not enriched for mental health disorders at the time of index exposure and, in fact, are consistent with prevalence rates in the United States. Looking at the data, regardless of the presence or absence of a given disorder, the sensitivity results showed no statistically significant differences. Likewise, specificity data showed no statistically significant differences by mental health condition. These data further remove uncertainty about test performance in the intended use population.

To explore if AvertD is classifying other mental health conditions or other SUDs, not OUD, we also examine the sensitivity and specificity of AvertD in OUD negative subjects. We selected three condition alcohol use disorder, depression, and anxiety, because they are the more common mental health comorbidities. These analyses removed OUD positive participants to avoid this confounding factor. The sensitivity and specificity of AvertD for these other comorbidities is essentially the same as the prevalence of these comorbidities in the underlying study population. These analyses demonstrate that AvertD performance seen in the study is for OUD classification, not these comorbidities.

In conclusion, the clinical study was designed to ensure that the study population matched the intended use population. This allowed for interpretation of the study results and demonstrates the applicability of the study results to the intended use population. The consistency between participants' self-reported oral opioid exposure and the corroborating documentation in the medical record over time further establishes the use of self-reported index exposure within the study design. There were no statistically significant differences in sensitivity or specificity between sites that did and did not offer OUD treatment. This provides confidence

in the study results and the applicability to the intended use population. Through our additional analyses, we confirmed that the AvertD study population was only enriched for OUD positive cases, not for comorbidity. Even in the presence or absence of mental health and non-op Opioid Use Disorder comorbidities, the test performance was consistent. Next, I'd like to turn the presentation over to Dr. Zacko for his clinical perspective.

Dr. Zacko: Thank you, Dr. Brower. My name is Chris Zacko. As a spine surgeon, I regularly have to assess the risks and benefits of surgical intervention, which includes postoperative pain management. I am a member of the Enhanced Recovery After Surgery, or ERAS Society, and my practice was the first to be highlighted on the ERAS Spine website. One of the core tenets of ERAS protocols is to minimize opioid use to enhance surgical outcomes. I'm also Professor of the Department of Neurosurgery and member of the Opioid Stewardship Committee, Vice Chair of Quality for neurosurgery, and former Surgical Director of Perioperative Medicine, all at Penn State. Like Dr. Garbely, I'm also not being paid for my time here Today. I'm happy to be here on my own accord to discuss Opioid Use Disorder and how I believe AvertD is an important step in providing physicians and patients with better risk assessment tools to inform safer opioid prescribing practices.

As Dr. Garbely mentioned, many patients with OUD start with a prescription for oral opioids. There have been many earnest attempts to reduce opioid addiction and elicit opioid use through implementing prescribing guidelines, limiting duration and dosage of exposure, and assessing risk. Most opioid prescribing guidelines, including those from the AMA CDC double AANS, and ERAS include risk assessment prior to prescribing opioids to determine if the patient has a high risk of developing addiction. But we know that existing risk mitigation strategies are

not working. People are still getting addicted to prescription opioids. Risk assessments using a clinical evaluation, including a patient's medical history or the Prescription Drug Monitoring Program identify patients retrospectively who may be addicted or have a history of addiction, but they do not prospectively evaluate a patient's predisposition to addiction, and risk questionnaires designed to assess the risk of future addiction are not sufficient on their own. Many are subjective and they do not account for genetic risk, which is a significant component of OUD.

The genes that people are born with can account for about half of a person's risk for addiction, so current risk assessment tools may be missing up to 50% of the equation. With AvertD, we be able to factor in the genetic component with confidence that those who are identified as high genetic risk have 18 times the odds of becoming opioid dependent compared to those identified as low risk. Having this information would be a significant addition to the tools clinicians have to understand an individual patient's risk and to ensure their pain treatment is as effective and safe as possible. Let's take a look at the potential risk of introducing a genetic risk assessment test like AvertD.

As I see it, there are basically three risks. There's the risk of a false negative, the risk of a false positive, and the potential for physicians to overly rely on test results, misinterpret test results, or to incorrectly act on test results. All of these can be mitigated through proper use, which will be reinforced through labeling and education. The risk of a false negative is mitigated through proper use in conjunction with clinical evaluation. Clinicians would still be following the current standard of care for prescribing oral opioids. The labeling an education program will emphasize the importance of following opioid prescribing guidelines, even for negative tests. A false positive could potentially mean that a patient at low risk of developing OUD is prescribed

an analgesic alternative to an opioid, of which there are many. Since physicians who treat pain continually reassess pain management for effectiveness, the negative risks of this outcome are also minimal. If a patient's pain was not adequately treated with non-opioid alternatives, the decision to avoid using opioids could be reassessed. Lastly, with any risk assessment tool, there's always the risk for overreliance on the results, or misinterpretation, or even incorrect action based on the results. These risks will be mitigated through labeling and educational materials that will describe in detail how test results should be interpreted and used.

In my daily surgical practice, I'm constantly making individual assessments of the risks and benefits of surgical intervention for my patients. This is the very foundation of the surgeon-patient relationship. The opioid epidemic clearly demonstrates it is impossible to advise an overall risk of a procedure without considering the effect of medication used for postoperative pain control. And, as you can imagine, most of my patients are now specifically asking about their own risk of opioid uses in the preoperative period. As highlighted earlier in the presentation, in the present day, my ability to assess that risk has limits. I always look to minimize opioid use, as it has been clearly shown that excessive opioid use leads to worse outcomes and tremendous morbidity in surgical patients. However, in certain cases, opioids may be required to achieve sufficient pain management. In these cases, I need a way to identify patients at high risk of developing OUD, including their genetics.

As a surgeon, I'm obligated to use every resource at my disposal to ensure favorable surgical outcomes. For example, data points derived from genetic risk information are increasingly utilized in clinical medicine to present comprehensive and thoughtful precision

medicine solutions to patients. Oncologists have BRCA1 and BRCA2; neurologists have a APOe4.

AvertD would enhance my ability to assess a patient's risk of developing OUD by providing me with information on genetic risk that is currently unavailable. It will be a critical resource in informed decision making regarding the use of opioids for perioperative pain management. The results of AvertD would be an important piece of the puzzle to optimize a multifaceted and precision medicine approach specific to the patient sitting across from me looking for guidance. If I determined a patient to be low risk for developing OUD, I'd follow current standard of care guidelines for acute pain. On the other hand, if a patient was determined to be at high risk, then I'd develop a personalized pain management plan to minimize or completely avoid the use of opioids. AvertD would give me the ability to have a more informed conversation with my patients about their pain management plan. Thank you, and I'll turn the presentation over to Dr. Brower to answer your questions.

Dr. Watson: All right. Is that the conclusion of the sponsor's presentation?

Dr. Brauer: Yes, it is. Good morning.

Dr. Watson: Okay, great. We have a lot of questions, so I'd like to open up the question and answer session to the panel members who have any questions to pose to the sponsor. Can we start with Jennifer Higgins?

Dr. Higgins: Thank you very much. I have a couple of comments and then a couple of questions, and my comments pertain to Dr. Donaldson's slide CO-49, 'Results are not impacted by age or race.' With a sample size of 385, 210 being negative and 175 being OUD positive, this seems low to me. I'm also a little bit concerned about the fact that there were only 47 over the

age of 65, and that seems low, given that we know there's increased sensitivity to OUD amongst older adults with complex health conditions and potential for polypharmacy rate of absorption, et cetera. Also, non-whites were significantly underrepresented, and those were some thoughts I had about the sampling.

I also wanted to point out on slide CO-52, the mental health comorbidities question that the FDA had. I wondered if there was any association found between different psychiatric diagnosis and positive AvertD tests. And what I'm thinking about is particular to age. Again, I know that the rate of depression amongst older adults is specifically a factor, and I wondered if there was any age based variation amongst the comorbidities that were assessed for sensitivity and risk.

And then finally, with Dr. Zacko's slide CO-68. As opposed to a retrospective and a claim of prospective identification of risk, I wondered what safeguards the sponsor might have in place to avoid results falling into hands of healthcare insurers who might deem subsequent use of opioids by positive testers as a preexisting condition worthy of insurance revocation or denial. And then, what prevents practitioners from relying heavily, or exclusively, on AvertD as an indicator of increased OUD risk and avoiding the conduct of a complete assessment, which is time consuming and labor intensive.

I'm wondering ultimately if the sponsor could respond to my questions about the sampling and then questions about labeling and education plans they have for mitigating these potential problems.

Dr. Brauer: Yes. Thank you. I'm first going to ask Dr. Mullen to please come up and describe the basis for the sample size for the study.

Mr. Mullen: Thank you. Chris Mullen, a statistician with NAMHSA. The study sample size was driven by the power requirements overall. So, for the whole group for both sensitivity and specificity. It was not specifically powered for subgroup analyses. Generally speaking, we consider subgroup analyses to be exploratory and hypothesis generating. I think if you do look at the P values that look at differences between the subgroups for sensitivity and specificity, including things like age and race, so far, we don't have any evidence that there is differential performance of the diagnostic test by those groups. So, any variation that we see so far is, is due to chance.

Dr. Brauer: Thank you. prior to describing our planned educational program and materials that we have developed to help prevent overreliance of the test, we would like to take a minute just to talk about the genetic information Non-Discrimination Act of 2008, which was embedded in one of your questions. This law prohibits health insurers from discrimination based on genetic information of genetic health test results, and AvertD should be covered as a genetic test. As such, as such, health insurers may not request or require individuals to undergo genetic testing or supply genetic testing.

You also had a question about our educational program, and we are going to pull up a slide about the details of that program, but we'll start with just giving you an overall view of the key messages of that educational program. As a summary, our educational materials will emphasize that many factors are associated with increased risk for OUD and that some of these interactions are not well understood; that determination of any individual's absolute risk is not possible; that the use of opioids involves risk and should be under medical oversight. Genetics is one risk factor and must be used with a clinical evaluation; that AvertD is not a diagnostic test,

and it does not predict that OUD will absolutely develop or will develop; that it is important to follow opioid prescribing guidelines, particularly for negative cases as well as positive; and that resources that provide information on non-opioid alternatives would be part of our educational materials.

Dr. Higgins: Just one follow up question. and forgive me if this is not — it's a question more for the FDA and it's probably something we should say for later — but would, with respect to education, would AvertD be part of a REMS program?

Dr. Brauer: I think we should save that for an FDA question. Is that okay?

Dr. Higgins: Okay.

Dr. Watson: All right. Did all of your questions get answered?

Dr. Higgins: Yes. Thank you.

Dr. Watson: Okay. Now, Elizabeth Joniak-Grant, do you have a question?

Dr. Joniak-Grant: Yes, I do. Thank you. I have a couple of questions actually. One thing I wanted to ask was, what is the timing of the getting the results? I noticed that it ships overnight, so how long would it take for clinicians to actually get the results in the intended? So that's number one. Number two is, with the intended population, what setting do you see this being used in? Is it primary care? Are you talking ERs? That can really impact what can be done during the course of an evaluation. Another question I had is: was there any demographic data for race, ethnicity, sex for OUD positive versus OUD negative, or low risk people, that you used in your study? In determining Opioid Use Disorder, what type of clinician used the DSM-5? Was it psychologists or psychiatrists, or was it primary care physicians? And, for the initial RX, it had

to be a prescription by clinician to the patient. Were participants told this? I realize all the training materials said the individuals doing the evaluations was told this, but were participants actually told this?

And then, one important thing I think with all of this is that some of this stuff is only as good as your algorithm is. So, given that, and from my understanding reading the documents, this wasn't talked about so much, but that you utilized a neural network and that requires data, right? The more data you have, as long as you're not overdoing it, the better. And so, from my perspective, and there are other people here that are way more experts in this than I am, why was the 1300 individuals seen as sufficient, given that you you're not using the tree approach where you can weight things and bias things? It really has to be about the numbers. Why was that seen as a sufficient number? And then, finally, it's been said a number of times that those that come back with the result of high genetic risk are 18 times more likely to develop OUD. What is that in real numbers? 18 times what? Thank you.

Dr. Brauer: Great. Thank you. That's quite a list for us to work on. We're going to start with the first one. The time to get a test result is approximately 24 to 48 hours from the laboratory so that the test results can be sent to a physician very quickly.

The settings in which the tests could be used include settings such as primary care, any type of surgery prior to surgery, also prior to a dental procedure, or in the potential use in an emergency room if the timing would allow.

You asked for a slide that compared the demographics of the OUD positive versus the OUD negative participants, and we will work on getting that for you.

Next, we're onto the clinical evaluation. The clinical evaluation was conducted by a healthcare provider at the site who had been trained in the DSM-5 criteria during the study initiation process.

Your next question was, were participants aware when they were being screened for this study that the prescription for the oral opioids had to be prescribed to them? The answer is yes. The sites were specifically trained to make sure they asked the patient that the oral opioids that they took was prescribed for that patient.

You had a question about the algorithm development and the sufficiency of the sample size, and I am going to ask Dr. Donaldson to speak to that process for you. And I'm also going to ask him to speak to the pre-test/post-test probability, which I think was getting at your question of what do these 18-odds ratios mean. Dr. Donaldson.

Dr. Donaldson: Thank you. Keri Donaldson. I'll speak to these two questions in sequence. I'll start with a question on the statement, in particular machine learning, that folks think sometimes bigger data sets are better and why we chose to use a well-characterized dataset. And it really speaks to, I think, two different approaches to better understand these polygenetic traits. The first one, utilizing large data sets, as has been described throughout numerous individuals in many diseases, is that you try to overcome for limitations in the dataset or perceived limitations in the genes by adding more and more numbers of individuals. That is one approach. Another approach is to make sure your data set, which is what we chose to do, is validated for both the exposure – in this case, we're talking about a very narrow population, patients who received oral opioids for their first time – as well as a standardized outcome. Everyone needs to be assessed by DSM-5. And a third point here is to make sure the genetics

themselves are actually correct. So, when we're talking about our patient population of 1300 or so, all three of those things are true. Larger data sets, that is not usually the case.

Now, to quickly pivot to your third question on how does the 18-fold difference in between a high genetic risk and a low genetic risk map to a probability difference pre-and post-test. I'm sharing a slide right now. So, this is assuming a 5% prevalence, which is a pretty fair assessment in the intended use populations, folks that would be taking, for the first time, short-duration oral opioids for between 4 and 30 days. And that's on the left hand side of this graphic. It starts at 5%. As you follow the blue line across the likelihood ratio, which was demonstrated in this study to be 3.98, you see on the right hand side of this graph that the post-test probability changes from a pre-test probability of 5% to a 17%. So that's that difference that we represented on a bar graph earlier. The converse is true when you're looking at a patient's post-test probability with a negative genetic test result. You star with the same 5% pretest probability. Instead of following the blue line, you follow the red line on this graphic, which is the negative likelihood ratio. And what you see at the right portion of the graph is the post-test probability of a negative or low genetic risk result of 1%.

Dr. Joniak-Grant: Thank you. one very quick follow up with this. With your patient population for developing the algorithm, do you have any demographic information about this group, especially age, sex, race?

Dr. Donaldson: We do have that information. It's going to take us one second to get it up. I would say, overall, there's no differences between the OUD positive and the OUD negative groups by any of the covariates. In fact, it was based specifically to do that in the clinical study. Are you referring to the clinical study population or the training data set?

Dr. Joniak-Grant: The data set for the algorithm.

Dr. Donaldson: Yeah. So that's a good question, and that's not exactly what I answered. So, I can show you right now a little bit about that information. So, we did follow best practices in this training and learning data set. I mentioned that a little bit earlier, and one of the things important to understand is, in this setting, versus a clinical study population, it is okay to overrepresent certain groups, in particular, certain groups that are known to have higher genetic diversity. So, in this 1300 patients, there were certain groups that were overrepresented to make sure that the training accounts for that greater genetic diversity. In particular groups like African Americans, African, and Hispanics were overrepresented in this group specifically to account for that greater heterogeny.

Dr. Watson: Okay. Were all of your questions answered? Okay. Dr. Dunn, do you have questions?

Dr. Dunn: Yes. Thank you. Hi, this is Walter Dunn. Three questions. So first, did you collect any information about how or why patients recall the use of these index exposures? I think the documents that you sent out, average, mean, time, meeting time, was 8 to 10 years. I can't even remember what I did last week. I'm wondering, did the patients comment because the medical event was especially memorable, because there was intense pain, or was there any comment about the reaction to the opiates? Did they say was especially effective, or perhaps especially pleasurable, and that's why they recalled it? Or perhaps the reverse. My concern is that not that their reports are inaccurate, but that potentially you're enriching for a population that either had a very pleasurable response or a very aversive response. And so that might not be representative of maybe the bulk of the population, where you're not getting such an enhanced

signal for opiate use. Or was there any evidence that patients reported a significant anxiety or concern about the prescription of opiates at that point, at that time, and that's why they remembered it? So that's my first question.

Second question. Can you just review about the methods in which the comorbidities were diagnosed, both for the time of the index exposure and also at the time of enrollment? My impression from the clinical trial design was that this was based purely on chart review, not patient interview, and that there was not a formal diagnostic interview of the patient for these comorbidities. So, my concern is that the diagnosis of OUD was very accurate; however, the diagnosis of these potential comorbidities potentially is not as high fidelity.

And then third thing, did you collect any information from the OUD positive patients about what they personally saw as the precipitating factor for developing the use disorder? Were they able to say it was the exposure to that index course of opiates, that's what kind of got me on this pathway of developing use disorder. That's obviously subject to recall bias, but was there any information about that? Thank you.

Dr. Brauer: Great. Thank you for your three questions. We're going to start at the top with your first question about the subject's ability, or participant's ability, to recall. And this slide might help with that a little bit. As you can see, the participants in our study were prescribed the oral opioids for a significant event in their lives, such as a surgery, a dental procedure, an accident or an injury, and a swath of other reasons in 'other.' It was something of an event for that individual person, which should help prompt with the memory or recall of something, rather than the example of what did I eat last night? Probably not as memorable. But a surgery that happened to you as a patient tends to be very memorable. And I think you see that in our tier

results where 83% of our subjects in their medical records had documentation of being prescribed oral opioids. So, we're very confident in the recall that we were asking for and the accuracy of the participants recall.

You next asked a question about the methods for the comorbidities. How did we assess the information that we showed to you about participants and whether they had any significant or any documented comorbidities? We did, in fact, review the medical records of each participant to determine if there was any evidence of the comorbidities, mental health comorbidities, and substance use disorder. Before the sites did that review, they were specifically trained on how to review the medical records, areas to look for, and given a comprehensive method and way to go about reviewing those records to help ensure that they were as dedicated and, and accurate as we possibly could achieve.

You also asked if we collected any information on the OUD positive participants. Did they give us any personal feedback of what it may have been about that index exposure that may have led them to the pathway of becoming OUD positive? We did not collect that information in our clinical study, but I will ask Dr. Garbely to speak to that just for quickly for you, as he treats these patients every day,

Dr. Garbely: Good morning, Joe Garbely. as a treatment provider treating OUD and other substance use disorders, the common pathway that we see on a daily basis, taking a careful history of patients coming in for treatment of OUD, is that they often start with a legal prescription for opioids and then start to go down a path because of the likability that occurs with opioids. And they begin to ask for more prescriptions. And once they're not available, they start to look at medicine cabinets of their parents, grandparents, and friends. And eventually they

outstrip their financial resources when they go to the street and try to buy these pills, and they start to entertain use of illicit substances, like heroin and fentanyl, which right now is the major concern in regard to Opioid Use Disorders. This is the common pathway that we hear on a daily basis when we are evaluating patients that come in for treatment for Opioid Use Disorder. I'll turn it back over to Dr. Brauer.

Dr. Brauer: Thank you. Just one follow up item on the methods and the procedures that we use to identify existing comorbidities from the medical records. We did, in fact, look at the specific condition of depression and did a sensitivity analysis to say how many patients we had to have improperly classified as going from no to, yes, to see an impact on our sensitivity. And we would've needed to misclassify 45% of our participants to see an impact on performance. So, we're very confident in our results with respect to that.

Dr. Dunn: Thank you. And just to clarify for that first question about their recall of their index course. So, there was no information about the valence of that initial exposure to the opiates, patients telling you that 'yes, it was an especially pleasurable response,' 'no, I remembered because it made me really nauseous.' No information about that?

Dr. Brauer: That's correct, sir.

Dr. Dunn: Okay. Thank you.

Dr. Watson: Dr. Gordon. Good morning, thank you. Adam Gordon. My question is with regards to the outcome of Opiate Use Disorder diagnosis. I'll preface this question by a comment that there has been some talk even from our NIH director that the criteria for Opiate Use Disorder may include a pre-addiction state of the mild severity of OUD. So, the two quick questions are, number one of, it seems like that your outcome assessment included having to

have at least 2 out of the 11 criteria for Opiate Use Disorder. Do you have a slide or any material that indicates what categories of mild, moderate, severe were in the study? And number two, more importantly, I did not see in any of the material that was presented prior to the hearing or even the presentation, that the diagnosis of Opiate Use Disorder was modified with criteria if they were on an opioid. It is common practice that if you are a prescribed opiate, you cannot include a physical dependence, or dependence, or tolerance criteria of the Opiate Use Disorder criteria. And I wanted to clarify that that was modified, or it was just all 11 criteria were assessed?

Dr. Brauer: Thank you. Great. I'm going to ask Dr. Donaldson to address the second part of your question while we're pulling up the slide that shows the number of individuals with mild, moderate, and severity. Dr. Donaldson.

Dr. Donaldson: Thank you. Keri Donaldson. I think it's an incredibly important question to understand the limitations of the current definition of the disease and how it's evolving. So, as Dr. Gordon has stated, the current way in which OUD is classified includes a list of criterions. Those criterions were used based upon the study protocol that were answered in series, and two or more would be considered consistent with mild, moderate, and severe. The specific question, Dr. Gordon, that you're asking about was, were participants that were on opioid medication therapy chronically, which is slightly different, they would be excluded from this study in the protocol. So, patients that took oral opioids for greater than four 30 days were excluded and didn't get into study. And then the last portion of this is the patients that were receiving opioid treatment or chronic opioid medication for treatment for OUD, were in fact included in this study. It seems like I'm not answering the question, so I'm going to ask you to, to re-clarify.

Dr. Gordon: Yeah. I'm more interested, not about the study inclusion, but the assessment, the incidents, of use disorder that was diagnosed after they had their incident event. Was the OUD diagnosis, if they were on an opioid after that incident event that put them on the opioid, was the criteria changed after they were on the opioid?

Dr. Donaldson: I'm looking Dr. Brauer, she may understand the question.

Dr. Brauer: I believe — and just help us out. Dr. Gordon, I believe that you're asking about the eleven criteria for the DSM-5 diagnosis — was the fact that a participant was on ongoing treatment for OUD taking—

Dr. Gordon: No, no. An opiate. If they were on an opiate, not OUD treatment, but an opiate prescription.

Dr. Brauer: We're trying very hard here. Whether the individuals who were participants who were OUD positive, whether they were on — we're trying, bear with us.

Dr. Gordon: Maybe I'm being not clear. So, if a patient is on an opiate prescription and you're assessing whether they have OUD. And in assessing them for that, for your product, was the criteria changed so that you could not include dependence and tolerance criteria for OUD if they were on a prescribed opiate. That's all I'm asking. Very simple.

Dr. Brauer: Thank you. That I think helped us very much. Give us one second here. Thank you for your patience.

Dr. Garbely: Joe Garbely. Thank you, Dr. Gordon. I will give it a shot here. I think what you're talking about is whether someone is following the prescription guidelines and taking the opioid and may be physically dependent on that opioid as a result of simply taking a prescription

as prescribed. In this setting, we're looking at someone who changes the delivery system, or someone that is using the opioid in a way that's not prescribed. So, overusing the opioid on a daily basis, perhaps running out of the prescription at some point in time. Does that help answer your question?

Dr. Gordon: That's fine. Thank you. All right. Thank you very much.

Dr. Brauer: Dr. Compton.

Dr. Compton: Thanks very much. One is a very positive comment. I really liked slide ML-88. It reminded me of an old slide ruler in terms of how you display that information. It took me a while to understand it, and I appreciate you sharing that technique. My questions. One is, I noticed in one of the case report forms that you assess for tobacco use, and yet I didn't see that incorporated into the comorbidity assessment. Tobacco use disorder has been identified as one of the risk factors for people that develop problems with opioids when they're prescribed. So, I was interested in whether you had done any analysis of tobacco use and tobacco use disorder in your samples. That's also generally a very easy bit of history to identify in patients, and so if it is a positive predictive factor, it can be incorporated clinically.

The second is a broader question. You've identified in your education materials that you expect AvertD to be incorporated into a thorough clinical evaluation. Well, what do you mean by that? What is a thorough clinical evaluation? Perhaps you could show slide ML-142 again and help us understand the increased information that AvertD may provide over and above typical screening tests, as one example. But I wanted to hear a little more clarification of that.

Dr. Brauer: Great. Thank you. First, we did not look at tobacco use as one of the comorbidities, and we did not analyze the data with respect to tobacco use. I would like to invite

Dr. Garbely to come up and address your question of how does AvertD compare to the other risk assessment questionnaires that are currently available for use during the clinical evaluation.

Dr. Garbely: Thank you, Joe Garbely. So right now, where are we? Currently, we are at a point where we are doing a comprehensive patient evaluation, a careful history. That's what we do as physicians. And also incorporating risk assessment tools. These are subjective in nature and because they are, they don't have the same reliability when it comes to discerning high versus low risk, for instance, or use in acute pain management. The other part of this is we will be querying the PDMP, Prescription Drug Monitoring Program of our state, which incorporates other states as well, neighboring states, to see about prescriptions that may have not come up or prescriptions that were not remembered by the patient. In addition to that, we will review the medical record that is available to us, and that's where we are right now. So that is a very subjective way of assessing risk currently. The AvertD device will help us move from the 50 yard line down the field with objective data. Not to replace what I just said is the current state of assessment, but to enhance it. This will fortify the therapeutic alliance, and it will help enrich the conversation we need to have as clinicians with our patients describing risk versus benefit of opioid prescription for acute pain management.

And I think it's important to understand that this particular device is all about acute pain management. It is not about chronic pain. Those with chronic pain were excluded from this study. And this might help Dr. Gordon's question, as well, because those that are on chronic opioids were not part of this study. These are patients that needed acute pain management.

Dr. Brauer: Did that answer your question, Dr. Compton?

Dr. Compton: Yes, it did. Thank you.

Dr. Watson: All right, James, I just wanted to point out, I realize we're over time now, but I think it's really important we get all these questions answered. So, I think we'll make up time at the lunch break. And with your permission, I'd like to continue on. Okay, Dr. Farrar?

Dr. Farrar: Yes. I have two questions and a comment. I very much like the M-88 likelihood slide as well, but I would point out that the base of that slide was a population risk of 5%, and it's been clearly stated by the presenter that the estimated risk is 1%. If you move the baseline to 1% and roll lines again, the likelihood ratio drops to 3. And I think that this is a key issue. A back of the Excel spreadsheet analysis of a prevalence of 3 in the population suggests that we are going to have false positives at the rate of something like a 50%. you're going to get a true positive of about 4 and a false positive around 180. So, you're going to misdiagnose many, many more people than you're going to be able to diagnose with this test. So, I wondered what the sponsor has to say about that.

The second issue is the comment about how this will be used. I think, as evidenced by the CDC recommendations that came out in 2016, recommendations are abused and grossly abused when they're applied in the general population, primarily from a lack of understanding. The very issue that was just raised by Dr. Garber was that this is an “objective” test, and I can see that clinicians, without thinking about it, will use the objective test and will not proceed any further and avoid opioids in those patients. Whereas what was said by the neurosurgeon was that they would then evaluate what they would do, but he himself said he would avoid opioid use, even in a population who are in severe pain, potentially, who are not able to take NSAIDs. And so, my concern is that, with the misdiagnosis process of a very small positive predictive value, and the tendency for medical professionals to depend on objective versus subjective testing, that patients

could be severely damaged or severely mistreated by this. And I wonder what the sponsor has to say about those two specific points.

Dr. Brauer: Sure. Thank you, sir. if we may, we're going to break your question down into three areas. We're going to have Dr. Donaldson do the pre-test post-test probability for 1% disease prevalence.

Dr. Donaldson: Thank you Dr. Farrar for the ability to talk about this. So, you're correct, and this slide shows a lower end of prevalence at 1%. Using the same positive and negative likelihood ratios, it scales through the same method we explained before, to post-test probability somewhere in the neighborhood with a high genetic risk result of 4% and low genetic risks result of less than 1%. Of note, the difference between those two, that 18-fold difference, is consistent and independent of prevalence, as we earlier said. Yes. The other note. The other note that I want to say here is that the 1% and 5% differences is talking about the 1% in the US population, which includes people that have and haven't been exposed to opioids. The reason we chose to put the 5% prevalence in front of you first is because that is the consistent estimate in the intended use population: patients that are opioid naive, that are first given an oral opioid prescription for acute pains. That's why we chose to lead with the 5%. But I'm happy to show this 1%, and I'm going to hand it off here to the statistician to answer a portion of your second question.

Mr. Mullen: Thank you. Chris Mullen. And if I have the wrong question being answered, please let us know. Just one thing to point out the intended use is actually for the product as an aid, not as a diagnostic test in and of itself. I think I would concur with what Dr. Donaldson said about the importance of prevalence. Of course, that is an assumption about how the test would be applied if approved. For the design of the study, of course, where we wanted to

avoid issues of prevalence and assumptions, that was the rationale for choosing sensitivity and specificity, as well as the likelihood ratios and the diagnostic odds ratio.

Dr. Brauer: Did that answer your questions, Dr. Farrar?

Dr. Zacko: One more. Chris Zacko. Yes. Two more. Yeah. Thank you for the question. So, what I want to do is go over how I see this test being used, and I feel maybe it is a misunderstanding. I see this test being important in both assessing a patient preoperatively and talking to them about the risk of a given surgery, plus the risk of potential for opioid problems afterwards. And then after surgery, as we try to control their pain, and particularly in complex surgeries, what do we do to do that? So. I always use an opioid sparing, not avoidant, strategy and use multimodal analgesia. So, where this test can be informative is the patient will know ahead of time whether or not opioids could be a problem for them. And then afterwards, as we discuss various options for controlling their pain, that conversation could be rehashed to a degree, right? So, then it would also help me know how long to stick to some of the other options other than opioids with that. But the important thing to know is, I need to control the patient's pain, because while opioid use after surgery leads to poor outcomes, so does poor mobility. So, I need to control their pain. I would not not give them opioids if that's what they needed. It just needs to be informed decision making with that patient.

Dr. Farrar: If I could just follow up on that specific, and then I'd like to come back to the positive predicted value with Dr. Donaldson. But this specific issue is you are very well versed in this product, and you are clear about how you might use it. The issue is and I think that your educational process is probably a good one, but we know that education of practicing physicians has relatively small effect on their behavior. My concern is that people will use this and

potentially deny opioid use for patients in severe pain, even though they potentially need it, primarily because this test will identify a risk, but it's not a hundred percent. And Dr. Donaldson's comment about 1% versus 5% I think is telling. I agree, 5% of people are at risk, but 1% end up being in the opioid abuse bucket. We have the potential for denying it. Observing these patients carefully and preventing long term opioid use and preventing opioid abuse is clearly a key feature of what needs to be done. What I would ask you is, isn't that what you're already doing, which is that you're using opioid sparingly, you're using opioids only when absolutely necessary. So how does this add to that, other than to give you a one in a four out of five chance of being correct?

Dr. Zacko: It's a great point. So yes, this is the strategy I already use. Where it can be informative is for the patient to know. And then what it can also do is shape potentially surgeries that we end up doing. Like you've already alluded to, the continual reassessment of pain after surgery is going to be really important. And I think there would be more comfort with myself, or the patients, if they had a negative test, to extend some of the opioid usage a couple days. But again, it's really important to continually reassess that. And again, I want to emphasize I would not leave a patient in pain. So even if they do test positive, but we can't control their pain, we would just have to have a very good conversation about how to manage their pain afterwards. And I agree with your concern. We've all seen tests that are taken to black and white, and the educational materials, and Dr. Brauer will speak to those soon, would be really, really important to emphasize it to practitioners.

Dr. Brauer: We do just want to share with you our plans for a controlled launch. We know this is a new test. It is the only tool available to assess genetic risk. The other tools out

there do not assess genetic risk. So, we are committed to making sure that physicians and patients have information about the test, and we are committed to doing a controlled launch in order to ensure that we are doing our best with our educational materials and providing individuals, physicians, and patients, what they need. Our plans are to work with a few select centers of excellence: to first work with them, to provide them our educational materials, to refine those educational materials and process, to make sure we are getting the user's needs and getting the key messages out there. We will collect feedback from this process and continue to further refine our materials and education as needed. We are committed to helping ensure that users use our product correctly.

Dr. Watson: All right. Thank you very much. Panel, I just would like to note we're very far behind. I think it's really important to get all these questions out. If we could keep them to clarifying questions and limit discussion until we have our open discussion, our discussion period, that would be useful.

Dr. Farrar: Yeah, that makes sense. And if I could just ask Dr. Donaldson to calculate positive predictive value and the number of true positives and false positives for 1 and 5% presented to us at some point, that would be useful. Thank you.

Dr. Watson: Okay. Thank you Dr. Zaafran.

Dr. Zaafran: Thank you. Sherif Zaafran. A couple of questions I had, just because of my concern of the unintended consequences, not in the physician community that was alluded to earlier, but actually in the regulatory community that will regulate physicians and how they are prescribing these medications. Number one: I still see that race — that there's a significant underrepresentation and overrepresentation of the white population. I think there was only 12

that was a non-white. I'm not sure what the breakdown of that 12 was. I have significant concerns there, especially when we're talking about genetic testing and variations with race as it as it alludes to that.

And again, my second question is around stigma, especially as an “objective,” quote unquote, test that, theoretically, will say that somebody's going to develop Opioid Use Disorder, and is that going to develop a stigma and have an under treatment of acute pain? I think that was alluded to earlier, and how you're going to address that.

And the third one, which I haven't heard anything about right now, is this seems like an all or nothing kind of thing. Was there any looking at the delineation of different types of opioids, or more importantly the amount of opioids and the duration of the opioids that were being used, and the risk stratification of people who are susceptible to that. And I'll leave discussion until later because of time.

Dr. Brauer: Great. Thank you. let's start with your question about race. And we are confident in our test performance in different groups of patients based upon race and ethnicity. Our confidence comes from our training and development sets, which included over 3000 individuals and overrepresented, Asians, African Americans, and overrepresentation of Hispanics. Our clinical study was designed to mirror the intended use population of adults being prescribed oral opioids for acute pain. And we did not see any statistically significant difference in test performance based upon race or ethnicity, although we recognize the sample size for certain groups was small. With that, we are committed to further evaluating AvertD in African Americans in the post-market setting should we be approved. We would be pleased to work with FDA in this regard.

As a quick answer to your question, all participants in our study had oral opioid exposure of four to 30 days. We do not know the more detailed level of who had five versus six, but we know their initial exposure was all 4 to 30 days. And I'd like Dr. Garbely to come up and speak with you about your question regarding potential stigma.

James Wang: Hi, this is James Wang. I just want to say we should start moving forward. We will give the sponsor more time in afternoon before the panel deliberations to answer more questions, but if the panelists just have a few questions they can provide to the sponsor, and the sponsor can get the questions answered during lunch, and then we can answer them during the deliberation. Does that work?

Dr. Watson: I'm not clear. You're saying stop questioning right now?

James Wang: I know we have a couple hands up. If the doctors who have their hands up can just briefly discuss their question, and then the sponsor can take note, and we'll give them time in the afternoon to answer them.

Dr. Watson: So, we'll have the questioner pose the question, but the sponsor will not answer right now. Just take notes and respond in the afternoon. Is that correct, James?

James Wang: Correct.

Dr. Watson: Okay, great. Okay, we have a question on the table, but sponsor, please take note of that and answer in the afternoon. Can we go now to Dr. Ness?

Dr. Ness: Okay, I'll try to make it quick, in the sense that the sponsor has done a very good job of saying this is a tool to use coupled with clinical information. We forget that we do collect a little bit of genetic information as part of that clinical history in the form of the family

history, in that there's a high relation between first degree relatives having substance abuse disorders and then the subsequent individual having it. So, my question is: they reviewed through these medical records, did they collect any of that data related to first degree relatives? Because that would tell us whether this test is actually any better than that clinical history. And it might also tell us if, with our false negatives, whether we're missing something, that there may be more than 15 genes that are important.

Dr. Watson: Thank you for that question. You'll have the answer in the afternoon. Dr. Wang, do you have a question?

Dr. Wang: Yes. Thank you. I have a couple questions. The first question is, I'm hearing conflicting statements regarding the clinical value of the test. On one hand, I'm hearing that this is 18 fold higher likely little ratio. But on the other hand, on the educational material slide, there is a hatch statement saying that this is not diagnostic. It's only an aid and does not predict development of OUD. And the intended use also stated very clearly that the result is going to have to be combined with a clinical information to make the final interpretation. However, I do not see this any of the clinical information combination or how to interpret the test result as being tested in the current clinical study. So, my question to the sponsor is how exactly is the test going to be combined with clinical information and how is it going to be tested in clinical study or in any future clinical study that’s being planned?

The second question is about the population. So again, the intended use population, as mentioned before, are patients who are going through primary care, pre-surgery, pre-dental procedures. The estimated prevalence of OUD in this and using opioid oral opioids for the first time as well. So, the estimated prevalence is pretty low in this population. However, we see the

clinical study population is highly enriched with individuals who are already who already have OUD and are undergoing treatments. So, I want to hear some comments from the sponsor how well this represents the intended use population, and also have concerns on the low predictive value of the test in this low prevalence intended population. I hope that is clear.

Dr. Watson: Thank you. Thank you, Dr. Wang. We'll discuss the answers in the afternoon. Dr. Walker, do you have questions?

Dr. Walker: Yes. Thank you very much. I have a clarifying question on the second response to questions from the panel. You indicated that in the blinded clinical study there were no difference in seeing between race and ethnicity. But I just want to have it clarified that that study was not indeed empowered to see those differences if they existed. Thank you.

Dr. Watson: Thank you, Dr. Walker. We'll have the answer in the afternoon. Dr. Bierut, do you have questions?

Dr. Bierut: Yes. Thank you. I have a question actually about the genetic variance that are selected. Knowing that the modern genetic studies are actually identifying different variants associated with the disease. So, given that these genetic variants were selected, and that they differ also by race, I'm curious what the distribution is of a positive test, or an at-risk test, is across the different ancestral groups. Because as I view it, I think that the population who really this is intended for is going to be probably the general population, because I anticipate that all of us will have a surgical procedure, a dental procedure, some type of procedure that puts us at risk for opioids and for Opioid Use Disorder. And I'd like to see the distribution in the general population to know, are there racial differences in ancestry differences that are seen?

Dr. Watson: Thank you, Dr. Bierut. We'll discuss that in the afternoon. Dr. Goldstein, do you have question?

Dr. Goldstein: Yes, Thank you. I'd like to know the exact number of African Americans in the study population, and I'd like to know what's the frequency of African Americans who declined to participate in this clinical study. We know that African Americans in general decline to participate in clinical trials.

Dr. Watson: Thank you. We'll have answers to those very important questions in the afternoon. Dr. Bateman, do you have questions?

Dr. Bateman: Yes. Thank you. I would like the sponsor to comment on the work by Hatum (phonetic) and colleagues that was included in the briefing book when they've looked at the association between the SNPs included in the predictive test and OUD, found that the association was likely highly confounded by genetic ancestry. I'm wondering why the sponsors haven't used the methods that have been developed to account for genetic admixture in adjusting for those potential differences as a potential explanatory factor. And I think this is a really concerning point, because it's possible that the introduction of this test could lead to disparities or exacerbate disparities in opioid prescribing that we know already exists, and pain management. I know there was data presented stratified by race, but I think a simple white/non-white stratification may be insufficient to account for the genetic admixture differences that might be driving some of the association that's been observed.

And then as a second point, it would be useful for them to comment why genome-wide association studies have not identified the SNPs that are included in the diagnostic test. I think, in the briefing document, the largest gene swab and other genome-wide studies that have been done

have identified only one of the 15 SNPs that are included in the test as being in independently predictive of OUD and with a very weak effect. So yeah, if they could comment on those, on those issues.

Dr. Watson: Thank you. Okay. these are all such important questions. I think, James, we'll need them before our panel deliberations. So, I guess before we begin that, we'll come back with the answers to these questions. James, is that correct?

James Wang: Yes, that is correct. After the Open Public Hearing session, we'll give the sponsor some time to answer these questions.

Dr. Watson: Perfect, perfect. Now, if we could just take a two to three minute bathroom break and then come back for the FDA Presentation to get us back on schedule, and then we'll have a lunch break, which we'll truncate a bit to try to — this is a long day — to try to get us to end at a reasonable hour. So, if everyone's okay with that, everybody break, and we'll meet back in two minutes at 8:30 my time, 11:30 your time.

Dr. Watson: Thank you everyone for returning from break. We are now going to proceed with the FDA Presentation. I'd like to remind the public observers at this meeting, while this meeting is open for public observation, public attendees may not participate except at the specific request of the panel chair. The FDA will also have 60 minutes to present. FDA, you may now begin your presentation.

Dr. Gussow: Good morning and thank you for joining us for today's meeting. My name is Dr. Keisha Melodi Gussow, and I will be presenting today on behalf of the FDA. I am the Team Lead for the Toxicology branch within the Division of Chemistry and Toxicology, DCTD. DCTD is in the Office of Health Technology Center, Office of In Vitro Diagnostics, within the

Office of Product Evaluation and Quality in the Center of Devices and Radiological Health, CDRH.

The topic of today's meeting is the de novo request from SOLVD Health regarding the AvertD device. The meeting agenda is as follows. I will first provide a summary of the device, regulatory review process, and the indicated disorder, which is Opioid Use Disorder. Next, I'll describe the device, summarize the regulatory history, and summarize the topics of discussion for today's advisory committee meeting. In the remainder of the presentation, I'll summarize the clinical study provided by the company in support of the de novo request, followed by a summary of the questions for the panel.

The subject device, or test, is called AvertD. I will use the words device and test interchangeably throughout this presentation to mean the same thing. AvertD is a first of its kind for a novel device that is intended to be used for the genetic risk prediction of Opioid Use Disorder in patients receiving prescription oral opioids for the treatment of acute pain for the first time. FDA review of novel moderate risk devices is conducted through the de novo review process.

The de novo review process is a risk-based classification process. During the review process, FDA determines if probable benefits outweigh probable risks, identifies probable risks to health for the device or product, determines level of control needed, with Class I devices needing general controls only, and Class II devices needing general controls and special controls. Clinical and/or non-clinical testing can be considered as part of the decision making and can be considered a special controls. Collectively, these provide a reasonable assurance of safety and effectiveness.

The device that is subject of this de over request is intended to predict the genetic risk of developing Opioid Use Disorder, hereafter referred to as OUD. OUD is typically diagnosed by a clinician using the most recent version of the Diagnostic and Statistical Manual of Mental Disorders, the DSM-5, which is the handbook used by healthcare professionals in the United States and much of the world as the authoritative guide to the diagnosis of mental disorders. The DSM-5 describes OUD as a problematic pattern of opioid use leading to clinically significant impairments or distress, as manifested by at least two of the following occurring within a 12 month period. While I will not read through each of the following, the list provides context for the condition that the device is intended for and highlights the complexities of diagnosing OUD. The list summarizes serious and dangerous situations where patients are failing to fulfill major obligations, having persistent social problems. and have recurring opioid use in situations in which it is physically hazardous for them.

Many factors may contribute to the risk of developing OUD, including genetic factors like heritability. However, the specific genetic associations identified so far explain only a portion of OUD risk. Given the ongoing opioid epidemic, a test demonstrating that the probable benefits outweigh probable risks could have some interesting public health benefits, benefits that could help limit higher risk opioid exposures, while maintaining availability for patients who really need it. The risks associated with false positive and false negative results should be carefully considered for tests of this type. For example, patients erroneously identified as being at high genetic risk of developing OUD could be deprived of much needed pain treatment while patients erroneously identified as having low genetic risk or not at risk could be exposed to opioids without appropriate precautions. There may also be emotional ramifications and stigmas associated with genetic testing.

The currently proposed indication for use of the device is as shown. AvertD is a prescription, qualitative genotyping test used to detect and identify 15 clinically relevant genetic polymorphisms in genomic DNA isolated from buckle samples collected from adults. The 15 detected genetic polymorphisms are involved in the brain reward pathways that are associated with Opioid Use Disorder, OUD, and identify subjects who may be at increased genetic risk for OUD. Information from AvertD provides subjects 18 years of age or older and healthcare providers with objective information to be used for informed decision making prior to the first prescription of oral opioids for acute pain. The information from AvertD is intended to be used in combination with a clinical evaluation and assessment of the subject.

The device is a multiplex genotyping assay intended for use in testing human DNA collected from buckle swab specimens. To generate a result, first, a healthcare provider collects a buckle swab sample that is sent to the lab, where the DNA is isolated, amplified, and purified prior to detection of the 15 SNPs on a microarray. The result of the genotyping of the 15 SNPs shown in this table are fed into a machine learning algorithm that yields a qualitative output for risk – either yes, no, or not applicable.

Currently, there are no FDA cleared or approved tests for identifying patients at genetic risk of developing OUD. An earlier version of the AvertD device that included 11 of the 15 single nucleotide polymorphs, or SNPs, that are included in the current version of the device received breakthrough designation in 2018. FDA granted breakthrough designation because the device met the criteria described in the breakthrough guidance. Namely, FDA determined that the device may allow for more effective treatments or diagnosis of a life-threatening disease or

condition. The device represents breakthrough technology, and there are no clear or approved alternatives, and the availability of the device could be in the best interest of patients.

Devices granted breakthrough designation are subject to pre-market review, in this case, in a de novo classification request. The company first submitted a de novo classification request for the 15 SNP device that was declined due to uncertainties in the clinical and non-clinical testing and interpretation of the results. The decision to decline the de novo request was upheld upon appeal in January of this year, 2022. After the appeal was upheld, the company proposed approaches to address questions raised by the FDA in the initial FDA submission review. FDA provided feedback on the proposed approaches, after which the company went back to the clinical study sites and collected additional information about the clinical study subjects from the medical records and medical histories that were available at those sites in an effort to address the questions raised by the FDA in the initial FDA submission review. The new information, along with the previously submitted studies, are currently under review in a new de novo classification request.

FDA states the Panel's assistance in interpreting the data from the clinical study in assessing its applicability to the intended use population and in providing their perspective on the benefit risk assessment. The two main topics of discussion are whether the results of the clinical study adequately represent performance in the intended use population and setting, and whether the probable benefits to help from use of the AvertD tests outweigh the probable risks for the proposed indications.

In assessing device performance, the FDA considers the analytical, or nonclinical, testing data, as well as the clinical testing data. FDA is still evaluating the analytical performance

characteristics of AvertD and expects to be able to resolve any outstanding questions regarding the analytical testing. We will not be seeking Panel input on these studies in this meeting. Today, we'll focus on the clinical testing to support the claims. A clinical study was conducted to assess performance of the device to support its intended use.

Before diving into the details of the clinical study, here is a summary of the main points. The clinical study conducted by the company was a prospective, observational study with one retrospective element, which was the initial, or index, exposure to prescription oral opioids prior to study enrollment. The purpose of the study was to determine risk of developing OUD following such opioid exposure. The study enrolled subjects with a self-reported index exposure at least 12 months prior to enrollment and included subjects with a self-reported index exposure ranging from 1 to 51 years prior to enrollment. The index exposure dates were based on the subject's recollection, raising uncertainty regarding the accuracy of the index exposure. Additional information about the enrolled study subjects was collected from the enrollment sites. After the study was completed to support the self-reported information. The additional information was also collected to support that clinical study subjects met the inclusion and exclusion criteria as they were written in the clinical study protocol and to collect documentation on comorbidities. Due to the low prevalence of OUD in the United States, an enrichment strategy was used to increase the likelihood of enrolling a sufficient number of OUD positive subjects. The enrichment strategy involved the enrollment of study subjects from sites that offer opioid treatment programs. Subjects for enrolled and buckle samples were collected from the subjects at 10 sites in the United States and tested using the AvertD to determine genotypes. A total of 385 subjects were included in the clinical study population, with 381 who had valid results from the AvertD tests included in the final clinical data analyses and estimates.

Results of the AvertD were compared to the OUD status determined by clinical evaluation during enrollment. A total of 174 OUD positive subjects had AvertD results and were included in the final analyses. The overall sensitivity for the percentage of true positives detected by the AvertD was determined to be 82.8%. A total of 207 OUD negative subjects had AvertD results and were included in the final analyses. The overall specificity or the percentage of true negatives detected by the AvertD was determined to be 79.2%. Additional information regarding the study findings and questions about the applicability of the findings to the intended use population will be discussed in coming slides.

Study subjects were enrolled while seeking routine care at 10 sites in the US. The vast majority of enrolled subjects came from practitioners that were familiar with the clinical history of the subjects, and presumably had a relationship with the subjects. 812 subjects were enrolled, and all were determined to meet the enrollment criteria. Due to the low prevalence of OUD, the company enrolled subjects from sites where subjects may be seeking treatment for OUD to enrich the population. 2 of the 10 sites, sites 10 and 11, named the Karen Pennsylvania Treatment Center and the Seven Hills Hospital, respectively, offer opioid treatment programs and are listed on the Substance Abuse and Mental Health Services Administration, or SAMHSA, Opioid Treatment Program Directory.

These sites that offered opioid treatment programs also had at least one healthcare provider who held a waiver to prescribe buprenorphine, which is used to treat OUD. To receive a practitioner waiver to administer, dispense, and prescribe buprenorphine, practitioners must notify SAMHSA's Center for Substance Abuse Treatment of their intent to practice this form of medication-assisted treatment. The Notification of Intent, or Buprenorphine Waiver Application,

must be submitted to SAMSA before the initial dispensing or prescribing of OUD treatment. Medication. Qualified practitioners who undertake required training can treat up to 100 patients using buprenorphine for the treatment of Opioid Use Disorder in the first year if they possess the waiver and meet certain conditions. One additional site, Site 2, named Clinical Research Associates, also had at least one healthcare provider who held a waiver to prescribe buprenorphine. However, Site 2 is not an Opioid Treatment Program site. OUD positive subjects were enrolled from all three sites who had at least one healthcare provider who held a waiver to describe Buprenorphine and OUD. Negative subjects were enrolled at sites 2 and 10, but not at site 11.

At the time of enrollment, all enrolled subjects were clinically evaluated to determine whether they met the DSM-5 criteria for OUD, resulting in a diagnosis of OUD positive or OUD negative. Subjects for groups into a higher risk pool if there was evidence of a substance use disorder, or SUD, or OUD. Subjects were grouped into a low risk pool if there was no evidence of an SUD or OUD. By definition, the low risk pool did not include OUD positive subjects, and, by definition, all subjects in the high risk pool either had evidence of OUD or evidence of an SUD. The purpose of the pools was to enrich the population with the higher prevalence of OUD positive subjects than is expected in the general population. Information from 689 of the 812 enrolled subjects were forwarded to a statistician, who selected 385 subjects for clinical study analyses. The statistician was not provided the OUD status of the subject but was provided the subject's demographic information and the risk pool assignment, which was based on the presence or absence of an SUD or OUD.

During enrollment, information was collected from the subjects using four different case report forms. Versions one to three of the case report forms did not include a list of inclusion and exclusion criteria. Version four included a complete, albeit different, list of inclusion and exclusion criteria. The differences are summarized in the table below. Version one was used in the enrollment of 61 subjects and stated that the minimum number of consecutive days for opioid use was five days, as opposed to the four days described in the intended use statement and used in following versions of the case report form. This version also did not include a maximum number of days. Subjects were also not asked to provide a self-reported index exposure date when enrolled using this form. The company stated that the index exposure date was identified by records associated with the study. No other inclusion and exclusion criteria were included on this version. Version two, which was used in enrollment of one subject, states the minimum number of consecutive days of opioid use is four days, and no maximum timeframe was included. The subject enrolled with version two was also not asked to provide a self-reported index exposure date, and records associated with the study were used to identify the date of index exposure. No other inclusion and exclusion criteria were included on this version. Version three, which was used in the enrollment of 41 subjects, states the timeframe for consecutive days of opioid use, plus 4 to 30 days as described in the intended use, and asked the subject to recall the date of index exposure. No other inclusion and exclusion criteria were included on this version.

Each of the 10 sites were trained individually prior to enrolling subjects using the study protocol and a training deck, or set of slides, that summarized the inclusion criteria, but did not summarize the exclusion criteria. The inclusion and exclusion criteria were discussed with each of the sites during training. The clinical study protocol, training material and case report form

version four included different criteria. Specifically, one of the inclusion criteria listed in the clinical study protocol states, "Subject was exposed to prescription oral opioids for a duration of 4 to 30 consecutive days, or a psychiatrist has diagnosed the subject as having OUD according to DSM-5 criteria." In the training deck, this criteria is described as a minimum exposure of four consecutive days to prescription oral opioids and never received medical care that included taking prescribed oral opioids for more than 30 consecutive days.

In version four of the case report form, this criteria is listed as, "Subject has taken prescription oral opioids for at least four consecutive days and not more than 30 consecutive days. Date subject first took prescription oral opioids for at least four consecutive days and not more than 30 consecutive days." Neither the training deck nor version four of the case report form include a provision for inclusion of subjects who were previously diagnosed with OUD and who may therefore have received oral opioid treatment for greater than 30 days. Notably, none of the criteria state that the subject took prescription oral opioids that were prescribed to them by a healthcare provider for the treatment of acute pain.

Additionally, one of the exclusion criteria listed in the clinical study protocol states, "Subject has never received medical care that included taking oral opioids for more than 30 consecutive days unless a psychiatrist has diagnosed the subject as having OUD according to DSM-5 criteria." This criterion was not listed in the training deck, and in version four of a case report form, this criterion is listed as, "Subject has never received medical care that included taking prescription oral opioids for more than 30 consecutive days."

The differences in the criteria listed on the four versions of the case report form the training deck and the clinical study protocol introduced uncertainty into whether different

populations were enrolled and whether the clinical study population represents the intended use population. To address this uncertainty, as well as others described in upcoming slides, the company collected additional information about the clinical study subjects from the medical records and medical histories that were available at the enrollment sites after the clinical study was completed. 'Medical records and medical histories' is defined by the company as including information that may be based on subject memory, such as the reason for the visit or chief complaint, past surgical history, past medical history, and prescription history, as well as other information such as the review of systems, procedure and operative notes, radiology reports, consults, current medications, and summary of findings.

Three forms for use to collect additional information, and each form collected different information to address different uncertainties in the clinical study. These three forms are distinct from the previously mentioned versions of the case report forms, which were used during subject enrollment. The subjects were not contacted to complete these forms, and only records available at the sites the subjects were enrolled were queried to complete the forms. Form One was used to collect information to address uncertainty surrounding the enrollment criteria. The goal of form one was to collect information from the medical record and medical histories to support that clinical study subjects met the inclusion and exclusion criteria as they were written in the clinical study protocol. Based on the information selected on Form One, the company asserts that all subjects met the inclusion and exclusion criteria as listed and as intended in the clinical study protocol. That is 100%, 385 out of 385 subjects.

As stated previously, subjects were enrolled from 10 US sites, two of which offer opioid treatment programs, and had at least one provider who held a waiver to prescribe buprenorphine,

sites 10 and 11, and one of which only had a provider with a waiver, site two. The company conducted sub-analyses of the performance of the AvertD in specialized sites versus unspecialized sites to determine whether the performance in the different subject populations was different than observed in the overall clinical study population.

This table summarizes the device performance when sites offering opioid treatment programs, Sites 10 and 11, are grouped together. The sensitivity of the AvertD in sites offering opioid treatment programs is 86.5%, with a 95% confidence interval of 79.6% to 91.3%, as compared to 82.8% in the total clinical study population. The specificity was 80% with a confidence interval of 49% to 94.3%, as compared to 79.2%. The sensitivity in sites that did not offer opioid treatment programs was 70.7% were a confidence interval of 55.5% to 82.4%. The specificity was 79.2% with a confidence interval of 72.9% to 84.3%. Of note, 76.4% of the OUD positive subjects in the clinical study were enrolled at opioid treatment program Sites 10 and 11.

This table summarizes the device performance with sites that had at least one provider with a waiver to prescribe buprenorphine, Sites 2, 10 and 11, are grouped together. The sensitivity in these sites was 82.7% with a 95% confidence interval of 76% to 88.2%, as compared to 82.8% in the total clinical study population. The specificity was 89.5% with a confidence interval of 75.2% to 97.1%, as compared to 79.2%. The sensitivity of sites that did not have a provider with a waiver to prescribe buprenorphine was 83.3% with the confidence interval of 51.6% to 97.9%, and the specificity was 76.9%, with a confidence interval of 69.8% to 83.1%. Of note, 93.1% of the OUD positive subjects in the clinical study were enrolled from Sites 2, 10 and 11.

Because the majority of OUD positive subjects were recruited at opioid treatment program sites, there was uncertainty surrounding whether the clinical study population included a higher prevalence of severe OUD cases than is expected in the general population. The DSM-5 provides guidelines for determining OUD severity based on the list of symptoms for diagnosing OUD. Mild OUD is described as having two to three of the symptoms for diagnosing OUD, moderate as having four to five symptoms, and severe as having six or more symptoms. Although prevalence estimates vary. One study in the chronic pain population found that the prevalence of mild OUD was 28.1%, moderate OUD was 9.7%, and severe OUD was 3.5%. In other words, the prevalence in that study was greater for mild OUD than moderate OUD, and the prevalence of moderate OUD was greater than severe OUD.

To address the uncertainty sub-analysis were conducted. As stated, 76.4% of the OUD positive subjects were enrolled at sites that offer opioid treatment programs, Sites 10 and 1. Of these subjects, information about the severity of the OUD was provided for a total of 132 out of 133 subjects. 126, or 94.7%, of the OUD positive subjects enrolled at opioid treatment program sites had severe OUD. 93.1% of the OUD positive subjects were enrolled at sites with at least one healthcare provider that held a waiver to prescribe buprenorphine, Sites 2, 10 and 11 of these subjects.

Information about the severity of the OUD was provided for a total of 160 out of 162 subjects. 129, or 79.6%, of the OUD positive subjects enrolled at these sites had severe OUD. In total, the majority of OUD positive subjects in the clinical study population, 74.1% had severe OUD. The high rate of severe OUD in the study population raised uncertainty as to whether the clinical study population adequately represents the intended use population, since the clinical

study population had higher rates of severe OUD than expected in the intended use population. No additional information was collected to address this uncertainty.

The subjects enrolled in the clinical study reported an index exposure date ranging from 1 to 51 years prior to the date of enrollment. The percentage of OUD positive subjects increased as time since self-reported index exposure increased, with 28.6% of the subjects who reported an index date between one to three years prior to enrollment being OUD positive, up to 75% of the subjects who reported an index date greater than 25 years prior to enrollment being OUD positive.

The use of self-reported dates to enroll subjects into the clinical study raised uncertainties about subject recall dating as far back as 51 years prior to enrollment and about the accuracy of the information that self-reported date was a legitimate exposure to oral opioids that were prescribed to the subject for acute pain. Additional information was collected after the clinical study was completed to address some of the uncertainty, as described previously. Form two was used to collect information from the medical record and medical histories to support the accuracy of the self-reported index exposures collected during enrollment of the study subjects. The information was captured in tiers that describe the type of information that was available for each subject.

Tier 1 includes subjects whom the company determined had information to support that they met the inclusion and exclusion criteria listed in the clinical study protocol. Tier 2 includes subjects who have documentation in their medical record of a procedure or event occurring, such as a surgery or an accident, for which oral opioids may have been prescribed for acute pain as part of medical care within a calendar year before or after the self-reported index exposure. Tier

4 includes subjects who have a description in the medical records of an oral opioid prescription for acute pain within a calendar year before or after the self-reported index exposure but may or may not have documentation of the actual prescription. An example of this might be a record that states a subject was prescribed seven days of hydrocodone for knee surgery, but the prescription may or may not be documented. Tier 4 includes subjects who have documentation of an oral opioid prescription for acute pain within a year of the self-reported index exposure. All subjects in Tier 4 also had documentation of a procedure or event, Tier 2. Tiers 5 and 6 describes subjects who either have no information in Tier 2, 3, or 4, or who have information in all tiers respectively. Since all of the clinical study subjects are described in Tiers 1 to 4, Tiers 5 and 6 will not be discussed in today's presentation.

I'll begin with Tier 4, which had the most information about study subjects. In other words, subjects with information in Tier 4 had prescription documentation, as well as information about a procedure or event that may be related to or oral opioid prescription. Prescription documentation includes a physical, electronic, scan, or photograph of a prescription that is stored in the medical records and medical history at the enrollment site. Prescription documentation was available to 35% of the study population. When broken down by time since the self-reported index exposure, prescription documentation were available for subjects across the range of self-reported index exposure dates, including for 37.5% of subjects who self-reported index exposure date was more than 10 years prior to the date of enrollment. The sensitivity of a burden was 70.7% in Tier 4 with a confidence interval of 54.5% to 83.8%, as compared to 82.8 in the total clinical study population. The specificity of AvertD was 84.8% in Tier 4, with a confidence interval of 75.8% to 91.4%, as compared to 79.2% in the total clinical study population.

The sensitivity and specificity of the AvertD in subjects with information in Tier 4 was compared to Tier 1 As a reminder, Tier 1 includes all of the subjects that were included in the clinical study. No OUD positive subjects were enrolled at sites 1, 4, 6, 7, and 9, and therefore, sensitivity could not be calculated at these sites. For sites for which sensitivity could be calculated, sensitivity varied with no clear observable trend. No OUD negative subjects were enrolled in Site 11, a site offering opioid treatment programs. Therefore, specificity could not be calculated for Site 11. For sites for which specificity could be calculated, specificity varied with no clear observable trend.

Looking now at Tier 4, no prescription documentation was available at sites 6, 7, 10, and 11. Sites 10 and 11 were sites that offered opioid treatment programs. Site 6 was a clinical practice site that participated in research, and Site 7 was a research-only site. Prescription documentation was available at Site 2 for all subjects, and therefore sensitivity and specificity estimates remained the same in Tiers 1 and Tier 4. Prescription documentation was available for most subjects at Site 3, and the sensitivity and specificity estimates are similar in Tiers 1 and 4. At Site 4, only one OUD negative subject was enrolled, and the prescription documentation was available, and specificity did not change. Prescription documentation was available for two OUD negative subjects at Site 9, and the specificity estimate was 100% in Tier 4. No prescription documentation were available at Sites 10 and 11. The overall sensitivity in Tier 4 was 70.7%, as compared to 82.8% in the total clinical study population. Specificity was 84.8%, as compared to 79.2% in the total clinical study population. Performance of the AvertD in subjects who had prescription documentation to support the self-reported index exposure decreased in comparison to the overall performance in the total study population.

Tier 3 included subjects who had a description of a prescription with or without documentation available in the medical record or medical history. 83.5% of the final clinical study population had information in Tier 3. The sensitivity in Tier 3 was 82.5%, as compared to 82.8%, and the specificity in Tier 3 was 79.6%, as compared to 79.2% in the total clinical study population. Performance in Tier 3 is similar to the performance observed in the overall clinical study population.

Tier 2 includes subjects who had a description of a procedure or event that may relate to oral opioid prescription. 94.75% of the final clinical study population had information in Tier 2. The sensitivity in Tier 2 was 82.7%, as compared to 82.8%, and the specificity in Tier 2 was 78.9%, as compared to 79.2% in the total clinical study population. Performance in Tier 2 is similar to performance observed in the overall clinical study population.

The last form used to collect information on the clinical study subjects after the study was completed was Form 3. Form 3 was used to collect information on the comorbidities of each subject. The AvertD detects 15 snips that have been associated with OUD as well as with a number of other substance use and mental health disorders. Because the SNPs are associated with other disorders as well, it raised uncertainty as to whether the device detected genetic risk of other comorbidities present in the clinical study population.

Form 3 was used to collect information on the comorbidities of each subject available in the medical records and histories. There was comorbidity information available for 97.9%, or 377 out of the 385 subjects. It is of note that the medical records and medical histories available at enrollment sites may not capture all of the comorbidities that the subject had, and some information may be missing for some subjects. Based on the company's analyses, there were no

clear differences in rates of comorbidities in the clinical study population compared to the US population. There were also no differences in the rate of comorbidities in the subjects at the time of self-reported index exposure compared to the time of enrollment. With consideration of the study design and execution, the overall sensitivity and specificity was calculated for the study population. The sensitivity was 82.8%, with a confidence interval of 76.3% to 88.1%, and specificity was 79.2%, with a confidence interval of 73.1% to 84.5%.

Additionally, based on the date of self-reported index exposure and the age of each subject, 85 of the 381 subjects in the clinical study analyses were prescribed their first oral opioid for treatment of acute pain prior to the age of 18. The intended use population is subjects 18 years or older who may be receiving their first oral opioid prescription. The number of subjects in the clinical study that were older than 18 years old at the time of self-reporting index exposure was 296. 121 of these subjects were OUD positive, and 175 were OUD negative. The sensitivity of the AvertD in the clinical study population who was older than 18 years old at the time of self-reported index exposure is 84.3%, as compared to 82.8%. The specificity was 78.3%, as compared to 79.2%. Performance estimates were similar.

In summary, there are several factors that contribute to the uncertainty in whether the observed clinical study results accurately represent the device's performance in the intended use population for the test. These include: uncertainty in how the study was designed, including the enrichment strategy of enrolling subjects from opioid treatment program site and grouping subjects into high and low risk pools based on evidence of an SUD or OUD prior to selection of the subjects for inclusion in the study; uncertainty in the study population due to use of inconsistent inclusion and exclusion criteria; use of subject recall to determine index exposure as

well as the enrollment of subjects who had a previous relationship with the enroller; uncertainty in the device design due to use of SNPs that are associated with comorbidities that are common in OUD positive subjects; and uncertainty in the clinical performance.

In conclusion, manufacturers may contribute to OUD risk. A test demonstrating that probable benefits, outweigh probable risks could have significant public health benefits. There are limitations regarding the clinical study design, study population, device design, and clinical performance. FDA faces challenges in making a benefit risk determination. FDA is therefore seeking expert opinions from our advisory committee.

Our questions to the panel can be summarized as one overarching question. Does the clinical study population adequately represent the intended use population such that the estimates derived from the clinical study are representative of the expected performance of the device when it is marketed and used in the intended use population?

We will have questions in the question and answer portion that will touch on: the impact of factors that contribute to uncertainty, such as use of different case report forms, confidence with which certain populations were excluded, index exposure based on subject recollection, recruitment sites, risk pool assignment, demographic makeup of the study population; device design and association of the 15 snips with other SUDs and disorders; clinical performance such as base sensitivity and specificity; benefits and risks of genetic testing to assess risk of developing OUD; clinical use of AvertD; and labeling mitigations that may minimize risk. Thank you for your attention.

Dr. Watson: Thank you so much to the FDA for that presentation. We will look for some clarifying questions that the panel has, but first, we received two additional questions from

the panelist after the Executive Summary was sent out. FDA, would you like to address those questions first before we open it up to the panel here?

Dr. Gussow: Yes. Thank you. I'll go ahead and address these two questions and a couple of questions that were raised earlier today. So, the first question we got was: were performance goals of 59.5% for sensitivity and 55.5% for specificity cited by the sponsor, pre-specified by, or agreed upon by the FDA, as the minimum threshold for the device to be clinically useful? Our response is: the company proposed the stated performance goals prior to the submission of the de novo classification request. FDA did not have sufficient information at the time to assess the proposed goals. FDA is now seeking the panel's input on the clinical significance of the study, results in performance, including the sensitivity and specificity estimates.

The second question we received was: does the FDA have concerns about the study design and what type of patients were enrolled by the inclusion criteria of retrospective self-reported opioid exposure? It would seem that patients who accurately and reliably recall their first prescription opioid exposure may have had an exceptionally positive or negative experience with their first opioid use. Thus, the study is enriching for subjects on two extremes of the spectrum and perhaps not capturing the majority of the population who do not have a notable experience and thus would not be able to fulfill the enrollment criteria for the study, as they would need to recall with some fidelity the approximate timing and circumstances of their opioid exposure. Our response is FDA is seeking the panel's input on questions along these lines. That is, the study design and the representativeness of the clinical study population of the intended use population.

We were also asked a question earlier this morning about FDA's thinking on polygenic risk scores. Our response is, we do not have cleared or authorized devices of this type as yet. We have authorized devices that measure multiple analytes by immunoassay and then input them into an algorithm to get a score. One such example is the Ova1 test. FDA doesn't have one expectation for performance because each test is considered for its intended use, and each intended use is considered as described. Generally, whether a test measures a single analyte or combines the result of several into one score, we consider what performance would be safe and effective for the proposed intended use. Acceptable performance depends on what the test is intended for. Acceptable sensitivity and specificity could depend on several factors such as if the test is for screening, monitoring, diagnosis, et cetera, and whether there are other mitigating factors. Examples of mitigating factors that we consider when we assess sensitivity and specificity include whether confirmatory testing is typically performed, or whether that test would be adjunctive and considered along with other tests as part of a clinical decision making. Alternatively, we may expect higher sensitivity and specificity if a test is intended as a standalone test that would be used to determine that a patient needs surgery or some other high risk procedure. We also consider the patient population and perspectives. This is not an exhaustive list of course, but just examples.

And then finally, a question was raised about whether AvertD would be a part of a REMS program. A Risk Evaluation and Mitigation Strategy, or REMS, is a drug safety program that FDA can require for certain medications with serious safety concerns to help ensure the benefits of the medication outweighs risk. While it's an important program, today, we're asking the panel to focus and provide feedback on the safety and effectiveness of AvertD for proposed indication. We certainly welcome any suggestions from the panel, but we'll not be asking the panel specific

questions about REMS. Thank you. Those were the questions we were prepared to answer, and now we can open to other questions.

Dr. Watson: Thank you so much. All right. I see some hands from the panel. Dr. Walker, do you have a question?

Dr. Walker: I do. On the slide where you were listing sources of uncertainty, it wasn't clear to me whether you were taking the possibility that the predictive capacity was influenced by genetic ancestry and that, if that confounder was taken into account, perhaps the predictive capability would be lower. Was that one of the uncertainties that was there and I just missed it?

Dr. Gussow: So, on the slide that I believe you're referencing, we didn't specifically point that out. We'd like the panel to consider these types of things even if we haven't specifically pointed it out.

Dr. Walker: Okay. Well, I know this will be the source of the public comment, so I'll reserve any additional comments until after that time. Thank you.

Dr. Watson: Thank you. Dr. Goldstein.

Dr. Goldstein: Two questions, please. First, was FDA able to determine the frequency or number of patients that declined to be tested genetically in the — I guess they're not really part of the study population in a sense, but they are important in the analysis, I think. And then the second question is, does FDA approval — is that likely to lead to physicians requiring genetic testing before prescribing opioids for patients with acute pain?

Dr. Gussow: I'll address the first question. We do not have information on the frequency or number of patients that declined. Our understanding is that all of the patients who are enrolled met inclusion/exclusion criteria. We might return to this question and have the sponsor weigh in on this, but we do not have this information. The FDA does not.

And then for your second question, does FDA approval lead to physicians requiring it? I'll ask Dr. Kelm to weigh in here. I don't believe we can comment on this.

Dr. Kelm: Good morning, and sorry for the noise coming back to everybody from my mic. I did want to answer the first question. I believe we had previously asked a question about people who had been approached and had declined. And I believe we were informed so that they didn't have that information and that everybody approached had enrolled. But I do think that's something that we can go back to sponsor, because it's been a few years.

This obviously is an assessment today getting your input on the device. And you had the proposed intended use, which talks about the purpose of the test. Obviously, then, if we were to authorize it, it would be up to physicians to decide whether or not to order it and use it how however they want in their practice of medicine. I don't believe, and there were no conversations, for example, with our colleagues on the drug side, that there would be at least — that was not part of our conversation that there would be any impact on, for example, the actual drugs themselves. And, obviously, whether or not there was any action taken upon this being available by any other group in terms of testing and tying that for being able to receive, for example, a prescription. That's also not something that that we typically would comment on.

Dr. Watson: Thank you Dr. Kelm. All right. Walter Dunn.

Dr. Dunn: Two clarifying questions. So, my understanding is that the label for this intended use is for the first opiate prescription for acute pain. So, if a patient receives that prescription and there's subsequent medical events that require opiate use, any use of this test would be considered off label. Is that correct? That's number one.

And then number two: in your initial briefing document, you had described most of the research sites, or most of the patients being recruited, from part of their typical or standard clinical care. But looking at the names of the sites, most of them actually had some research element in their description. So, it sounds like a lot of these research sites also provide clinical care, but there's some focus on actually recruiting patients for studies. Thank you.

Dr. Gussow: Thank you for your question. Questions, addressing the first: yes. The proposed intended use is for the first prescription for acute pain, and yes, it is our understanding that any use other than that would be off label. And for your second regarding the types of sites, we have limited information about the treatments offered at these sites, if there are treatments offered at these sites. Yes, a lot of them have research in the name. We might also want to return to this question and have the, the sponsor weigh in. What we do know is that two of the sites were opioid treatment program sites. These are listed on SAMHSA's website, and presumably these are sites where people are seeking treatment for Opioid Use Disorder or others. The remaining seven out of eight, we are not clear on what type of treatment or research is done at those sites, and we know that one 1 of 10 is a solely research site, and only seven people were recruited from that site.

Dr. Watson: Thank you, Dr. Farrar.

Dr. Farrar: Yeah, I wanted to ask if the FDA had calculated positive predictive values for the estimated prevalence of the disorder in the group, and whether it uses positive predictive value in its assessment of the potential benefits of particular diagnostic tests. Thank you.

Dr. Gussow: Right, thank you for that question. I apologize that we don't have our statistician here, but we did do some calculation of the PPV. The sponsor had proposed several different PPV the calculations with different estimates of prevalence. We didn't formally look at PPV. We spent a little bit more time looking at sensitivity and specificity. Whether we use this calculation in our assessment, like I said, we, we looked at sensitivity and specificity more closely. In some devices we look more closely at PPV and NPV, but I think one of the reasons why we didn't look at PPV as heavily in this case was because the incidents of OUD isn't exactly clear. It was brought up earlier that it's about a 1% incident in the general population, but 5% is calculated. I hope that answered the question. Dr. Kelm, would you like to add something?

Dr. Kelm: Yes, I would like to a add a little bit more. Because of the study design and the enrichment — so the prevalence, obviously, in the study would not reflect the prevalence in the intended use population. And I think we found that it was hard to get a good estimate, at the time, of what the prevalence of OUD was. And we found some varying numbers depending on what publication. So, in this case, based on that information, we often will just use sensitivity and specificity when looking at the performance of the test and thought that was appropriate here. I will say that for most diagnostic tests where based on study design as well as information on prevalence and studies reflecting prevalence where they best can, is that we do prefer to use positive and negative predictive value when providing information about tests and using that to evaluate tests. I hope that's helpful.

Dr. Farrar: Yeah. If I might just follow up, I might recommend looking at — the sponsor indicated that they felt that the rate was 1% in the population and 5% potentially in the operative group. And you know, just generating positive predictive values for 1, 5, and 10 all demonstrate that the

likelihood of a positive prediction being well below 25%. I think that that's an important topic for us to discuss later, since you're going to be wrong, even in the best of situations, you're going to be wrong in four out of five of the people that have a positive diagnosis. So, I think that is very key in terms of the estimate of risk versus benefit. Just a comment. Thank you.

Dr. Watson: Thank you for bringing that up. We will definitely discuss that during panel deliberations. Dr. Bateman.

Dr. Bateman: Thank you. I'm wondering if the FDA has a perspective on the concerns raised by Halpman (phonetic) and colleagues in their submission to the panel as well as in their article in Drug and Alcohol Dependence in 2021, suggesting that the test is highly confounded by genetic ancestry, and if genetic ancestry is balanced, that these SNPs, these single nucleotide polymorphisms that constitute the test, don't predict risk any better than chance alone. It to me it seems like a really major, major issue regarding the validity of the test and, and whether it has clinical utility. I'm just wondering if FDA has looked at that issue and what their thoughts are.

Dr. Gussow: Thank you for raising that question. I think these are along the lines of the types of thoughts we had regarding uncertainty in the data we are looking at, and we're hoping that the panel can weigh in on questions like this and deliberate on things like this.

Dr. Watson: Thank you very much. This has been a fantastic session. We actually got ourselves back on track, so I would like to suggest that we break for lunch now. We are scheduled to be back at 1:00, and I think we can do that. Is everybody okay if we just have a 24 minute lunch break? Okay. So, we can get right back on track. So, we'll break for lunch now.

Dr. Compton: 1:00 will not work for me. I need a little longer to run out.

Dr. Watson: Okay. That's fair. 1:15?

Dr. Compton: That'll work.

Dr. Watson: Okay. So, Panel members, please do not discuss the meeting topic during lunch amongst yourselves or with any member of the audience. We'll reconvene at 1:15. Thank you all.

Dr. Watson: Welcome back, Panel. It's now 1:15, and we're going to resume this Panel meeting. We'll proceed with the Open Public Hearing portion of the meeting. Public attendees are given the opportunity to address the panel to present data information reviews relevant to the meeting agenda. James Swink will read the Open Public Hearing Disclosure Process Statement. James?

James Swink: Both the Food and Drug Administration and the public believe in a transparent process for information gathering and decision making to ensure such transparency. During this Open Public Hearing session of the Advisory Committee Meeting, FDA believes that it is important to understand the context of an individual's presentation. For this reason, FDA encourages you, the Open Public Hearing speaker, at the beginning of your written or oral statement, to advise the committee of any financial relationships that you may have with any company or group that may be affected by the topic of this meeting. For example, this financial information may include a company's or a group's payment of your travel, lodging, or other expenses in connection with your attendance at this meeting. Likewise, FDA encourages you at the beginning of your statement to advise the committee if you do not have any such financial relationships. If you choose not to address this issue of financial relationships at the beginning of your statement, it will not preclude you from speaking. Thank you.

Dr. Watson: Thank you so much, Mr. Swink. The FDA has received 19 requests to speak prior to the final date published in the Federal Register. Two of the speakers are live, and the rest are recorded. The first live speaker is Andrew Kolodny. So, Mr. Kolodny, please begin.

Dr. Kolodny: My name is Dr. Andrew Kolodny. I'm the Medical Director for the Opioid Policy Research Collaborative at Brandeis University, and I'm President of Physicians for Responsible Opioid Prescribing. My clinical work over the past 20 years has focused on the treatment of Opioid Use Disorder. This slide may look familiar, this is a CDC slide that was shown for many years. The chief speaking point from the CDC in using this slide was that the medical community needed to prescribe opioids more cautiously. The green line shows the increase in opioid prescribing, the red line, prescription opioid related deaths, and the blue line, prescription opioid related addiction. The message from the CDC was that the green light needed to come down in order for the red line and the blue line to come down. In the early 2000s, after these trends became evident, opioid manufacturers began to change their messages a bit. Instead of citing Porter and Jick (phonetic) and saying, "Nobody gets addicted to opioids," the messaging became what was termed "the new paradigm." The idea being that prescribing should continue to go up, but doctors should use risk assessment tools, like the opioid risk tool, to identify risky patients and monitor those patients more closely. And, in doing so, we could have our cake and needed too, that the prescribing could continue to go up, but we wouldn't see these increases in adverse public health outcomes.

Unfortunately, that didn't work. And you can see that as the prescribing went up, we did continue to see deaths and addiction go up. And this didn't work, in part, because it was based on a flawed assumption. The assumption being that the problem with opioids is that there are some risky patients out there, that there's a subset of our population that's prone to addiction, and that we can identify them. And if we're more careful with them, we'll have good outcomes. That doesn't work. The problem with opioids is not that there are some risky patients, it's that the drugs themselves are inherently addictive. One of the most popular tools that's been promoted

was the opioid risk tool, which in a recent, AHRQ review published in 2020, it was shown that it doesn't work. It doesn't have the sensitivity and specificity. A tool that can allow us to predict in advance who will or will not become addicted to opioids does not exist, and SOLVD has not invented one.

I would point the committee to a letter that's in their materials submitted to the docket by a group of genetics professors, including Joel Grolenter (phonetic). Their letter has pointed out that this notion that AvertD can predict addiction is impossible. And they've pointed to a paper they've published in Drug and Alcohol Dependence on this very test showing that it doesn't work. And I would urge the committee to take a look at both their paper and their letter. One of the statistics that was used repeatedly in the sponsor's presentation was that, with addiction, there's a 50% genetic risk. That was presented over and over again, as if, with all addictions, the genetic risk is the same. That's not true. Opioids are not like alcohol. With alcohol, most of the population can be repeatedly exposed and doesn't become addicted. It's a subset of the population that becomes addicted to alcohol, and therefore genetics are likely to play a very significant role. Opioids are more like nicotine. With highly addictive drugs. The genetics begin to play much less of a role, and it's the exposure to the drug that plays a more significant role. There was a point in time when more than 50% of the adult men in the United States were addicted to nicotine.

This is a paper that was published just two days ago. it's a paper comparing NSAIDs to opioids in emergency rooms in patients with severe pain. And in this systematic review and meta-analysis, they found that opioids were not better than NSAIDs for treatment of severe pain in an emergency room setting. This adds to a growing body of literature showing that, even for

severe acute pain, NSAIDs are often as effective, if not more effective, with less adverse effects than opioids. We can be reducing our exposure of patients to opioids.

Lastly, this is a slide that comes from a paper published a couple of years ago, comparing use of opioids postoperatively in the United States to other countries. This was a site that looked at appendectomy, colosectomy, and hernia repair. In all of these different countries, they compared to the United States. In the United States, the median number of pills when patients were discharged after these procedures was 20. In all of these other countries, including the Netherlands, the median number of pills was zero, because patients are not routinely sent home from an emergency room or postoperatively with opioids in these other countries. NSAIDs are used instead.

Preventing Opioid Use Disorder does not require a genetic test. It requires more cautious prescribing. We should avoid opioids when we can, and we often can. And when we can't, we should prescribe the lowest possible dose for the shortest duration possible.

Last point I'd like to make is for the committee to consider what happens when you create a financial incentive for a prescriber to do a genetic test in order to prescribe an opioid. What happens when that prescriber gets more money because they're ordering this test and prescribing an opioid than if that prescriber had prescribed an NSAID instead? That will very likely contribute to increases in opioid prescribing. And if you look at this slide, you'll see, and as you've already heard, it takes 24 to 48 hours for the results to come in. According to the CDC, for most patients with acute pain, three days are usually more than enough. So, are these prescribers going to prescribe an excess duration of opioids in order to be able to order this test?

I would be very concerned about the financial incentives that would be created by this product if it was actually approved. Thank you.

Dr. Watson: Thank you so much for your comments. Dr. Kolodny. Michael Abrams, are you present? You have the floor now.

Dr. Abrams: Thank you. Good afternoon, everyone. I'm Dr. Michael Abrams, a Senior Health Researcher with Public Citizen's Health Research Group, and I have no financial conflicts of interest on this matter. We oppose FDA marketing authorization of the AvertD test for identifying patients at increased genetic risk of Opioid Use Disorder prior to the use of opioids for acute pain, and we do so because we share concerns, some of which were expressed in the FDA summary document for this meeting, and I'm going to quote directly from that document. FDA said, "Numerous factors impact the interpretation of test performance and raise uncertainty about the applicability of the observed clinical trial study test results presented by SOLVD." We agree, for example, that there are concerns about the standardization of data collection, the reliance on years past self-reports, and medical record abstractions. We also note, as others have today that the study population is racially quite homogenous, right? 92% white.

However, our main concern about this test's effectiveness is its reliance on flawed predictive modeling. Flawed because of the limits of genetics to predict OUD risk, in no small part because of confounds related to ancestry. And these concerns, as Dr. Bateman brought up initially in this meeting today, were summarized in comments submitted to this committee by the researchers Alexander Hatoum and colleagues, who stated the following, and I'm going to quote directly from them. "Current knowledge about OUD genetics is strong enough for us to be clear that predicting a patient's risk of OUD using AvertD or any test like that is impossible," with an

IM at the beginning of that. These researchers further noted that, quote, "Even a full genome's worth of markers, roughly 6 million SNPs, are not sufficient to predict OUD in a clinically useful way." Close quote. Now, notably, Hatoum et. al directly analyzed an earlier, similar version of AvertD, with 16 markers instead of 15 markers that you are considering today, and they found that only one of the 15 markers were supported by current well-powered gene discovery studies. None of the five algorithms used for AvertD predicted OUD any better than chance. Once ancestry of the subjects was balanced, the markers showed large differences in allele frequency in different populations. And due to these differences, the algorithms deployed predicted race and ethnicity, rather than Opioid Use Disorder. Accordingly, these researchers concluded that tests like the AvertD, and I'm quoting again from their document they submitted to you, quote, "are not only of no predictive utility and give false sense of confidence, they could lead to widespread harm by biasing the decisions about the treatment of pain."

So, in conclusion, we agree with that assessment, and we encourage the committee to vote no on the voting question one before you today. And we urge the FDA to deny the request to obtain de novo marketing authorization for AvertD, because the device's benefits are poorly established in this case and the overall device suggestion here is scientifically implausible. Thank you very much.

Dr. Watson: Thank you, Dr. Abrams. We will now play the prerecorded comments from the public.

Dr. Suzette McKinney: Good afternoon. My name is Dr. Suzet McKinney, and I am a longtime public health professional. Having spent the past 20 years as a public health executive, practitioner, and advocate. Throughout my time in public health, I have seen the devastating

impacts of the opioid crisis on individuals, families, and communities. I've also witnessed the stress that this crisis has caused to my colleagues in the healthcare profession who are providing direct service to patients. The opportunity for our society and our economy to have a diagnostic tool available that can give providers an understanding of those who are more susceptible to opioid addiction would be a game changer in our fight against this horrific disease.

The benefits are significant. We have been working to resolve the opioid epidemic for over 15 years, and while this test is not a silver bullet, it can help to identify patients who are at risk for addiction. We need new and novel approaches to addressing this epidemic, and genetics are now shown to be a key component of addiction, and we should be looking for ways to use them proactively to identify those who may be at higher risk. Guidelines call for risk assessment prior to prescribing, but currently available tools are limited. An informed decision will always be better than an uninformed decision. And the best part is that it's a simple cheek swab. But this simple cheek swab and the results that it provides can give a physician and their patient the opportunity to decide for themselves if it is appropriate to minimize or eliminate opioids in their treatment regimen. At a minimum, this test creates an opportunity for the discussion and makes the patient and their physician more careful. I hope you will consider this in your deliberations today. Thank you.

Dr. Newland: My name is Brand Newland. I hold a Doctor of Pharmacy degree from the University of Iowa College of Pharmacy, as well as a Pain Management Certification issued through the American Society of Health System Pharmacists. I'm currently pursuing a similar certification in pharmacogenetics. In 2018, I co-founded Gold Finch Health, a company with a mission to help patients understand, demand, and ultimately experience better surgery, faster

recovery, and less exposure to opioids. I believe that the opioid crisis is the number one public health issue of our time. It's true even in 2022 that far too many people receive far too many opioids, especially after surgery. The vast majority of surgical patients receive opioids postoperatively. Nearly 10% become persistent opioid users. Surgery is the number one gateway to long term opioid use, but it doesn't need to be this way.

We have evidence-based approaches today to stop opioid addiction and overdoses much earlier. A pre-therapy risk assessment would be a significant addition to our clinical armamentarium. The benefits are significant. Guidelines call for risk assessment prior to prescribing, but currently available tools are limited. It's known that genetics play a role in addiction. We need to have as much information at the fingertips of both healthcare providers and patients before therapy begins. This is key to creating patient-centric pain management plans that are both safe and effective. An informed decision will always be better than an uninformed decision.

My assessment of the risk of this testing is low. The test entails a noninvasive cheek swab. For the provider of this test, it represents an easy to implement step within workflow and a step that yields meaningful information. The intervention, if the test comes back positive, is simple and low risk itself: minimize or limited opioids by taking advantage of what we know about multimodal pain management approaches. At minimum, the patient, his/her family, and healthcare providers can be more watchful. There is no stigma associated with this. This is patient empowerment in the 21st century. I support review and approval of this testing. It is critical to patient care and a much needed part of the solution to ending the opioid crisis once and for all.

Dr. Hamid: Hi, my name is Kamran Hamid. I'm an orthopedic surgeon, and I practice at Loyola University Medical Center. My job is to take people's pain away. Unfortunately, in order to do that through surgery, we often have to give them a little bit more pain. Orthopedic surgeons are known for unfortunately contributing to the opiate crisis in the United States, and the reason we do this is not because we don't care, but it's specifically because we do care. It's hard to look at somebody who's hurting to not want to help them. There's a saying that pain is a more terrible lord of mankind than even death itself. But it's not just pain that we try to take away. It's also suffering, and there's a difference. Pain is the body's physical response to a no susceptive stimuli. Suffering is our mind's perspective on this, and addiction is suffering.

Addiction is also a complication of surgery that is like anything else that we take care of: postoperative infection, postoperative blood clot, postoperative arthritis. However, for the rest of these things, we prepare before surgery in order to manage them as well as possible after surgery. We ask questions like, “Do you smoke? Do you have diabetes? Do we need a CT scan or an MRI?” And we prepare ourselves as much as possible before surgery to take care of the complication afterwards. Postoperative opiate addiction is no different from this. Unfortunately, up until now, we haven't really had very good tools to address this.

Now, we have an opportunity to do this, as technology has reached where we need it to be. The ideal instrument for predicting whether somebody will need help after surgery should be one that is simply done, that has a simple answer, and is simply actionable. And I think the technology that we're discussing today addresses all of those issues, would be something that I would really love to have in my practice so I can best help my patients manage their pain and suffering. Thank you.

Dr. Fox: Hi, my name is Eric Fox. I'm a board certified oral and maxillofacial surgeon practicing in southeastern Florida and practice full-scope oral maxillofacial surgery. I'm well aware of the risk of opioid addiction, especially in my state of Florida, and try and limit the amount of opioids that are prescribed to all of my patients. I'm keenly aware that third molar wisdom teeth removal, or third molar surgery, is a patient's first exposure to opioid or narcotics. And if we had a test that was safe, rapid, easy to administer and cost-effective, it would identify those patients at higher risk. It would be extremely beneficial for the prescriber, the parent, and the patient. A cheek swab that can be safely performed at the time of initial consultation can be discussed with the parents and patients prior to surgery, helping us either limit or completely eliminate opioid exposure and try and provide a different medication for postoperative pain management. I strongly believe that a tests like this would help alleviate the opioid crisis, and especially in my state of Florida. Thank you.

Dr. Chacon: My name is Guillermo Chacon I'm a board certified oral and maxillofacial surgeon practicing in the state of Washington. I'm an affiliate professor at the University of Washington School of Dentistry, and I'm in private practice in Pierce County. It seems to me that this will become a very valuable tool for those of us who deal with surgery in the early stages of life. As you can imagine, many of our patients that come in to see us in our practices in oral maxillofacial surgery are young teenagers that are being exposed to surgery for the first time in their early teen years for removal of wisdom teeth or removal of bicuspids for orthodontic treatment, or things along those lines. These are individuals that require pain management after we perform these procedures. You can imagine that the way things are with the opioid epidemic, having a tool like this would be extremely valuable because there is really no way for us to know who is at risk of becoming dependent on this substances that we can prescribe for pain

management after surgery. So, if we have a test, it would be an invaluable tool to make sure that we're not going to expose children or young adults who may be at high risk of developing dependence of opioids that we can prescribe and avoid putting them at risk. Now that doesn't mean that this is the only way to prevent this from happening, but it certainly gives us an extra level of safety that we could use in order to minimize the risk of exposing somebody who probably shouldn't be exposed to these substances.

Mr. Jones: Good afternoon. My name is Rich Jones. A little bit about myself: I'm a professional counselor. I've been working in the field of behavioral health and substance use disorder and treatment and recovery services for about 21 years. I'm also a person in long term recovery myself from an Opioid Use Disorder, and I've been in recovery for coming up on 22 years. So, I'm going to talk today about a solution in a way to attack the opioid epidemic, if you will, that is really important to me from a personal and professional level, and from a common sense, pragmatic point of view. I've been looking at this issue for 20 plus years, and what we have seen is a great deal of emphasis on what I would call downstream interventions, identifying people who have a problem, trying to find a way to increase treatment access, try to find a way to increase recovery access. And I want to support that, and I understand that.

But what we really need is, we need to go upstream. We need to find a way to talk to people about the issue of Opioid Use Disorder, about the potential problems associated with opioid us before you reach the point where you've entered into dependency or Opioid Use Disorder severe. It's just simply too hard to have the conversation at that point. It's much more reasonable to go upstream and have the conversation there. Well, to me, this solution solves that problem, because it's able to help us identify people that are at high risk, people that have a

special concern around this. And I know. I was one of those folks. At that time, I was not working in the industry. I didn't know anything about it. If somebody would've approached me and said, “We ran this test, Rich, we did this evaluation, we took this cheek swab, and we looked at what the results are. You are at higher risk for opioid dependency for the following.” I, for every part of me sitting here, I absolutely believe I would've listened to that. Instead, what happens is, we go on to the bitter ends. We go on until we quote unquote “hit bottom” and we have to listen. So, I would ask, let's consider solutions that get there earlier. Let's consider ways to start the conversation. This is a perfect example. There are many things that could be added to that. But it has to start with identification, and it has to start with awareness. And this product is all about identification and awareness. Thank you very much for your time. We'll talk soon.

Mr. Kaufman: Hi there. My name is Ken Kaufman. I'm the co-chair of the Substance Free Coalition of Northwest Michigan, but I come to you today speaking as an individual, and I'd like to start with a little bit of background about my family and our journey with prescription opioids and why I think prevention is so important. In 2007 when our son, Brian, first came to us as a junior in high school, and said he was addicted to prescription opioids, we knew nothing about them. We didn't know what Vicodin was, we didn't know what Oxycontin was. We put our son into a rehab facility, and he was out within a few days, and clean, so we thought. A year later he came home from college during welcome week and said he was now addict addicted to Oxycontin. We put our son into a rehab facility for a longer period of time. It wasn't long though, a year or so later, and eventually he turned again to opioids, and that eventually led to heroin. Our story ends tragically when Brian overdosed and died in a hotel room on April 23rd, 2016.

Since 2016, I've devoted my time to prevention. This is where I can see the benefits of having a test like AvertD. When our son first had his wisdom teeth removed and was prescribed Vicodin, we would never have allowed him to use them had we known his propensity for addiction. I can see the benefits of knowing the risks in advance, and although there is currently no way to accurately predict who is and who is not at risk of becoming addicted when taking prescription opioids, it is known that genetics play a role. And we need to have whatever information we can access to determine if someone in someone is genetically disposed before they are exposed. Any informed decision will always be better than an uninformed decision. The risks of a genetic test seem very, very low. It's just the cheek swab, and it's a non-invasive test and takes little time. As I understand the genetic test, if you are found to be a high risk, then just minimize or eliminate opioids. There are alternatives to opioids to manage pain. Knowing if you are at risk or in danger has to be beneficial and much in a much better alternative than going through what our son, Brian, and our family did once he became addicted. Thank you for your time.

Ms. Barber: My name's Jodi Barber. I'm from Laguna Niguel, California. On January 8th, 2010, I lost my 19 year old son and three of his close friends in the same year to an overdose. They quickly became addicted to opioids that were prescribed by a doctor. I soon heard of thousands more dying. A beautiful young 17-year-old high school football star was injured, and the doctor prescribed him Norcos. He soon began his addiction, which led to more pills, and later, sadly, he died. He asked me, “Jody, why didn't the doctor ask me if addiction runs in his family?” Both his parents were addicted to prescription drugs. Sadly, I personally know parents who lost more than one child and other family members to opioid addiction. Last year alone, 107,000 Americans overdosed and died. Most of these deaths were from fentanyl

poisoning. Many who died began their addiction from opioids prescribed by a doctor. The opioids were expensive. So many turned to heroin on the streets, which led to fentanyl. These can become preventable deaths if we can identify people at high risk for Opioid Use Disorder and keep new addictions from happening.

Opioid patients need to better understand their risk prior to taking an oral opioids for acute pain management. For a genetically predisposed high risk individual, they can be headed towards Opioid Use Disorder in as little as three days after taking that first pill. New technologies that provide objective risk assessment are in front of the FDA right now for approval, and there is an immediate need for them to become part of any long term solution. The 'SUP' in support stands for Substance Use disorder Prevention, and we must make sure that the funds from the Support Act are being used for that purpose. Identifying high risk patients and closely monitoring. Prescribing practices will most definitely reduce addiction and overdoses. Narcan will save some lives by reversing an overdose, but unfortunately, you have to overdose. It is important that we all work together to stop addiction before it starts. I feel this type of prevention will keep other families from going through the devastating loss that my family and thousands of people are experience experiencing every day. I'm asking you to please authorize AvertD, the first of its kind genetic risk assessment for Opioid Use Disorder to market. Thank you.

Mr. Daniels: My name is Ken Daniels, television Voice of the Detroit Red Wings, but more importantly, I'm the father of Jamie Daniels, whom we lost to a fentanyl-induced overdose December 7th, 2016. He was just 23. If only we knew a decade ago what we know now. If only we knew then the harm and tragedy that would take place. If I were a young parent today, I'd

want my child tested to see if there was any risk of addiction. That is why we tell Janie's story to Congress and to public forums. We must end the shame and stigma that surrounds substance use disorder, this mental illness. There are alternatives to opioids to manage pain. Let's just figure out who should use them. Patients in concert with their physicians can make an informed decision. Again, if only we knew then what we know now. A test being available to gauge opioid susceptibility. 90% of those 20 million plus Americans who suffer from substance use disorder began using opioids prior to the age of 18 and many from an unnecessary prescription. If 4 million young adults have their wisdom teeth pulled annually and prescribed opioids, 50,000 of those will have a new, persistent opioid use from what is basically an unnecessary prescription.

We started the Jamie Daniels Foundation in 2018 and have raised more than 1.2 million dollars to help end the shame and stigma surrounding substance use disorder. For those who hear of us, we say empathy is the highest form of knowledge. Why? Because it comes without judgment. But knowledge to keep you from ever having to deal with it? Priceless. Thank you.

Ms. Barry: Hi, my name is Megan Barry. I'm a mother who lost her only child to an overdose. His name was Max, and he was 22. Since we lost Max in 2017, I now deeply understand the heartache and grief that new families and loved ones are experiencing every day when they lose someone they love, because this epidemic hasn't subsided. Why? Because we haven't done enough. And I have come to appreciate so much more about the disease of substance use disorder and the myriad of things that we can and should be doing to treat this like the health crisis that's become. And like a lot of parents, I ask, “What can I have done different? What could I have done better?” There is no answer, but I do know I wish I'd had more resources or tools, more of an understanding on how I could have helped. Like every good parent, I wish

we had talked about this. I wish we had asked for that. I wish we had known Max's potential risks. When he was little, we took him to the doctor for his annual well baby checkup. We measured his growth. When we checked the boxes of his developmental timeline, we had a vague knowledge of childhood diseases, but we never, ever, ever talked about substance use disorder, any potential risk that Max might have had. What if we had? What if we had been given to evaluate Max's propensity for addiction? Would it have helped. Would it have saved him? I don't know. But every parent wants to give their child a fighting chance when it comes to fighting disease.

Ms. Kannaka: Hi, my name is Whitney Kannaka, and I'm a single mom to four young girls. Prescription opioid addiction destroyed my family, and this is my story. It started with low back pain. He went to see his GP, who started him with steroid injections and physical therapy. He continued to suffer from pain that was moving up his spine, and he could hardly walk. It was impacting his quality of life, and he needed to work. His doctor prescribed him Tylenol 3 with codeine, and that is where our lives were changed forever. He went to various doctors over the course of three years and little did I know he was shopping doctors, and he was getting prescribed various opioids, including Vicodin and Oxycodone. It wasn't until he began acting erratic and manic, not sleeping, and lying about family finances that I clued into his addiction. He was sly and deceptive, and after an attempted intervention went awry, he was arrested for a DUI. Upon the release of his time in court, he hid his guns and destroyed the family home, and ultimately was put under an order of protection. This is when I discovered, through phone records, that he had been buying Oxycontin from a street dealer in the suburb not too far from our house. It was for more than two years he was doing that. I know now that this is typical

behavior once oral opioids hijack someone's brain. I moved my girls, and I gave him an opportunity to save our marriage and our family. Sadly, he chose pills instead.

Since then, I have been grieving the loss of a living person. My daughters don't know their real dad, the one before prescription opioids. He remains paranoid and untrustworthy, a marginal father at best, and we have no relationship. When I learned about the newest research and testing that could be available for patients who need significant pain management, I knew my story had to be heard. If one non-invasive cheek swab can identify even the slightest potential for addiction prior to being prescribed opioids, it can save multiple lives, not just one. The hell my girls have been through is a tragedy they should not have had to endure. This didn't just impact their father. It will impact our whole family and most of our relatives for our entire lives. This research and presurgical test would have provided key information to assist in making an educated decision about my ex-husband's treatment and his pain management plan. Our family could still be together, not a statistic. The jury is out on if he will end up an opioid death count in here in Illinois. I pray every day that my girls won't lose their dad. Fighting opioid addiction isn't a simple task, as we all know. We need to use all the tactics we have at our disposal to prevent abuse. Why would this not be another weapon in our battle gear? We can't win without tools. This one is a no brainer. The time to act is now. Thanks for listening.

Mr. Fox: Hello, my name's Chris Fox. I'm the Executive Director of a Washington, D.C.-based coalition called Voices for Non-Opioid Choices. Voices of our over a hundred members have come together under the shared belief that we can do more to prevent addiction before it starts, particularly for short-term, episodic, acute pain patients. In short, roughly 4 million patients will go on to misuse opioids following an acute pain incident. This is a travesty that

doesn't have to take place. We know that there is the widespread availability of FDA approved, safe and effective non-opioid approaches that can help these patients manage their pain without the need for prescription narcotics. Unfortunately, to date, we have not yet made available these therapies to too many perceived providers, patients, and their families and loved ones. What's more is, despite repeated efforts from voices and our allies, we continue to meet with resistance when we try to make sure that these patients get access to these therapies so that we can, together, work to prevent opioid addiction in this nation.

We know we're at a crisis. There are 81,000 Americans that died last year from an opioid related drug overdose. When we launched Voices in 2019, we were losing 115 Americans every day to an opioid-related drug overdose. Well, now, we lose 220 Americans every day to an opioid-related drug overdose, and SAMHSA data tells us that 80% of those who now abuse or are taking heroin initiated their habit via prescription opioids. We need novel approaches, and we need better opportunities to identify those patients who may be at risk for misusing or becoming addicted to opioids or developing an Opioid Use Disorder. That's why we are in favor of any opportunity that we can work with manufacturers or industry or others to make sure that those patients who may have a predisposition to a substance use disorder get special access to non-opioid based therapies. That's why we're here today to let you know that if there is an opportunity to identify these patients, to better understand who of those 4 million Americans we can save, that we do that. Thank you for your time and thank you for allowing me to be here.

Mr. Janasek: Good afternoon. My name is Joe Janasek. I have been a licensed physical therapist for the last seven years and currently act as Regional Manager for Revolution Physical Therapy Weight Loss in Chicago, Illinois. It is a pleasure to speak with you all today and to shed

light on how this risk assessment screen, AvertD, has tremendous benefits to assist patient care from the PT perspective. As we know, musculoskeletal conditions are the most common diagnosis associated within opioid prescriptions. With that being said, physical therapists are considered movement experts who treat patients with musculoskeletal conditions day in and day out. AvertD demonstrates significant application in our profession to provide education and information to our patients who are preoperative, postoperative, or even struggling with long term opioid use.

As physical therapists, it's in our DNA to empower patients with the tools required to live a fully functional and pain-free life. We do this through intervention in the clinic, but more importantly, through the education we instill in our patients for long-term successes. In my opinion, AvertD is an essential piece of the puzzle for patients to more individually address their conservative, preoperative, and postoperative plan of care. This risk assessment will put the power back into the patient's hands and give them the opportunity to make a justified and confident decision if prescribed with opioid for pain relief. In the physical therapy world, we use risk assessments to determine many of variables that put our patients in danger. For example, dynamic gate indexes, functional gate assessments, timed up and go measures. These all determine if the patient is at risk for falls in the community or at home. If they fail the examination, we provide them with the education and tools to reduce the risk of an accident while in the community. AvertD is no different. It is a clinically validated genetic test that helps patients and their providers better understand their genetic risk for developing Opioid Use Disorder. From their test results, we then can ethically and morally provide our patients with a more encompassing outlook on their medical pain management process.

Now, physical therapy is just one alternative to combat the opioid epidemic, and with the continued diligence of the APTA advocating for therapists and patients alike, we have made tremendous steps in the right direction. I encourage all of you to continue taking the correct steps forward to returning our healthcare model to a more patient-centered approach where we, again, empower our patients with sound decision making tools that will affect their physical, mental and emotional wellbeing. Thank you.

Mr. Beyer: Hi, my name is Cal Beyer. I am speaking on my own behalf. I'm calling from the suburb of Seattle Maple Valley, Washington. I am the father of five children aged 30 to 21. I was born into a family with multi-generational substance use disorder involving both alcohol and opioid prescription pain medication. Having the ability to test for a genetic predisposition for Opioid Use Disorder will make a difference for families, employer health plans, and individuals like me. Due to a knee injury requiring surgery when I was in fifth grade, I started taking opioid pain medication. With re-aggravation of the injury and additional surgery, I was prescribed additional pain medication when I was in high school. I experienced persistent use. This included misuse by taking more pills than were prescribed. It included abuse through diversion by topping off my pill bottle by siphoning off a parent's prescription supply. During a two year pity party, I started misusing pain medication to numb my feelings of sadness. It compounded my problems. I experienced withdrawal and dark thoughts. I even considered suicide. Fortunately, these pills were only Tylenol #3 and Percocet, and not Vicodin or Oxycontin. The nightmare of fentanyl did not exist then either. I feel blessed.

I have 35 years of professional work experience in risk management, insurance, and safety, health, and wellness. I've consulted to industry in various sectors, including public

entities, healthcare systems, construction, manufacturing, and service industries. As risk manager for a large heavy construction firm, I helped launch the Mental Health and Suicide Prevention initiative in the construction industry in 2014. I've been doing presentations and webinars on opioid risk reduction for various construction alliances and associations. I title it "Waging a counterattack against opioids in the workplace and at home." The construction industry continues to battle opioids and persistent use from two leading gateways, prescriptions for muscular skeletal injuries, the National Council of Compensation Insurance and the National Safety Council report 55% of workers receive opioid prescription. That's twice what it is for any other industry, and doses are 20% stronger for 20% longer duration. And the second gateway is surgery. In the 2018 Plan Against Pain, research showed an average of 8.7% of patients having surgery resulted with persistent opioid use. Approving an easy to use swab device to detect the genetic predisposition of Opioid Use Disorder would be a major step forward to help prevent persistent opioid use and addiction, especially for opioid naive patients. The FDA knows the scourge of opioids in our country for well over 20 years. Lives of innocent Americans can be saved by the approval of a device like we're talking about today. Thank you very much for your time.

Mr. Large: Thank you for allowing me a couple minutes to speak, as this is an incredibly important issue to me. My name is Brett Large. I'm the business manager with Heat and Frost Insulators Local 19, located in Milwaukee, Wisconsin. I'm also vice president of the building trades here. The building trades here in Wisconsin represent approximately 30,000 union construction workers and many more non-union, as well. My job is to protect my members and to provide healthy, safe, and productive workforce for my employers. As you may or may not know, construction workers report almost twice the national average of Opioid Use Disorder

and are six times more likely to fatally OD on opioids than any other industry. Why? Construction sites are dangerous places. It's incredibly physical work. We abuse our bodies, experience a variety of injuries from worn out joints to broken bones, and et cetera, et cetera. Members go to doctor so they can return to work and provide for their families. The doctors write up a 30 day script of painkillers and send us on our way. My members don't choose opioids. Opioids choose them. If there was some way to find out if we potentially have a predisposition to opioid addiction, every one of my members would want to know before taking anything. We love our families and we're proud of our jobs. With something like this, we can make an educated decision, and an informed decision is always better than an uninformed decision. This could be a potential game changer in the opioid battle.

Without something like this, there's no way to know who is and who is not at risk. My members and the providers deserve more information before taking or prescribing these potentially dangerous drugs. The risks of this seem very low. Something as simple as a 32nd cheek swab could result in a lot less funerals and a much healthier population. Shame on us all if we don't take full advantage of every opportunity to protect anyone who could potentially be at higher risk for addiction. Anyone who has seen what this addiction does firsthand knows just how truly important and vital something like this is. I, for one, am tired of going to funerals for my friends, my family, my members, my members' families. I've seen too many lives lost and too many livelihoods destroyed. Opioid addiction destroys lives, families, careers, futures, if it doesn't outright kill you first. I have seen some truly great workers, foremen for their employer, happy family men, throw it all away due to an opioid addiction after an injury. I implore you. We need to do the right thing here. Help me protect my members. Opioid recovery centers are full. They have waiting lists to get in. How about, instead of waiting until a life is destroyed, a family

is lost, a career is over, how about we take a proactive approach? We try to help people before they become addicted. Thank you for your time.

Mr. Sorte: Greetings. My name is Brad Sorte, the President and CEO of CARON Treatment Centers. As a nonprofit organization that's spent more than 70 years dedicated to improving the lives of those impacted by substance use disorder through treatment, recovery support, education, prevention, and research, I implore the FDA Advisory Committee to authorize AvertD to help us save lives in preventable disease of Opioid Use Disorder. The best way to address substance use disorder is to prevent a substance use disorder. While the prescription monitoring systems, insurance company restrictions, physician education and awareness initiatives have all helped to reduce the number of prescriptions for opioids. According to SAMHSA's national survey on drug use, the health of 9.3 Americans was jeopardized by them using prescription painkillers, and 2.3 met the clinical criteria for an Opioid Use Disorder.

We need to do better to prevent those most likely to misuse and develop an Opioid Use Disorder from being introduced to opiates. Advancements in genetics and genetic testing are changing the way diseases are prevented and treated. It's the same with behavioral health and substance use disorder. AvertD is a milestone in preventing the chronic disease that is substance use disorder at addiction. The fact is we recognize that there are some patients who require opioids for acute pain, and this life saving test would allow individuals with a high risk to make informed choices about their pain management. Specifically, with AvertD for the first time, patients will have information about their individual risk for Opioid Use Disorder before they're prescribed an oral opioid for acute pain. It also gives healthcare providers the information they

need to create a truly customized care plan so that if a person at risk of Opioid Use Disorder needs to be prescribed medication, the provider can better plan for the resulting physical, mental, and emotional side effects. This awareness and early intervention can save lives, and that's why we are a research partner for AvertD. This is why I'm here to call for authorization, representing the thousands of families we serve, and we believe advert will save countless lives, families, heartbreak, and ever experiencing the Opioid Use Disorder at all. This test will save lives and we strongly encourage the Panel to authorize its use. Thank you very much for your time.

Mr. Horwitz: Hi, my name is Jeff Horwitz, and I am the Chief Operating Officer of Safe Project. Safe Project is a national nonprofit dedicated to ending the addiction epidemic currently facing our country. Last year, over 108,000 Americans lost their lives as a result of an overdose. That's one person every five minutes. In some cases, in many states, nonfatal overdoses eclipse fatal overdoses at a rate of 15 to 1. In most cases, an overdose, or the pathway to addiction which led to the overdose, was paved with prescription opioid. Yet, despite all of this, few Americans appreciate the fact that they and their families are at risk.

Opioids play an important role in pain management. It's especially true in the management of acute and chronic pain. However, there are also risks associated with the use of opioids. In fact, we are all well aware that an individual can become addicted to an opioid after as little as three or five days of use. Given the ongoing opioid epidemic, the horrendous trends of overdoses, and the lack of appreciation for the danger of opioids at every level of our society, a risk assessment tool is needed now more than ever. An effective risk assessment tool will not only provide doctors with additional information on the decision to prescribe an opioid or to provide a non-opioid alternative, but it will also educate and empower patients to be more

informed and more engaged in their own pain management. On behalf of Safe Project, our partners and all of those that we represent, we strongly encourage you to approve the use of this tool. The ability to authorize the use of a risk assessment resource that can actually provide accurate, actual data in real time that's not based solely upon previous care will better inform physicians and lead to more robust conversations between doctors and their patients. This is indeed an opportunity to manage pain without creating the risk of future overdose. In fact, it may be the one and only new and effective way to effectively stop the addiction fatality currently facing our country. I encourage you to support this product. Thank you for your opportunity to speak, and all the very best

Dr. Watson: Thank you to all the speakers for their comments. That was the final comment that we received, so I now pronounce the Open Public Hearing to be officially closed. We're going to go on now with our agenda. We'll begin the Panel Deliberations. Although this portion is open to public observers, public attendees may not participate except at the specific request of the panel chair. Additionally, we request that all persons who are asked to speak identify themselves each time. This helps the transcriptionists identify the speakers. Now, during the next hour, we'll open up the floor to questions for both the sponsor and the FDA from the panelists. But first, I'd like to circle back to the questions that the panelists sent to the sponsor earlier this morning that we didn't get time to address. Sponsor, are you prepared now to address those questions?

Dr. Brauer: Yes. Thank you, Dr. Watson. We are prepared to address those questions and prior to addressing the question, we would just like to clarify our intended use, as several of the questions relate to that. This is a high level of our proposed indication for use that AvertD is

a genotyping test. It detects 15 clinically relevant SNPs to identify patients at increased genetic risk for OUD. It is for adults 18 and older being prescribed oral opioids for acute pain. It is intended to facilitate shared informed decision making between the provider and the patient, and most importantly, it is intended for use as part of a clinical evaluation and assessment. That is, it is not a standalone test. We do have answers to your questions to help facilitate the process. We're going to repeat the question and then the answer.

Dr. Goldstein asked how many African Americans were in your study and how many declined participation? There were 14 African Americans in our study. We did not collect any demographic information on those patients who declined to participate in our study. We are committed to further evaluating AvertD in African Americans in the post-market setting and look forward to working with FDA with respect to this.

Dr. Dunn asked, "Did the research sites in your study also provide clinical care?" Nine of our sites provided clinical care for patients. Their names included "Research" in the title of the site because that is the portion of the practice that we contracted with to do the work. Importantly, 98% of the participants in our study were coming to the site for routine care. They were not recruited to participate in the study.

Dr. Ness asked, "Did you collect data relative to first degree relatives with substance use disorders?" No, we did not.

Dr. Walker asked, "Was the study powered to detect differences in race or ethnicity?" No. The study was not powered to detect differences.

Dr. Gordon asked, "Did you modify OUD diagnosis criteria if patient was ongoing on opioids tolerance or dependence, and what was the severity of the OUD cases in your study?" To

the first part of Dr. Gordon's question, Criteria 10 and Criteria 11 about tolerance and withdrawal were not considered to be met for those individuals taking opioids solely under appropriate medical supervision, such as medication assisted treatment for OUD. We do have information on the severity of OUD for the participants in our study who were OUD positive. That information is shown on this slide, as well as the sensitivity of AvertD for patients with mild, moderate, or severe OUD. We also performed a statistical test, which demonstrated there was no difference in sensitivity between mild versus moderate and severe and mild/moderate versus severe. I am now going to ask Dr. Donaldson to come to the podium.

Dr. Donaldson: Thank you. Keri Donaldson. I'm going to be answering two questions in sequence. The first is by Dr. Bierut asking about what is the difference in distribution of positive and negative as at risk test scores across different ancestral groups, and are there racial and ancestral differences in the general population? This is a multi-part question, and it's very important to understand before we start that, as opposed to some of the other comments we heard today, we're actually evaluating the performance of our test, not a test that's different from ours in different populations and in different data, but this is the most salient question in terms of de determining the importance of AvertD in different ancestral groups and racial groups.

The first few slides that I would like to show is talking about the fact that there is significant literature to support. These genes have been described previously in other substance use disorders as well as OUD. This does not mean that they're only for OUD, and that is not required when you're doing machine learning to use genes or targets to separate two groups.

The second thing that I would like to get into before we get to the actual distribution of tests, is one of the comments earlier that in fact there are allelic frequency differences between

different ancestral populations. This is shown on the slide that's up as well as this slide. It's important to note that a comprehensive machine learning process must account for these allele frequencies differences to see equivalent performance across different ancestors.

Now, to Dr. Bierut's question. We use the mean allele frequencies from several populations: African American, Latinx, non-Finnish European, Finnish European, South Asian, all from genome ID, as well as African, African American, East Asian, European non-Finnish, American, European, and South Asian from a thousand genomes, as well as African American and European American from NHLBI. We used this frequency data to generate a synthetic, in silico population consistent with the us intended use population of over 1 million individual genotypes per ancestry. All distributions of AvertD are consistent across all ancestral groups, including American, Latinx, Finnish European, South Asian, and East Asian. Of note, there is greater variability in the distribution of test results in groups including African and African Americans, where there's a less pronounced peak of the distribution. This is not surprising given there's limitations of data sources and variability of heterogeny native genetics across individuals of African ancestry.

The next comment that I would like to talk about goes actually one step further. We also evaluated the ability of our test not to diagnose OUD, but to diagnose, or categorize, populations by ancestry. Folks have commented that the test could possibly confound indications by classifying by ancestral group. We took the synthetic population we just discussed from the same data sources and asked, “Can we change the probability of being an individual ancestral group by AvertD test results?” Definitively, the answer is no. Across all ancestral groups, the diagnostic odds ratios only changed in between negative two and positive two. That's important to contrast

into the diagnostic odds ratio changes that we discuss and highlighted for AvertD of 18.1. That ends my questions on Dr. Bierut.

I want to transition a little bit to talk about some of the studies that have been mentioned in the literature and contrasting those studies with what we've seen in our study. This question was from Dr. Bateman, as well as a few others, referencing an article and asking why we didn't adjust for potential ancestral differences., worried about tests could exacerbate disparities for opioid prescribing, not satisfied with the white/non-white breakdown of the study, and talking about the 15 SNPs and why 15 SNPs are significant.

So, to answer this question, you're really looking at two very different populations falling two very different methods with two very different results, not surprisingly. On the left hand side, we highlight our approach. Our approach to diagnostic test design follows a very tried and true method that has been accepted with the community of talking about very clear delineations of populations. In our setting, we're talking about all of our training and learning, as well as our clinical data set we presented today, occurs in opioid naive individuals, excluding chronic pain. That is very important because the genetic signature or contributed between patients that develop Opioid Use Disorder after an acute exposure for acute pain may be different than those that develop OUD after chronic pain. If you do not exclude these populations, as seemingly the other paper did, you can confound and mis-discrete genetic signals.

The other important distinction to make is, in our paper and in our studies, we validated that every patient took oral opioids for between 4 and 30 days. It's consistent with our intended use population. That is unclear in the other study, whether it was controlled for exposure.

The other interesting thing that to note when you're thinking about comparing these two very different approaches and papers is, our disease was very tightly defined. We're talking about confirmed DSM-5 OUD. It's essential to confirm outcomes and have a standard outcome assessment when you're trying to develop genetic classifiers. The paper printed in 2021 from Hatoum is looking at opioid dependence. These are not the same things.

The next thing that I would talk about, which was referred to a little bit earlier, is the genetic information included in all of our patient population and our training subjects was performed in validated and clinically accepted methods of genotyping. A very large difference here in between GWAS studies and validated data sets is commonly SNPs, including the ones used for our assay, may be imputed in those GWAS assays and can be wrong 10 to 20% of the time.

The last thing, or clear distinction to make, which gets to some of Dr. Bateman's questions, is, what is the best practice or approach in machine learning? Broadly, there's multiple different validated approaches. Specifically, to our approach, we take an unsupervised dimensionality reduction to identify the targets versus a supervised approach. You can see varying levels of classification accuracy with supervised versus unsupervised approaches. Lastly, and most importantly, when you're trying to determine the difference between research data sets and research findings and clinical data sets and clinical findings is really, the proof is in the pudding. Performing independent assessment of the performance of a classifier in a prospective, blinded, well-controlled clinical study is the way to assess for performance. That is what we did today, and that information is not available on the other paper. I thank you for your time and I'll turn it back over to Dr. Brauer.

Dr. Brauer: Dr. Wang asked us whether our study population represents our intended use population and whether the higher prevalence of OUD positive participants impacted that. Our study population aligns with our proposed intended use population. Our study population consisted of participants who took oral opioids for the first time for 4 to 30 days. That is identical to our intended use population, which is individuals who are considering their first prescription for oral opioids. The age was the same. Our study population was greater than 18 years old, or equal to 18. That is the same for our intended use population. In our study population, our participants were enrolled as part of a regular clinical visit. In our intended use population, it would be patients who were seeing their healthcare providers, and the results of AvertD would be used in combination with an overall clinical evaluation and a pain management plan. Mr. Mullen is now going to come to the stand to talk about the enrichment part.

Dr. Mullen: Thank you. Chris Mullen. Dr. Wang also asked about the role of enrichment in the study. The approach used by the sponsor was valid and consistent with principles and guidance documents. In particular, I think it's important to note that the enrichment was done blinded to outcomes, so it did not have the ability to bias things. The stratification and sampling that was used by the sponsor was another tool that was used to reflect and mirror the intended use population. Finally, I would note that the motivation for enrichment was really to gather enough positive cases to ensure that we could assess properly the sensitivity of this test with a reasonable study size.

Next, I'd like to address Dr. Farrar's questions about predictive value. He asked about predictive value for difference prevalence levels. We were able to calculate this quickly over the break. So, you see here, for AvertD, the predicted negative and positive predictive values by

prevalence levels, from 1% in the first row, 5% and second, 10% in the third. You can see negative predictive values are 97.6% or higher. The positive predictive values range from approximately 4% at 31%, depending on prevalence. But I think it's important to put those positive predictive values in some context. In particular, we found a reference to a currently cleared genetic test for warfarin genotyping for bleeding risk, as noted in McClain in Genetics and Medicine 2008, this positive predictive value of 7% was based on a 5% prevalence rate.

Additionally, I'd like to put AvertD results in the context of currently existing tools for opioid use, non-genetic tools, but tools that are in use, nonetheless. Here, you can see for our AvertD test, as well as the other tools listed below, the sensitivity/specificity results, as well as the positive predictive values for those prevalence rates of one five and 10%. And you can see that no other currently tool in use, again, non-genetic tools, provide the combination of sensitivity and specificity that AvertD does. And further, the positive predictive values, no matter the prevalence level, are an improvement upon these other tools. Again, though AvertD is not intended to be used as a diagnostic standalone, it's intended to be used to identify patients at high genetic risk. And for some further clarification and clinical perspective, I'd like to pass the mic to Dr. Zacko. Thank you.

Dr. Zacko: Hi, Chris Zacko. There are several times questions came up, and I'd like to go back to the safeguards put in place against false positive and false negatives and overreliance on the test. I think it's really important to emphasize these are in alignment with the patient education materials that will be supplied. And this will, again, go over the indications for the test. It'll go over the interpretation of the test. And also, more importantly, emphasize using current practice guidelines when treating the patients, and that always comes first and foremost. In my

practice, again, I would use this preoperatively to get a better sense of overall global risk for a patient and not just the risk of a given surgery. Postoperatively, I think it's also important to emphasize, again, as stated, I use opioid sparing methodologies to control someone's pain, but we have to continually reassess that pain. There are several tools at our disposal that will do that, and this test will help inform you which tool to use, which drug to prescribe, how much to prescribe, how long to prescribe it, and which ones can be used in synergy with others. That does not mean we won't prescribe the opioids, and if an opioid is a right medication for a patient, we would use that medication.

I think clinicians are used to using tests and understanding that no test is perfect. For example, before I do any major surgery, we test coagulation parameters and platelets, and that would tell me a patient is not an increased risk for bleeding during a case. However, it's not uncommon for me to be doing a particular procedure and that patient is bleeding more than I might expect from a given test. It doesn't mean I abandoned my techniques to control hemorrhage based on a test that told me they shouldn't be bleeding. So, in a similar way, I assess the patient. If they're in pain, I treat the pain, and if opioid is the most appropriate medication, then that's what we would end up using.

In terms of other tests that we're used to using, I want to touch on the genetic risk assessment tools that we're also used to using. For example, there's tests like Factor V to find an increased risk for blood clots. They're 10 times more likely to develop a lifetime incidence of pulmonary embolism than your base rate. This can inform decision making in such as quitting smoking, hydration levels, and oral contraceptive choices. There’re other tests, like VKORC1 and COIP29, that accounts for 18 to 30% of the variant in stable warfarin doses in the European

population. This allows clinicians to titrate the appropriate dose or choose an alternative agent for anticoagulation.

Lastly, I'd like to comment on something else I haven't touched on in the importance of this test in that it could be a steppingstone for research in the future. There's no test that's perfect for solving complex problems, particularly something like Opioid Use Disorder, particularly the first test in the space. I think we can all think back to APOe4, when that was used, and it had limited use until years later. We were able to get PET scans, and those two pieces of data together are more powerful than one test alone. And I think this test, when used properly as prescribed, is a safe test. It could be quite efficacious, and it really starts to shine a light to the dark corners of a complex problem like Opioid Use Disorder. Thank you. Next, we'll have Dr. Garbely come up.

Dr. Garbely: Good afternoon, Joe Garbely. I want to just dovetail on what Dr. Zacko was saying about false positive and false negative tests and kind of reframe it as, what is the proper reliance on this test? This test, AvertD, is not a diagnostic test. It does not diagnose Opioid Use Disorder, but rather, it gives us a genetic risk assessment tool that we can utilize when we are encountering a patient for acute pain management. Physicians are already using many different risk assessment tools. That is the standard of care coupled with comprehensive physical examination, careful history, querying the PDMP, and also looking at medical records that are available, as I said before. This is our standard of care. This particular test will enhance that standard of care, not abandon it. It's not meant to replace that standard of care. And that will be clear in the educational materials and how we train healthcare professionals to use this test.

Now, I want to talk about stigma, and this is a very important topic, one that, unfortunately, I've encountered for all too long in the addiction medicine space. Addiction, and Opioid Use Disorder in particular, has encountered a significant amount of headwinds in the form of stigma. So, stigma associated with Opioid Use Disorder and also with addiction in general, has the negative perception of a moral failing, a lack of willpower, that it's a choice rather than a chronic disease with known biological contributing factors. AvertD is not here to diagnose Opioid Use Disorder. It's just, again, providing a genetic risk tool that we do not have right now. AvertD could help reduce that stigma. Acknowledging that Opioid Use Disorder is a chronic disease and a complex disease, I see it as an opportunity to further information and to help with a nuanced conversation that we all have to have with our patients prior to managing acute pain and prescribing opioids. This opportunity to me, is long in the waiting. We need to have a test like this so that we can have a more nuanced conversation.

And lastly, I would like to summarize and give a summary statement. I can't stress enough that AvertD is not a diagnostic test. Again, it does not diagnose Opioid Use Disorder. It identifies if a person is at high or low genetic risk for Opioid Use Disorder. Clinically, having these test results will not stop physicians from considering opioids for acute pain management. AvertD will be used along with at least five other risk assessment tools: the pain medication questionnaire, the opioid risk tool, the brief risk questionnaire, the brief risk interview, the screener, and opioid assessment. In addition, we will continue to review the PDMP to assess a patient's prior use of opioids. We are not going to abandon best practices because we have a new tool. Perhaps most importantly, having this test will be a catalyst for having a conversation that's much more nuanced than it currently is between the physician and the patient about the risk of opioid use and the benefits of managing pain in an opioid sparing way, especially if one is at

high risk for Opioid Use Disorder. Does this happen today? It does, for sure, but it doesn't happen enough. The sponsor is committed to continued data collection and analysis in the African American population, and they're committed to a robust education program and a controlled launch to allow them to maximize the effectiveness of the educational materials.

In the end, we need to ask ourselves, is society better off with or without this genetic risk assessment test? Since genetics largely influence an individual's risk for Opioid Use Disorder, I would like to have it available to me, and other physicians like Dr. Zacko would like to have it available to him, and I know all of our patients would like to have it as well. Thank you for your time and consideration.

Dr. Watson: I think we're done with the sponsor responding to our questions. Thank you very much.

Dr. Brauer: Thank you.

Dr. Watson: This is really the time for the panel to deliberate. The purpose of this is so that each and every panelist can get any and every question, discussion point that they want out, out. So, I will ask to hear from everyone. Questions for the sponsors or the FDA. I'm going to start by going around the room, and there are some questions in the chat. I think I will start with those. We have a question from Dr. Walker about the number of African Americans. Dr. Walker, will you please start?

Dr. Walker: Yes. I just wanted some clarification, that in the 1300 study participants, there were 14 African Americans. Is that what the sponsor said?

Dr. Watson: Sponsor? Would you answer?

Dr. Donaldson: Yes, I'll answer. Thank you for the opportunity. No, that is not correct. So, there were 14 African Americans in the clinical study population. In the 1300 training and learning data set, which is how I understand your question, approximately 30% of that population was African American, as determined by ancestry profiling.

Dr. Walker: What was the size of that population, please?

Dr. Donaldson: The specific population that I'm talking about is the 1300 that you were talking about. And the 3000, I'm happy to provide that information as well. That'll just take me a second to pull it.

Dr. Walker: Okay. Thank you.

Dr. Donaldson: You're welcome.

Dr. Watson: Any other questions, Dr. Walker?

Dr. Walker: No, thank you very much. Next, we have a question from Dr. Compton.

Dr. Compton: That's all right. My question's been answered because we've moved into just the deliberations.

Dr. Walker: Okay. If you have any other questions, we can…

Dr. Compton: Well, I do, but I was going to wait my turn with my hand up. But I'll go ahead. My question relates to the potential risks in a negative result, because one of the points that was made is that a negative result would mean that standard of care for opioid prescribing can then be readily applied without having to make modification. And yet, what we've all too often noticed, is that standard of care is a high level of opioid exposure compared to the actual

needs of patients and is an evolving process. So, I'm curious if there's any data or approach to understanding potential risks of continuing to over-prescribe in the face of low-risk patients. This may not be a concern for the patient themselves but is a major concern for their social networks. Because we've learned many hard lessons over the last decades that opioids are not just used by the patient to whom they're prescribed, but they are very frequently shared with family, friends, and others. And so, this is not just a concern for patients, but is a potential risk for overdose risk and for harms to a much larger population than simply the patient in front of us. That was a little bit of a comment and a bit of a question. If there's any information about the risks of changes in prescribing patterns in response to a negative test, and it may not be known.

Dr. Watson: Sponsor. Do you have that information?

Dr. Donaldson: The answer to your question, I think it's very valid. We did do a market assessment for what's currently being used in terms of risk prescribing in the intended use population for acute pain. We saw about 64%. Probably highlighting the comments from Dr. Conley, that, this population, certain people are doing what standard of care should be, and certain people are not doing what standard of care should be. Meaning about 40% of the prescribers are not performing risk assessments prior to prescribing. We see this as a significant opportunity for our test to inform that discussion and decision making, aligning that percent of the population that's not currently performing what CDC would say standard of care is to what standard of care should be. So, in our presentation when we reference standard of care, we're really talking about what standard of care should be, meaning how people should be minimization of opioid prescribing in this intended use population, or maybe even opioid sparing strategies. Thank you for the opportunity.

Dr. Watson: We have some hands raised, and I'll go through the hands first, and then I will make sure we go through the entire voting panel one last time, so everyone is able to make comments. But I'm just going to go left to right in order of who appears on my screen. So, Dr. Dunn, do you have a question?

Dr. Dunn: Yes. Walter Dunn. So, I actually have a question maybe for my fellow panel members. I think someone could comment. My concern is about the composition of the study subjects and accurately it reflects the clinical use population. Certainly, the study design was meant to enrich for opiate positive individuals. The percentage of comorbidities, or the prevalence of comorbidities, that's kind of my interest here. In questions to the sponsor, and also from their document, they note that, at time of index exposure, the prevalence of comorbidities was not any greater in the opiate positive population compared to general population statistics, and even at time of enrollment. So, after the OUD had been diagnosed, there's still not a significant difference. I'm wondering, for those of you who are experts in the field, it is it unusual to see that, even early on at time of opiate exposure, that you would not see a greater incidence of comorbidities? And even at time of opiate use disorder diagnosis, not a greater incidence of comorbidities? Because my understanding is that does not seem to comport with what the sponsor has demonstrated with this subject population.

Dr. Watson: Would any panel members like to weigh in? Dr. Bierut.

Dr. Bierut: As a psychiatrist who works with people with substance use disorder, comorbidity is incredibly common, and we expect a lot of the comorbidity to start at a young age. We know that anxiety and affective symptoms start at a young age. Smoking starts at a young age. Drinking starts at a young age. What I am assuming is that their assessment for this

was not fully capturing. That would be my expectation. Wilson Compton is also an expert who may wish to comment.

Dr. Watson: Dr. Compton, please offer us your expertise.

Dr. Compton: Well, this is Wilson Compton. I would absolutely agree with Laura that the rates of comorbidity looked low to me, but I assume that that's due to the fact that they're based on medical records and other assessments, as opposed to a personal, thorough interview for all of those potential psychiatric conditions. And with the question I raised earlier about tobacco, it doesn't even fully cover the whole range of comorbidities in the assessment.

Dr. Watson: Would the sponsor like to respond?

Dr. Brauer: Yes. Thank you. We do have data on comorbidities at the time of enrollment. This is after the participants had been exposed to oral opioids and had the opportunity after exposure to develop OUD. And as you can see, certain comorbidities are in fact higher in the OUD positive population at this point in time. For example, 43% of our OUD positive participants also had depression. 24% also had alcohol use disorder. 34% had anxiety. This is a very different profile than what we saw from these individuals at the time of index exposure. Thank you.

Dr. Dunn: Just a quick follow up to my colleagues. It sounds like, if they had done a more thorough job of identifying these comorbidities at time of index exposure, you would have anticipated a higher prevalence of these other SUDs, either maybe cannabis or tobacco or alcohol, that could potentially inform your clinical assessment of their risk, even before they were exposed to OUD. Would that be accurate?

Dr. Watson: Colleagues?

Dr. Bierut: I'll say yes.

Dr. Compton: I'm not sure. I mean, partly this is a somewhat retrospective prospective kind of study. So, it's hard to tell from this particular study design. And it'll depend on the age and certainly those rates at the cross-sectional rates for OUD look a lot more like what I would've expected in the data we just saw.

Dr. Watson: Thank you. Dr. Bateman?

Dr. Bateman: Great. Thank you. I just want to draw the panel's attention to some of the points made in the letter we received in our briefing packet from Dr. Hatoum et al., from WashU, Dr. Engelberg from Indiana University, and Dr. Gaertner from Yale. These investigators are some of the leading human geneticists working on Opioid Use Disorder. And they make the point in their letter that they've conducted some of the largest, well-powered in state of the art genome-wide association studies of Opioid Use Disorder to date. And they say that current knowledge about Opioid Use Disorder genetics is strong enough for it to be clear that it is not possible for patients risks for Opioid Use Disorder to be predicted from the 15 SNPs that are included in the assay. And they point out that it's very likely that the associations that are observed in the sponsor studies are highly confounded by genetic ancestry. And they found, when they did an evaluation of these SNPs, that if ancestry is balanced, it does not predict Opioid Use Disorder risk any better than chance alone. They cited a recent genome-wide association study of individuals of both European and African ancestry looking at Opioid Use Disorder. They pulled together — this is the study in Molecular Psychiatry from 2022. They pulled together seven large cohorts, resulting in over 20,000 cases and 600,000 controls.

And, despite doing a genome-wide association study across all SNPs and this very large cohort and correcting for ancestry, single nucleotide polymorphisms accounted for only 3.8% of Opioid Use Disorder variants. So, you know, I think that that study and the points they make about the potential for confounding by ancestry really calls into question if the performance of this this test. If, in fact, their explanation is correct, which it seems likely that it is, this confounding by ancestry could exacerbate racial and ethnic differences in opioid prescribing. So, it carries really significant risk. I just wanted to make sure everyone had noted those points, and, to me, they're really quite compelling.

Dr. Watson: Thank you. Dr. Wang, do you have some questions? Thoughts?

Dr. Wang: Yes. I want to follow up on a couple of different points I had earlier. One of them was whether the study design would actually represent the target use population. And I'm curious to hear our panelists' thoughts on this as well, besides myself. I know the sponsor responded saying that yes, the clinical study did represent the targeted population. However, I want to point out that the target population is an opioid naive adult population, while the clinical study here is looking at a heavily recruited population, enriched population, with people who are already on opioid OUD treatments. So, I'm curious to hear your thoughts on whether you think this would represent the intended use population. So that's one of my questions.

The second comment was the analogy has been drawn by the sponsor between this test and the Warfarin genomic test, the pharmacogenomic test for warfarin to predict the warfarin dose. However, I also want to point out, a difference here is that, in the warfarin dosing situation, there was actually an equation out there that will take into account both the genotype information as well as the clinical information as well as demographic information and combine everything

together to come up with a recommended dose, rather than an experience-derived dose. In this case, AvertD, we know the genotype information, the genotype result, but as the sponsor mentioned, this will be used together with at least five other tools, clinical tools, together. But to me it's still unclear how these other five tours will be combined with the AvertD information to derive the final decision making for either the surgeons or the primary care providers to decide how much opioid or whether even to give the opioid to the particular patient.

And finally, I have a concern regarding the possible false sense of security here that may be elicited by a negative result. We hear a lot of the comments from the public regarding, if we knew about the results 10 years ago, then we could have avoided prescription and possibly avoided the funerals, et cetera. But I want to also point out, the clinical study was not designed to look at whether the AvertD result would increase actually the outcome or decrease mortality in the opioid use population. It was merely designed to tell whether they can predict OUD development in the future. So, I'm a little bit concerned about whether a negative result will actually lead to a false sense of security.

Dr. Watson: Anyone have comments about that? Okay, Thank you. Any other questions, comments, Dr. Wong?

Dr. Joniak-Grant: I have a brief comment about that. I think there would be a bit of a false sense of security in the fact that, just looking at the public comments today, there was sort of this built-in assumption that everyone who had an issue would've tested as high risk and it could have been avoided. I think that right there shows us — and that was from doctors as well — what could be a real risk with having a false negative in this situation.

Dr. Bateman: Yeah, I think that comment is spot on, and I think, especially given concerns about the validity of the test, I think it could really give people a false sense of assurance. If the GWAS study's correct and there are only 4% of the inter-individual variability is explained by genetics, we could really go down the wrong path in thinking people are safe to prescribe opioids to when, in fact, they're just as high risk as the rest of the population.

Dr. Farrar: And if I could just make one brief comment and then I'll comment more in a minute. But you notice that in the slide shown the negative predictive values were very, very high. And I just want to make clear — this is probably clear to those of you who look at these things regularly — but the negative predictive value of not doing the test in 1% prevalence is 99%, and in 5% prevalence is 95%.

So, not doing the test is a pretty good test, given the structure and the current characteristics of this test.

Dr. Watson: Thank you. Dr. Wang, did that discussion cover your concerns?

Dr. Wang: I think so. Thank you.

Dr. Ness: Well, actually, can I make this one statement again? A question is — I'm kind of new to the advisory panel sorts of things. The company had said that how they were going to report these things was high risk/low risk. Can it be a recommendation from the FDA from the advisory panel if you took out that a person is low risk, and they're just within normal limits? We treat things differently in medicine when someone's labeled low risk. That somehow, again, gives us a false sense of security. Whereas if they're within normal, then I make sure I look at all those other factors that come into abuse, and environmental, as well as other sorts of things. Are we able to make that kind of a recommendation, that they've given some evidence that there is a

high risk predictor with this, but I have troubles with coming up with a low risk designation, because that implies, “Oh, phew, I'm able to do whatever I want now.”

Dr. Watson: So as a point of our job here is to answer the questions that the FDA has presented to us, so we will, within the questions that are answered, we will give our answers, and that's kind of what the major job of this panel is. All right. Did that explain that, Dr. Ness?

Dr. Ness: Yes.

Dr. Watson: Okay. Dr. Goldstein, you had a question?

Dr. Goldstein: Yes. Thank you. I want to point out that the clinical study population looks nothing like the American population. And so, just to remind you, when one does machine learning with a lot of data and a lot of variability, you don't end up with an algorithm that you can just easily explain to somebody. You do a clinical trial to test whether that machine learning algorithm is really telling you ground truth. And I just want to know from the sponsor, why didn't you do a proper clinical trial with a study population that looks like the American population, because that's where this is going to be used?

Dr. Watson: Sponsor, do you have an answer scene?

Dr. Donaldson: Yeah. Thank you so much. So, the intended use population for this test was derived and specified with the FDA prior to study and trial design, and it uses the portion of the population that is prescribed oral opioids in the US as specified in a few consensus documents, including those from CDC. As Dr. Goldstein has noted, that's different from the demographics of the total US population. Those are two different values, Dr. Goldstein. So, the way we did it, the way we addressed the intended use population is the percent of the population

as specified in several source documents, including those by the CDC, that specified by age, sex, and a few other things. What we could as best approximate, which you made the point is that's how you test the generalizability of your machine learning classifier. But that population is different from a representativeness than the general demographics of the total US population.

Dr. Goldstein: I find it hard to believe that the frequency of people of color in the intended use population is really that low.

Dr. Donaldson: Well, it's about half. There's numerous studies that say that people of color, or African American population, are underrepresented in the prescribing of oral opioids in the US. Many people on the panel could probably tell you more about that than I can, but the average population of the US that that would classify as African American by race could be somewhere around 13%. Our goal was half that given the prescribing guidelines. So, the target goal for enrollment for African Americans pre-specified was right around 6%. And we tried very hard to get that 6%, and we came close, but did not achieve it.

Dr. Walker: Do you have those numbers for us now in the two study populations, the number of African Americans?

Dr. Donaldson: You're asking for the training sets?

Dr. Walker: Either the 1300 or the 3000.

Dr. Donaldson: In the 1300 it was 30% African American, and I'll get you the 3000. We have the back room pulling that information for you,

Dr. Higgins: But it was much lower in the clinical sample.

Dr. Donaldson: It was specified in the clinical sample to accurately reflect the intended use population in the US. And I can reference the CDC guidance documentation on that as well.

Dr. Goldstein: But you ended up with an underpowered subpopulation.

Dr. Donaldson: I'll ask Chris Mullen to answer that question.

Dr. Mullen: The power of the study was based on the total sample size and the endpoints, right? So, there was no specific claim or hypothesis test that was pre-specified. Sample size was small. I think as the sponsor noted earlier, there's a commitment to exploring addressing this further with the FDA and post-approval studies, if that's a possibility.

Dr. Watson: Thank you. Dr. Goldstein, did you get your questions answered?

Dr. Goldstein: Yes, I got my questions answered, but not my concerns.

Dr. Watson: Okay, thank you. we have a question from Dr. Zaafran. Oh, I'm sorry. I'm gonna come back to Dr. Ruha next, and then Dr. Walker. Dr. Zaafran?

Dr. Zaafran: Sherif Zaafran. I know a lot of questions have been asked about the demographics, and here's what my concern is. I know the sponsor's been asked this question many different ways. First of all, I'm an anesthesiologist and I'm in the operating room all the time, and I can tell you that the population of African Americans is far larger than the sample size that you have on there. The second part of that is that the “white” quote unquote population is really represented in a very monolithic way, when that is not necessarily the case, especially when you're talking about things genetically. People from Mediterranean backgrounds have certain specificities around genetic makeup. People of a Jewish background might have

something that is very specific to them. That kind of lack of variability in the sample size, to me, is very, very concerning, especially if we're going to be drawing conclusions out of it.

And that's kind of the second point I wanted to make and ask is that — I know a lot of concerns have been raised about sampling and taking the results of these tests and potentially, if you have a negative propensity, that there may be overuse of opioids. I think it goes both ways. You have a propensity of overuse of opioids, and I think there's also a propensity of underuse of opioids. I mean, people who have surgery and have acute pain will need pain control. And, with all due respect to those who are out there saying that there are options for no opioids, I know that there's opioid sparing techniques, but that's what it's called, opioid sparing techniques. There may be a decrease in the amount of opioids, but I think it's unfair to assume that that means zero use of opioids. And I worry that that stigma, and I'm going to use that word very deliberately, is going to cause a lot of, again, regulators or other entities out there to frown on the appropriate use of medications for people in that type of setting.

When we were on the pain management task force, a lot of things that were being addressed were the CDC guidelines from 2016, where it talked about 90 MMEs, not as a hindrance to use opioids, but as a mechanism for primary care physicians to consult with chronic pain physicians when more than 90 MMEs were used.

Well, regulatory agencies, including my own, started actually penalizing even the chronic pain physicians for using greater than 90 MMEs, which was exactly what was not intended. So, we can talk about the fact that we're going to have concerns raised out there by saying that it's not intended for this use, for that use. But the practical point is that it will be. And unless we're very careful about the data being valid and making sure that the screening test is appropriate,

even if you get to the point where it's appropriate, there's going to be a lot of unintended consequences of how patients are going to be treated based on these results. There will be stigma, there will be undertreatment, there will potentially be overtreatment, and we've got to be very, very careful with that.

Dr. Watson: Thank you. Dr. Zaafran. Would you like to respond, sponsors?

Dr. Donaldson: So insofar portions of this, we can respond to, myself, and portions I may ask one of the clinicians to come up to. Insofar as choosing the race and ethnic groups for the clinical trial, they were pre-specified and they follow FDA guidance documentation and they're consistent with US government standards on how, in diagnostic tests, you have to specify these populations. We agree that the diversity that is represented by those racial groups, the broad categories, we can and we are committed to in a post-market way further exploring that. But insofar as a regulatory submission, we have to follow the specified race and ethnicity categories that are current with FDA guidance.

In terms of overuse or underuse of our test, we believe that the best way to appropriately roll out a diagnostic test like this is through measured metered centers of excellence, where we will work with the practitioners that are ordering and the patients to hone the use and the information over time, prior to a broad launch of the test.

Dr. Watson: All right. Have they answered your concern to Dr. Zaafran?

Dr. Zaafran: I do think that they're talking about what the minimal standards are. I think that's very different from what a lot of us, as panelists, are thinking about what may be the appropriate standards of what is reflected out there. So minimal standards, to me, is not necessarily the best standards. But again, that's my opinion.

Dr. Watson: Thank you. Dr. Ruha, do you have questions, comments?

Dr. Ruha: It's more of a comment, and it's actually not that different. There's a been a lot of focus on the false sense of security we might get with a false negative. But I, too, had wanted to point out the potential for undertreatment. I hear a lot about how this is a tool to add in risk assessment and that perhaps surgeons will be using this risk assessment tool along with the several others available. I can't help but feel that that's somewhat unrealistic. I don't really think that before prescribing opioids, it's common right now for physicians to really be going through all these available risk assessment tools. And I'm afraid that this will be it. And I'm actually even more concerned with not only the false positives, but the real positive results of patients getting undertreated for pain.

And even when I look at slide CO-72, I think from Dr. Zacko's presentation, where it was how to mitigate the potential risks, the false negatives were the clinicians are following standard of care. But yeah, if you look at the false positive, it's like, well, if pain management with non-opioid alternatives is insufficient, well, the worst pain is really in that very acute post-op period. And that's when the patient may not have access to opioids because their test was positive and be undertreated. And I think in the real world, they may be at home suffering, unable to get in touch with their physician or get an on-call person who isn't going to prescribe them the opioids or show up in emergency departments. And I feel like patients are going to be undertreated and suffer as a result of that. And I did hear quite a few of the public speakers, in addition, including physicians — it sounded to me like they were thinking, “Hey, this is also going to tell us who not to treat with opioids.” But really, I think everybody needs to be approached with extreme

caution. Everyone should only receive opioids if they really need them. And if you really need them, you really need them. So, I just wanted to add that.

Dr. Watson: Thank you, Dr. Ruha. Dr. Walker, did you have a question?

Dr. Walker: Yes. I did. I wanted to reinforce what Dr. Bateman was saying about questions now about test validity. I think this is a real concern, particularly now that there is in the published literature that correcting for genetic ancestry makes the predictability of the test basically a coin toss. I'm not a statistical geneticist, so I know the sponsor had some points to address that. But I feel that what really needs to happen now with — in fact, the issue is that the group that did not see the linkage was using a different algorithm, not the sponsor's algorithm. I think now, though, the onus is on the sponsor to do the correction for genetic ancestry and show that the test is still valid. I feel like that is the main thing that that needs to happen for this to go forward.

Dr. Watson: Would the sponsor like to respond?

Dr. Brauer: Well, we'd like to respond to two things, if acceptable. We do have information on prescribing practices of 165 physicians. That seemed to be a concern that was expressed. If I may go ahead. We did a survey of 165 physicians to understand how they are currently doing risk assessment before prescribing opioids. We had prescribers from across the country from different areas and different types of practices. It's important to keep in mind that over 90% of prescribers today use some form of risk assessment prior to prescribing opioids. So, the thought of using tools to use for risk assessment is part of what these prescribers are doing today, and that was specific for things like elective surgeries, hip surgeries, knee surgeries, and spine surgeries. And of interest, over 50% of the surgeons who participated in our study describe

themselves as not satisfied with the current tools that are available for their use, indicating that they really need something else. Thank you for allowing us to share that information with you, because that may be helpful for you in your deliberations. And for one follow up about the Hatoum paper. Thank you.

Dr. Donaldson: Thank you for this question, again, and I think it's incredibly important to spend a minute talking through this because there are some concerns still being raised. Fundamentally the paper that was written by Hatoum versus the paper that we have published in our previous work are focused on two very distinct populations using two very distinct types of information. And it's not surprising you get different results. The comment on, "You should adjust for ancestry like the Hatoum paper": Our approach adjusts for ancestry in the dimensionality reduction at the very beginning of the machine learning. We do that in an unsupervised fashion. It's one of the reasons why, secondary to the question that I answered earlier, we do not see with our algorithm differences in classification by ancestry across huge in silico data sets. So that's best practice on how to evaluate your test.

I do think it's very common that the first step in risk assessment development is bigger is always better, and, from an analytical approach, try everything. It is what you see in this paper. Commonly, what you get with unclear disease definitions, unvalidated data sets is seeing no difference. The absence of evidence in that setting is not the evidence of absence. It's exactly the opposite. Some of the previous discussions talking about GWAS studies fundamentally in OUD versus other diseases, it's important to understand GWAS, in things like mental health diseases like schizophrenia, originally found no difference when they didn't subtype the population. Groups like this Hatoum paper over several years, when they started to characterize different

subtypes of schizophrenia, the genetic predictors became more clear. Our work represents a later version of what's on the right, meaning that we have validated data sets, clear exposures, clear populations, and we've seen generalizable results, as well as results that can be used to disprove that our classifier confounds by ancestry.

Dr. Watson: Thank you. I think Dr. Bierut, we have you next for questions.

Dr. Bierut: Thank you. I'm making more of a comment about where I'm sitting with this discussion. I see two different issues, as a group, as I'm thinking about this. One is, what is the validity of the study, which is an important issue of how does it work, is it really predicting, is there confound by race? And then the second part is the implementation of the study. Let me talk about the implementation first. I think that we really don't have an idea at this point about how it will be implemented. The test was really done in a more retrospective manner of seeing if it could identify differences with individuals with OUD versus not OUD. But when this test is implemented, what we believe will happen is there will be a change in physician behavior, and there will be a change in patient behavior. And at this point, I don't think we have a good idea how this behavior will change. We have concerns with this implementation that there may be reduction of prescribing to people who are at high risk, and maybe inappropriate reduction of prescribing to that group, and the individuals will suffer pain and not get analgesia. We also have concern that, for those individuals who are at normal risk, may be prescribed opioids more readily without much concern. There are concerns either way. I will say though, overall, I am a believer that as we understand our genetics — and I know the genetics, and I know there's strong genetic drivers to these diseases — it is good to have this knowledge, and for an individual to

have this knowledge, and a physician. But I think we are still in this beginning area of how it gets implemented.

Returning to this issue of the validity, what I know from the GWAS studies is that the percentage of the variants explained by the GWAS findings to date is very small. I am surprised at the strength of the findings in this, what I would say is a small sample of 380 some odd people, that the individuals are classified so well into negative and positive for OUD. I just find that surprising, knowing how the genetics works with these GWAS studies. What I believe the sponsor will say is that they have a proprietary algorithm here, where they're not just doing an additive 'this SNP adds this much risk, that SNP adds that much risk,' but they're doing some type of combination, multiplicative type of adding together of SNPs, which is their proprietary thing. I would really like to see this in another data set so that I feel confident about the findings. And I would also say that there are genetic data sets that are out there that I think can be used. And thinking of the big data emulated trials now, where there's UK biobank genetic data, there's All of Us genetic data, and can an emulated trial be done in those types of settings? So, I have more, but those are kind of where I'm thinking going forward.

Dr. Watson: Thank you, Dr. Bierut. Dr. Gallagher, you had a question?

Dr. Gallagher: Yes. I don't want to be repetitive, so I won't give any of my repetitive kind of questions, but one of the things that concerns me is that machine learning, the machine learns more the more data you put in. And so, to identify these 15 SNPs that seem to have so much overlap with other identified mental health and other diseases, I'm just concerned a little bit that, how do we have confidence in the fact that this is the right 15? Because there is that large overlap, and I'm not seeing something that goes, "Oh, these are the two or three out of these 15

that do this, and then there's all these other additional overlapping ones." So, I'm just concerned about that, for one.

And then a basic question is, if you've only done 3000 or 1300 or 385 or whatever it is, when do you expect — to the sponsor — to be able to say, "Oh, we now have enough that our test may have to change, or we'll have something different." Because I'm also looking at the risk of that negative response and how that might impact. Thank you.

Dr. Watson: All right. Sponsor, I think the question was posed to you.

Dr. Donaldson: It was and thank you for the opportunity. So, a couple things that I want to cover here. So firstly, one of the comments earlier was, can you change the labeling from low risk to normal or normative risk? We're completely open to that labeling change in terms of the test report or result report. I just wanted to note that earlier. I think it's a very important point that came up as part of one of these conversations.

The second thing that just came up with Dr. Gallagher, we were talking about the ability of data sets and do bigger data sets contribute to more knowledge. I think overall the answer to that is correct and true. My caveat is that the data sets have to be well-characterized and validated. So, one of the challenges, and it took us a fair amount of time to get to 3000 total results where we did some of our development work is because the vast majority of data sets out there, in particular, some of these GWA studies do not have validated outcomes, meaning DSM-5 OUD, what we're talking about here, nor are the genes determined in a way where the veracity of those genes are significant enough to be used in clinical development. So, it's important to understand that what we have today really represents the largest amount of validated data sets in our intended use population.

The last portion of that that I think we did not share, but I'm happy to go through, is what is the contribution of the different genes to the total amount? And this gets to one of Dr. Bierut's questions as well as Dr. Gallagher's questions. So, what you see here is the 15 SNPs and the genes of where those SNPs are located in terms of how it impacts a classification ability of our result. And Dr. Bierut was correct. This is not an additive cumulative model like a polygenetic risk score. And that's one of the challenges when people see this and say, "How can you use 15 genes to derive such clear distinctions between populations?" In reality, these 15 genes are combined in around 1600 different ways to give us a differential division between two populations.

How do they relate to each other, which was Dr. Gallagher's question. On this slide, what you're seeing is all 15 classified into two groups, what we call a high impact group and a low impact group. The division, or the line in between the high and low impact group, is that results or genes or genetic signals. If you remove one of those targets in the high impact group, it causes reclassification of 20% or more of the results in between as contrast between the three genes of the three SNPs and the three genes below the line. That causes a change in impact somewhere in the neighborhood of three to 10% of those results. So that's how this is classified. It's the best that we can say in terms of what does an individual gene contribute to the overall score.

The last thing that I want to talk about is how long this type of work takes people. We're talking about new data sets coming out recently, large data sets that we, as a commercial entity, don't have access to. We would love to have access to high quality data sets with validated results. We've been doing this stuff for a long time and we've gone through and followed a standard approach to developing a clinical test, and that includes multiple iterative populations,

characterized, validated, and followed. And what we're talking about today is really the end of a very long process that's taken about 5 to 10 years.

Dr. Watson: Okay. Thank you. Ms. Joniak-Grant, do you have a question?

Dr. Joniak-Grant: Yes. Hi. This is actually to address some things that Dr. Beirut brought up about how it's implemented, saying that we don't really have a good idea how this will happen. I think we can look at clinician responses to other opioids and make some educated guesses. I mean, we've seen time and time again how guidelines get shifted into mandates and, on the ground, how a lot of things that are meant to apply, for example, to acute and postoperative patients, for example, prescribing limits legislation. In my home state of North Carolina, we've actually seen the greatest decreases in prescribing for chronic pain patients, both with cancer and without cancer. We really are in sort of an era of opioid pharmaco-vigilance. Opioids are stigmatized. People who prescribe opioids are often stigmatized. And so, I think we have to be really mindful of the context that this is going to be happening within and look to what's been happening so far.

I keep thinking of it from the patient community. I'm really concerned about people who get positives, an assumption of having an addictive personality, so to speak. How would that impact getting meds beyond opioids? How would that impact other things regarding food? For example, how will that translate into interactions with family members, especially children and about what is appropriate care for them or not appropriate care for them? I think if we really kind of pause and look at what's been happening, this focus on legal liability, how will legislatures get involved? How will those who refuse to take the test be seen by their clinicians? Right? There's a lot of times that behaviors where people question opioid prescribing clinicians can have a

tendency to view that as drug seeking behavior. So, if you refuse this genetic test, how will that be interpreted?

And I think the reason, at least for me, it gets kind of extra mucky is the fact that they want to say it may do this, and it may do this, and it may do that. And then the yes/no makes it become a will. Right? It becomes very bifurcated. It becomes very sort of clear cut. And we saw that in public comments from not just potential patients and parents of patients, but also from the doctors, is this assumption that we're going to know who's at risk and who's not at risk. And we can't. Perhaps it'd be better if it presented like prenatal testing results. There's a 4 to 17% risk to develop into Opioid Use Disorder, versus, yes, this is high risk, or yes, this person is low risk. That might be one way to address it.

But I think the big point is that we've heard time and time again, an informed decision is always better, but if 20% of these cases are going to be an error, and even when they are correct, we're getting a 17% at most kind of prediction. Is that really informed? For me, as a patient, I would say no. This is very gray, and the test is trying to put it there in a way that's simple. And I think it'll, based on what we've seen in the past with how this information is used in clinical interactions, we have to be extremely cautious about this going forward.

Dr. Watson: Thank you very much. Ms. Higgins, please show us your comment.

Dr. Higgins: Jennifer Higgins. I want to go back a little bit to the design of the study, again, and I'm wondering, maybe I'm incorrect in this, but it seemed to me that they were largely east coast research centers that were used to recruit. And I also wondered about this could be the FDA or the sponsor. Why did you decide to use subjects who were with practitioners that knew

them? What was the rationale behind that? To me, it just seems sort of strange, but I wondered if there was a reason behind that. Maybe I misunderstood.

Dr. Watson: So, sponsor, please answer.

Dr. Brauer: Sure. Our clinical sites were selected for geographic diversity in the US, and we did have sites representing all four regions, including the Northeast, the Midwest, the South, and the West. And in fact, we wanted to recruit participants from sites where they were receiving clinical care as part of the process. We wanted to do an all comers design study, where patients who were coming into their normal healthcare providers would have the opportunity to participate in our study. We saw this as a good practice with respect to eliminating potential bias in patient selection. Does that, does that answer your question, Ms. Higgins?

Dr. Higgins: Yeah, I, I have follow-ups, but I'll let it go.

Dr. Watson: Oh, please, get all your thoughts out. It's what this is for. I don't want anyone to have any lingering questions when we finish this hour.

Dr. Higgins: I look back to the FDA's analysis of the study, and I just took note of the fact that there were three sites that had sensitivity and that were sort of shown to be reliable. And I'm wondering about, for some reason, when I read the packet of information, it struck me that there was some geographic bias, and that's all I'll say, towards the east.

Dr. Watson: Thank you for that comment. We have three more hands up, so I'm going to go Dr. Walker, Dr. Farrar, then Dr. Dunn. Dr. Walker?

Dr. Walker: Thank you. I wanted to address two, two things. One is, I would like to reinforce what Dr. Bierut was saying. If we could know what someone's risk was, and if we

could know someone was at higher risk than someone else, I don't think we would not want to put that into the clinic and let the clinic put the guardrails around it. Let the community put the guardrails around it. I also feel a little less concerned that the community will know what to do with this information if it is, in fact, truly correct information. But this comes back to my — I did have a question about the sponsor's answer about the genetic ancestry, because I think I understood that that was queried at the front end during the machine learning process, but that the actual test, as it's designed to now, the actual algorithm has not been challenged, was underpowered to even look at race and ethnicity. I find it hard to believe that it could have actually been challenged with a genetic ancestry, but I just would like to hear from the sponsor that that understanding is correct.

Dr. Watson: Sponsor?

Dr. Donaldson: Thank you for the opportunity to clarify this, because I do think it's one of the major important factors in this entire discussion. So, you're correct, Dr. Walker, it is incredibly important, as people have talked about here, is to correct for ancestry where there's a genetic difference in allelic frequency. A hundred percent agree. There's multiple ways to do that. Our approach represented or overrepresented ancestry to correct for the additional allele frequency distributions in the learning and training data sets, both of the 3000 discovery patients that we used, the 1300 patients that we use to set the final parameters and lock that in. You are also correct that in the clinical study we did not use ancestral information because the adjustments for ancestry were baked into the model and it's not necessary. The last thing that I want to correct, which I think I've stated before, so I apologize if it was stated slightly incorrectly. There's a lot of questions coming this way, and a lot of them are fairly high level.

The best way to determine if AvertD classifies different from ancestry we evaluated. And the way that we did that is we took the AvertD classifier, the actual tests that we have that's locked, that's already developed, and we deployed it into several large in silico data sets derived from what most would say are consensus allelic frequencies across the world. I mentioned that. We said, do you get different distributions in AvertD results by ancestry? Overwhelmingly, the answer is the distribution of the AvertD test results by ancestry are equivalent with one exception. Right? That exception is if a patient or if an allelic frequency is drawn from an African or an African American group, instead of seeing a distinct peak like the other groups, we tend to see a more muted plateau of the distribution of the scores. A broader base peak. This is research, it is not clinical data, but it was drawn, and I'm highlighting it on the slide now, from simulated populations of 20 to 30 million individuals.

Dr. Walker: Thank you.

Dr. Donaldson: I think it's an incredibly important point. I'm going to put it up there and it does answer, I think, the question that you were asking as well as the questions Dr. Bateman was asking before, is, given the limitations of everything out there, how do you determine how this test will perform in an ancestral diverse population? And this is how we arrived at doing that, is using the best information on the allelic frequencies over numerous different ancestral groups looking at the distributions. And there's no difference in patterns overall. There's a slightly broader based peak in the African-American or African groups.

And then the last thing that we did, which may have been missed in my communication as well, is we actually tried to classify people by different ancestries using our result. And the answer there versus the 18 fold difference that we see in between the pre- and post-test

probabilities with AvertD, we only saw a plus or minus two with a different ancestry. So that definitively tells us we are not classifying individuals by ancestry. And I'll just put that up one a second time, because I think I may have gone over it really, really quickly in my presentation. So, it'll just take one second for you to see it.

So, this is slightly different than the question of the distribution of score results by ancestry. This is, does AvertD, our test, classify by ancestry in different populations? And what you're seeing in the second and third bullet points is there's no real way for us to distinguish who is African in origin versus who is Asiatic in origin from an ancestral perspective with our test results. And this is the best way that we know how to do this given the current information. Thank you for the opportunity.

Dr. Watson: Okay. Thank you. next, Dr. Farrar, did you have questions, comments?

Dr. Farrar: Yeah. And if I could ask a real quick question of you before I do this, are we going to, in our voting or discussion of the question that we're being asked, give our justification for it at that point?

Dr. Watson: After the voting at the end, they're going to ask everyone to sort of justify their answer after the vote is done.

Dr. Farrar: All right, I'll keep this shorter. The first thing is just that clearly the opioid use epidemic is a huge issue, and we really need to get on, and the presentation by the patients was very moving. And having a test would be wonderful, but I would argue this is not it. As I have sort of harped on, the positive predictive value here means that we are going to be wrong when a positive test is achieved. You're going to be wrong 96% of the time, between 92 and 96,

depending on what you think the prevalence is. So, you're going to be labeling 95, let's say, percent of patients incorrectly using the positive test.

Dr. Walker: Could you explain a little more about where you get that number from? I have to be honest, it's not obvious to me how that number is derived, John.

Dr. Farrar: Sure. Yes. The issue is that sensitivity and specificity are clearly very important, but the sensitivity is and its ability to predict is dependent on the prevalence of the disease. If the prevalence is 50%, then having an 80% sensitive or specific test can be a very useful test. When you're talking about something with a prevalence of 1% in our population, or even 5% in our population, then the sensitivity of 80% results in a prediction of 3.9%, meaning that for these, 96% of them are going to be wrong. It's going to predict correctly in 3.9 of the patients and who get a positive test. And I'm happy to — I teach a class on this, so I apologize. I don't want to be lecturing about it. But the point is that the sensitivity/specificity is not adequate when you're talking about a test that is going to test for things that are relatively rare. Low positive predictive values can be used, but they're always used with a secondary test. Take, for instance, the home HIV testing. Doesn't do terribly well but does well enough to send you to go get a PCR test or something else to know it. Warfarin, all of these are used with additional testing to supplement them. And the large NP, negative predictive value, as I said, really is no better than not doing the test. We know already that 95 to 99% of patients are not likely to have this, and so they only do slightly better, and we need to keep that in mind related to the fact that there are 40 to 50 million surgeries in the US every year. And so, the number of possible incorrect answers is huge. I want to commend though the sponsor in doing a very interesting

study and outstanding study, but for all of the various reasons that have been raised by my colleagues, there are a lot of study issues related to them, and I won't go through those.

I do want to speak though for two seconds about the clinical issue, which is that we're already doing, or should be doing, opioid sparing processes in taking care of our patients. And the fact that we don't is a real travesty. And the fact that people with third motor extraction go home with any opioid ought to be close to malpractice, frankly, because we have very good evidence that the non-steroidals work outstandingly well. If the patient cannot take the non-steroidal, then giving an opioid is reasonable, and knowing what their risk is very important. But I think what we're talking about here is that having the good test would be really, really good, but that this test is not going to improve what should be our standard of clinical care, and it's going to end up with stigmata of positives when they shouldn't be. It's going to raise risks about what happens to people who are negative. It's probably only explaining a little bit or a portion of the genetic risk, mind the socio-environmental risk. I'll stop there and simply say that I think that it's a great thing to undertake and we need to get a lot better at it before it's going to be useful.

Dr. Watson: Okay. Thank you. We have two more hands up. I'm going to take those, and then I'm going to go very rapidly around the table to see if anyone else in the panel has additional comments. Then I know I, as well as many of you, need a bio break. So then we will take a bio break. All right, Dr. Dunn.

Dr. Dunn: Hi Walter Dunn. Thanks. I wanted to comment on two reoccurring themes. One, that more information is better when you're assessing risk. It was often mentioned that informed is better than uninformed. I think my concern is what happens when physicians or patients are misinformed, because if they're incorrect interpretation of this test — and this also

relates to how effectively physicians can be educated to incorporate this information in their management of pain. Even in the best case scenario where this test is actually detecting some increased genetic risk is presumably only a small fraction of that genetic risk, right? Yet all of us today, sponsors and panelists alike, have been using the terminology high genetic risk, right? Or even worse, the patient is a high risk. So, to Dr. Ness's point, a more accurate characterization and perhaps terminology may be, quote unquote, "the presence or absence of a signature of risk in the reward pathway," right? So that we move away from the implication that this test is assessing all genetic risk.

So, to the point about physician education, if you look at the figures from the sponsor, one of the figures shows that genetic risk contributes to 50% of developing OUD, but that doesn't apply to this test, right? This test is only a, again, best case scenario, very small fraction of that. Yet, when you're presenting that information, I think the easiest and immediate interpretation is that "Oh, I'm going to be capturing 50% of this patient's risk with this test." So even if you listen to some of the public comments from the physicians and surgeons, right? One of them called this a simple, actionable tool. I think this is complete opposite. This is not a simple tool. I think it's a very nuanced tool that I don't know if we're able to convey that nuance to be applied in a meaningful and thoughtful manner in the clinical setting. Thanks.

Dr. Watson: Thank you. Dr. Bateman?

Dr. Bateman: I'm still struggling to square the results from the sponsors' study with the GWAS study. The odds ratio from the sponsor study is 18 for these 15 SNPs. And even if you think about interactions and additive effects, multiplicative effects, the fact that when the investigators look in samples of 20,000 patients with Opioid Use Disorder, 600,000 controls,

only a single one of these SNPs emerges as a significant predictor, and there with a very weak effect. It's just hard to reconcile and makes me really concerned that there's confounding here that's unmeasured and unaccounted for in the analysis. And I take the sponsor's point that they are looking prospectively. They're looking at opioid exposure in an opioid naive population and then the development of Opioid Use Disorder.

But even if you think that's a subtype of Opioid Use Disorder, more generally, you'd expect there to be common pathways to Opioid Use Disorder. And the fact that these very carefully conducted genome-wide association studies are just not seeing the strength of effects that are observed in the test, it is, to me, quite concerning.

Dr. Watson: Would you like the sponsor to respond, or was that more a comment?

Dr. Bateman: I mean, I'm happy for them to respond. I think we've heard their response points. But if, if they want to.

Dr. Watson: Sponsor, do you have any additional response?

Dr. Donaldson: I would love to just share one additional point. Dr. Bateman's struggling through comparing very, very different studies using different methods in different populations. The statement that he made to begin with is that he's conflicted between taking study results done in one way and one population and extrapolating them to another. That is not surprising, okay, so I'm not conflicted by that. And I'll just use one example. So, a very large study that has imputed SNP designs, which are common to GWAS, meaning they don't measure the genetic signature, they impute it. An acceptable number for an imputed snip in terms of reliability can be in between 5 and 25%. If I misclassify 5 to 25% of my genetic signature, no matter how big my population is, I will not be able using machine learning to accurately develop a classifier. So

that's just one example. I know it's an oversimplification, but I'm trying to help you understand how the differences in methodology and veracity here can drive two very different results.

Dr. Bateman: Yeah, I take your point. But you're going from an odds ratio of 18 to essentially a null.

Dr. Donaldson: It's a very good point. So, so let me take you to the next part, and I'm not being argumentative, I just want to make sure I convey this because I know we're pressured in time. The other thing, the way in which signals are detected in GWAS, the math behind that, where you're saying they find a significant association, although a low effect with a particular target, and I believe the one that you're referring to is OPRM, right? That particular target, the way the signal reduction happens in a GWAS analysis is distinctly different than a machine learning environment. It is two different pathways, and we're trying to reconcile things that are true and true, meaning that you can't, to date, in the largest GWAS studies, using the methods and study design they have done, you cannot identify genes that have popped above that significance level. And so how that's addressed in GWAS is you increase the study sizes, and you increase the populations, which is what normal GWAS studies do. It's sort of a meta of meta of populations. In machine learning, the approach is exactly the opposite. In machine learning, what you do is you define a very clean phenotype in a defined population. You take that phenotype, and you train the hyper parameters of that phenotype on a larger population, more diverse, and then you blindly deploy it. It is two very different approaches, and I appreciate you're struggling through it because it is very complex.

Dr. Bateman: One quick follow up. Are there examples of machine learning based approaches applied to candidate genes that have high predictive value for other complex traits?

Dr. Donaldson: There are. There are, but these have been done in, I mentioned one in schizophrenia. So how this works over time is, in general, they clarify the phenotype in the GWAS population. When schizophrenia started to do genotypes, it was very broad. Everybody was either positive or negative in terms of a schizophrenia outcome, where they started to see better predictive values using machine learning techniques as well as others, as where they defined early onset schizophrenia, and even with much smaller patient populations. All of a sudden, because the phenotype is much clearer, you start to see those targets pop up both from a GWAS perspective as well as from a utility of a machine learning perspective. And I know this is a short time for this type of data, but I do appreciate the repetitive questions. I really do.

Dr. Watson: Okay. We are 30 minutes late for our break. So, we can either continue on now and die, like I am, or we can take a five minute break now, come back, and finish. I see you, Jennifer. We will get your question, and we'll go around the table. So is that okay with the group if we take a five minute break now, come back, finish all the discussion and comments so everyone is fully satisfied that everything has been covered. Is that okay with the group?

Dr. Walker: Karol, no one wants you to die.

Dr. Watson: I appreciate that. So we'll take a five minute break now. Please remember, do not discuss any of the content of our meeting with anyone. Additionally, whenever — five minutes, don't discuss, just come back and we'll finish up.

Dr. Watson: Welcome back. I call this meeting back in order as we finish up our deliberations. Ms. Higgins, please, with your question.

Dr. Higgins: Yeah. I just was thinking about what Elizabeth had said, and I thought about the public health ramifications, and it is a prescription and I'm wondering if there are

access issues. Those who don't have insurance, would they have availability of this test? Is there going to be a have or have not? And if they do have access, will that put a strain on public benefits to pay for this?

Dr. Watson: Sponsor?

Dr. Higgins: It's maybe rhetorical. Just a thought, being rhetorical.

Dr. Watson: Okay. Thank you. All right. Now, we just are going to go around to every panel member and ask if you have any further comments, concerns, questions. Shake your head no, quickly, we'll move on because we are running behind. But if you have any lingering concerns, I want them to get out. But, Dr. Compton, any remaining concerns? Ms. Higgins, any remaining concerns? Dr. Zaafran, Any remaining concerns, Dr. Bierut, any remaining concerns? Dr. Dunn? Any remaining concerns? Dr. Ness? Any remaining concerns? Dr. Ruha, any remaining concerns? Dr. Gallagher, any remaining concerns? Dr. Goldstein? Any remaining concerns? Dr. Walker? Any remaining concerns? And Mr. Wreh, any remaining concerns?

Mr. Wreh: Well, I don't have any concern, but I'm not sure if it's time for me to go ahead and speak on behalf of industry. So, you just let me know when I can speak for industry.

Dr. Watson: I think this is the correct time. Am I correct, James?

James Swink: Yeah, Elijah will be given time right before the vote, as well as the consumer rep and the patient rep. Okay. Dr. Farrar, did I get you any remaining concerns?

Dr. Farrar: The only thing I forgot to say, my earlier comment, was that one of the concerns about a quote “objective test” is that it will gain substantially more traction in the medical environment than the subjective testing that we currently doing. And in the absence of

adequate evidence for it doing a better job, I have great concerns that it may be adopted when it shouldn't be. So, I said the rest of it before.

Dr. Watson: Okay. Thank you. Dr. Wang, any remaining concerns? Dr. Gordon? Any remaining concerns? Dr. Bateman, any remaining concerns? Have I gotten to everyone on the Panel? Great.

Dr. Joniak-Grant: I have one remaining concern.

Dr. Watson: Oh, I'm sorry. Ms. Joniak-Grant.

Dr. Joniak-Grant: Yes. I think my other concern I just wanted to mention briefly is that there's a lot of talk about the importance of training and education but no details. And I was like, "Oh, let's look at the package insert." And I went to it, and it was basically a paragraph with a lot of statistics that I don't — you know, clinicians aren't statisticians, so they might not always know what exactly those numbers mean. And so, I think if something like this were to go forward, I would really want that information spread out, because there's lots of issues that arise in clinical communications, and then them actually becoming actionable on the ground. And so that was just one thing I wanted to raise. Thanks.

Dr. Watson: Great. Thank, thank you. Everyone has gotten a chance to voice any concerns, correct? Great. Let's move on to the questions now. At this time, we're going to discuss the questions that the FDA needs our feedback on. So, panel members, electric electronic copies of these questions have been emailed to you and posted to the FDA website. So, I ask each panel member to identify themselves when they make a comment. But please go ahead and show the first question.

FDA: As described in the FDA and sponsor executive summaries and panel presentations, there are several factors that contribute to the uncertainty in whether the observed clinical study results accurately represent the device's performance in the intended use population. For the test, for each of the following factors, please discuss its impact on a) Clinical study subject enrollment and the resulting clinical study population, b) Clinical study test performance interpretation, c) Applicability of the study results to the intended use population. A: Use of different CRF versions during the study to collect the data, including completion of an additional CRF after study completion to support the subjects met the inclusion and exclusion criteria specified in the protocol. B: Confidence with which the study excluded subjects whose index oral opioid exposure was illicit and or for treatment of chronic pain. C: Recruitment of subjects both from treatment sites and from non-treatment sites. D: Determination of index oral opioid exposure based on subject recollection and the additional information available in the medical records and histories at enrollment sites. E: Assignment to a risk pool based on SUD and OUD status, absence of OUD, positive subjects in the low risk pool, and subsequent use of risk pools to select study participants. F: Demographic makeup of the study population with regard to race, ethnicity, age, and sex.

Dr. Watson: All right. can we keep that question up, please? It's very complicated. I need to summarize the panel's feelings about each of these. So, I'd like us just to go around one by one. So start with A. there were different CRF versions used during the study to collect data, so what impact do you think that made on the enrollment, the clinical subject enrollment, the study test, performance, and interpretation and applicability? Please make your comments.

Dr. Gordon: This is Adam Gordon. Can I ask a question about that question that you're posing to us? Are you only asking us to decide or to have an opinion about the different CRF versions, or the broader question of clinical study subject enrollment?

Dr. Watson: So, FDA, I believe we are to take it at what they wrote, but if FDA has any clarification, I would appreciate that. I take it as just the different CRF versions. But FDA?

FDA: That is correct.

Dr. Farrar: This is John Farrar. I think that the difference in the CRF, while not ideal in a clinical study, and clearly can use substantial improvement above what was done in the final one, but that the differences made little or no difference, actually, in the assessment of the data for this study.

Dr. Watson: Thank you. Any other thoughts?

Dr. Dunn: I would echo what John said, just so it's not just his thoughts. I don't think their different CRFs made a difference to what the final assessments were.

Dr. Gordon: This is Adam Gordon. I'm just concurring with that. I think this happens a lot in research studies that sometimes there's modifications of forms. It wasn't a crucial aspect of the data or data analysis.

Dr. Watson: Great. So, the feedback I'll give to the FDA is that the different versions of the CRF were not ideal, but we don't think it substantially affected the results of this step. Is that fair? Dr. Compton?

Dr. Compton: I agree with this perspective. I was going to say that I was also convinced by some of the sensitivity analyses that helped confirm that it didn't make a big difference.'

Dr. Watson: It didn't make a big difference. Thank you. All right, let's move on. Unless anybody has any further comments, we'll move on to the next question. The confidence with which the study excluded subjects whose index oral opiate exposure was illicit and if for treatment of chronic pain. Does anyone have a thought about how that affected the study results? Interpretation, applicability?

Dr. Joniak-Grant: I have a thought. Dr. Elizabeth, Joniak-Grant. I, overall, I feel pretty good about it. I would've preferred if the medical records or the history had been reviewed outside of the enrollment site just to go a little bit further back into their history, especially when people are saying that they were exposed 50 years ago. Not, I don't know if there's those records, but I think that would've just been useful information to help increase my confidence level.

Dr. Watson: So if I'm understanding you correctly, you say you were not completely confident that the study excluded subjects who had illicit exposure. Is that correct?

Dr. Joniak-Grant: I, personally, and others can comment on this, I feel it was fairly sufficient. I'd feel more steadfast in my decision if there'd been more information outside of just the enrollment sites.

Dr. Watson: Dr. Zaafran?

Dr. Zaafran: Yeah. What’s really striking to me about this is that there's too much selection bias in the way the study was designed. And if you're truly trying to make this into something that is going to cover all patients who are in the perioperative setting, then you need to have more of an example up there.

Dr. Watson: Can we just focus on the question of the confidence with which the study excluded subject is in the question?

Dr. Zaafran: In B, it was very selective. And by excluding subjects, it basically made it extremely biased. And my point is, is that it needed to be a random sample as opposed to excluding specific entities or including specific entities. That's my point.

Dr. Watson: But were you confident that those who had illicit exposure were excluded for —

Dr. Zaafran: No, that's my point.

Dr. Watson: No, okay. Dr. Farrar and then Dr. Gallagher.

Dr. Farrar: I also have limited confidence in this, although given having tried to do such studies that it's very hard to get more confident about it. I this what Zaafran was just describing is that you do a much different kind of study design, but that's not what we're being asked here. These patients should have been included and then looked at separately after more thorough data collection, as opposed to being excluded without adequate data collection.

Dr. Watson: Thank you, Dr. Gallagher. So, while I'm confident what they collected is good information and geared toward the marketing of their product, I'm not so confident with the overall aspects because these other patients were left out. And so, we don't have adequate information to say anything about longitudinal effect of whatever is learned.

Dr. Watson: Thank you. Dr. Compton?

Dr. Compton: Well, again, just sticking with your exact question, I'm reasonably confident, and what reassured me was the secondary approach to looking at medical records, at

least to correspond some procedure and some likely prescription at that same time that helped convince me that the exposure was more likely than not, that their work, prescription opioids, that that patient took around that period of time. Like others, there's some questions about this design, but that's separate from the question that FDA asked us.

Dr. Watson: Thank you, Dr. Compton. I completely agree. The question that they asked us was specific to how confident are we that the excluded subject whose index oral opiate exposure was a loose elicited or for the treatment of chronic rather than acute pain. So, I think we have other issues about their exclusions, but for that specific question, I have one person saying fairly confident and others saying reasonably. But for that specific question.

Dr. Walker: People are also addressing that Part C though of the applicability.

Dr. Watson: I really would like us to go one by one. I can't keep them all straight. And so just on B. So right now, I have barely confident that those whose exposure was I elicited or for chronic versus acute pain. Fairly confident in that. Is that fair?

Dr. Zaafran: Dr. Zaafran, I feel like I didn't completely get your meaning correct.

Dr. Zaafran: You're asking me now, right? Yes. Yeah, So, I mean, maybe I just don't like the question, but if you want me to like, answer it

Dr. Watson: I don't like the question either, but that's what we were asked to answer.

Dr. Zaafran: Yeah. So maybe that I have some reasonable confidence, but I have concern about the design. Maybe I'll just say it that way.

Dr. Watson: So, I will actually add that to our, because I've heard that from most of you. I will add that to the end of our question. So, reasonable confidence that those whose

exposure was elicited or for chronic rather than acute pain, that's, we're reasonably confident in that, but we have problems with the design. Is that fair?

Dr. Zaafran: That's fair.

Dr. Watson: Okay. if there are no other comments on B, we'll move on to C. Recruitment of subjects both from treatment sites and from non-treatment sites. Are we okay with that?

Dr. Farrar: This is John Farrar. They presented data, there was not huge differences between them. That being said, there is some concern about it. However, given their need to enhance the population of OUD patients they needed to recruit from the treatment sites, it does raise questions about the potential bias there that we would not be able to analyze. But in general, I think it's okay.

Dr. Watson: Dr. Wang?

Dr. Wang: I'm not particularly excited about that design because I think really biased their selection and led to a final group. And that's totally different from what their intended use group is.

Dr. Watson: Thank you. Dr. Gordon.

Dr. Gordon: Yeah, Dr. Wong just said exactly what I was going to say. I think the issue with me with C the applicability to the study population and the intended use population.

Dr. Watson: Thank you. I think I hear you all saying the same thing. So, I'll give feedback saying that we have some concerns about using those two populations because it may introduce bias.

Dr. Compton: I would add one. This is Wilson Compton, and I'd add one more bit of evidence to that, in that when you look at the severity that was shown, it's sort of the reverse of what you might have expected if this had been a prescription opioid exposed in a prescription Opioid Use Disorder population where mild and moderate would've outnumbered the severe. And we saw the opposite here because of the recruitment strategy. I did like the way they stratified and still showed more or less the same sensitivity and specificity in the different severity; that helped reassure us. But it's not the strongest design.

Dr. Wang: I echo that, but at the same time, I think the sample size in the mild and moderate is so small that I'm not sure you have 100% confidence in the statistical power in that comparison.

Dr. Compton: We're going to get to that whenever we talk about any of the subgroups. This was not powered to look at subgroups, but it is at least suggestive that they didn't see huge changes.

Dr. Watson: But yeah. And I'm pleased that all of you guys are giving pretty much the same feedback, which what I heard was that there are concerns about the use of these two populations because it can introduce bias. And when you look at the data, that it's so heavily skewed towards the severe, rather than the mild or moderate, that suggests that we probably did introduce some bias. Is that fair?

Dr. Farrar: That is fair. Let me say one other thing relative to this. Dr. Donaldson talked about the way in which you generate artificial intelligence learning procedures, and you do need to have enough of the issue that you want to study to be able to do the learning. So, I understand why the need was there and, but it suggests very strongly that they were just missing a whole

group of people as Dr. Donaldson was just saying, and that either another study or some other way of dealing with that needed to be there in order to overcome the potential bias.

Dr. Watson: Thank you.

Dr. Compton: I'm sorry to question that, but Dr. Farrar, the artificial intelligence and the machine learning was used to develop the AvertD test, but it wasn't used to study the 380 subjects that were the enhanced group that we're talking about right now. Unless I'm missing something.

Dr. Farrar: No, you're, you're exactly right. And, and I apologize. Yeah, I was getting those two things confused.

Dr. Watson: Thank you. Okay. if everybody's okay with that, let's move on to the determination of index oral opioid exposure based on recollection. Are we concerned about this?

Dr. Higgins: I am actually. I'll share a personal story. I feel like family members might be able to better corroborate an experience that someone has when someone's using opioid. My mother's experience with her knee surgery was not at all what I witnessed as someone who was not using opioids.

Dr. Watson: Dr. Dunn?

Dr. Dunn: Yeah. This was the one aspect of the study design I had the most problem with. I don't question the accuracy of the patient's recount. In fact, I think the problem is that this retrospective component is actually enriching for patients who have, again, a very strong memory of that incident for one reason or another. I think that's related actually to the reward pathway. So, I don't think it's actually representative of the intended use population. I think

you're selecting for folks on the extremes of the spectrum where they either have a really positive experience with the opiate or very negative experience, and nothing kind of like in the neutral, especially when you're looking at 10 to 25 years back and they talk about the mean incident as — 10 to 25, or sorry, 8 to 10 years. They do go back and look at the medical records, but it's not an exhaustive search, right? They only look a year before and a year afterwards. So then, we don't know if something happened actually even earlier, but I think just the retrospective component enriches for something that may confound its applicability in a clinical use population.

Dr. Watson: Thank you, Dr. Joniak-Grant.

Dr. Joniak-Grant: I think that subject recollection for specific events can be really good. I don't know if I would trust it on number of days, whether people remember if they had it three days or five days or seven days. I also have a little bit of doubt for people many years ago before there was sort of a lot of public understanding of what are the names for opioids. Do they always know that something they got perhaps in the past was in fact an opioid? Was there considerations of codeine for when you had a bad cough? I know that's one thing they used to give us all the time, at least me as a kid, and things like that. And so, I know a lot of times, patients aren't always aware of the generic name and what that would've been. I think that's especially true 15, 20 years ago.

Dr. Farrar: I just wanted to comment asking about the index oral opioid exposure. Maybe I didn't understand some of the presentation, but it seemed to me that the index oral opioid exposure wasn't something that was necessarily that long. Because if it was, then it really

is going to select for people who have a distinct memory of what happened to them and not others, and that would be a real problem.

Dr. Watson: Well, I think they showed a fairly large range of past exposures. Some as long as decades ago.

Dr. Dunn: The mean and media were like 8 and 10 years, right, and it's up to 25 plus years.

Dr. Watson: Yeah. Decades.

Dr. Wang: The maximum number was like 50 years.

Dr. Watson: Well, I hear fairly consistent comments from everybody. Yeah. We are concerned about that subject recollection and what biases that can induce. Is that fair? Okay. If everybody's okay with that, let's move on to E. Assignment to a risk pool based on SUD and OUD status, absence of OUD positive subjects in the low risk pool, and subsequent use of risk pools to select study participants. How concerned are we that that influenced our test enrollment, performance, and applicability? Dr. Dunn.

Dr. Dunn: This was the second part of the study that had kind of concerns about. I don't know exactly how the assignment of the pools may have influenced the outcome. My concern with this was how they define the high end risk of high and risk factors, right? So obviously they spend a great deal amount of time and carefully identifying OUD status. But in terms of the other SUDs or other comorbidities, that was a fairly low bars looking at the at the medical record. And I think just looking at the comorbidity incidents, I think we've already concluded that it did actually did not do a good job of collecting that information, which did ultimately inform, or

partly informs, which risk group they went into. But that being said, I actually don't know. I don't know how to interpret that they use those risk pools as far as you know, the generalizability of the, the study outcomes.

Dr. Watson: Thank you. Other thoughts? Anyone else? So, what I heard from Dr. Dunn is there's concern about how they define the high and low risk status and other comorbidities and don't really know how to interpret the risk pools. So, it's a little confusing. Is that fair?

Dr. Dunn: Yeah, that's fair.

Dr. Watson: Okay. Any other thoughts? Comments?

Dr. Compton: This is Wilson. I would just suggest that that concern overlaps a little bit with our concern about the enrichment by using treatment sites, because that's so confounded with OUD status.

Dr. Watson: Good point.

Dr. Wang: I concur.

Dr. Watson: Okay, great. So, let's go on to F, if everyone's okay with that. Are we okay with the demographic makeup of the study? Like sex, race, ethnicity, et cetera? And I see Ms. Higgins shaking her head now. Do you want to comment?

Dr. Higgins: I know it was not intended to be a subgroup analysis, but as mentioned earlier, if it's going to be used, generalizable and just across the regular mainstream population, I think it should really have done some better subgroup analyses. In fact, I would recommend that, for the next phase, if there were another study, strongly.

Dr. Watson: Dr. Joniak-Grant.

Dr. Joniak-Grant: I definitely have a problem with it being 92% white. I think the other issue that I have with it too is, is kind of what Melissa Higgins was saying, is that I could see where this could potentially be used. Because if it's touted as this objective thing, right, it could be used to say, well now we can get away from these issues that we have with equity and unfair treatment with pain management and we can go to this genetic test. So, I think it's really important with that, that we have all groups represented. Because I think that's where it's probably going to be applied on the ground.

Dr. Watson: Thank you. Dr. Gallagher.

Dr. Gallagher: I'm concerned only in that there's such a health care inequity in terms of insurance ability and things like that for groups of people. And if the study was enriched by going to treatment centers, the likelihood of who was at that treatment center is likely higher to be somebody who's white, who has health insurance to cover those kind of things, et cetera. It just kind of confounds that issue of what are the demographics of the study. And so, I think in that sense it was skewed.

Dr. Watson: So that kind of goes back to our concern over the recruitment procedures as well.

Dr. Gallagher: Right. But I think that recruitment element impacts the ability to get a demographic that is wide enough.

Dr. Watson: Yes. Dr. Farrar.

Dr. Farrar: Given that we've all agreed that the genetic pool that is currently being specified in this test is identifying certainly not all the genetic risk, and maybe only a small part. The fact that there's now socioeconomic status and other issues, I was very surprised. Given the fact that there are lots of things, as was well said, in a number of cases that lead to interest in and ultimate the use disorder. I was surprised that they weren't measured as part of the study.

Dr. Watson: The fact that they didn't report any SES or social determinants of health, things like that?

Dr. Farrar: Exactly. Or sexual preference, or, I mean, you could come up with quite a long list. And all of those lead to increased anxiety, increased conflicts in life that lead to a desire to have an out.

Dr. Watson: Thank you. Dr. Zaafran?

Dr. Zaafran: So, like I was saying earlier, that lack of diversity in the demographic makeup, I think kind of skews it in a way that is unintended. That population is quote unquote 92% white. I don't know what white means. We're all a mix of different races. We're all a mix of different ethnicities. And that's not really delineated out there. So short of actually generalizing the population that's being studied, you're never going to really get it right. And the other thing is that there's no really easy way to delineate populations that have genetic predispositions where we talked a second ago about social determinants of health. There are other social determinants of health that may trigger people who have a genetic predisposition versus those who have a stronger genetic predisposition that may not be affected by social determinants of health. So again, short of having a random sample of the demography out there, you're really going to get a skewed set of data out there that's not really very helpful.

Dr. Watson: Thank you. All right, Dr. Kelm, can I give you our summary of question number one? We believe the different CRF versions were not ideal, but probably didn't make much difference in this study. For B, we have reasonable confidence that those whose exposure was elicited or for treatment of chronic disease were excluded, but we have a lot of problems with their study design and their recruitment procedures. For C, we had some concern because we definitely thought that using people from treatment and non-treatment centers may introduce bias. And when you look at the severity, you actually see that there is a lot more severe patients than you might have expected in a general population rather than mild or moderate because of this recruitment strategy. For D, we are concerned about using subject for collection because we really do feel like it enriches for patients who had either a very good or very bad experience with opioids and not really consistent with just the general population. For E, we are concerned about how they defined high and low risk status because they didn't do a great job of collecting comorbidity data, and we actually don't really know how to interpret the risk panels. And that again, overlaps with their concerns about the recruitment strategies. For F, we are clearly concerned with the lack of diversity. we're concerned that they didn't report anything about SES, social determinants, other things that clearly can affect this. Panel, did I summarize that fair? Okay. Dr. Kelm, is this adequate?

Dr. Kelm: Yep. We thank the Panel and think that is helpful. Thank you.

Dr. Watson: Thank you. Okay, so can we move on to question two?

FDA: Two. Given the device design, in which 15 SNPs that are associated with OUD as well as other mental health and SUDs are evaluated, and the clinical study design, please discuss the following: A) Does the clinical study provide sufficient information to understand whether

the device is detecting risk of OUD specifically or risk of OUD in addition to other comorbidities? B) Does the information collected following initial study completion, for example, Form 3, clarify whether the device may be detecting comorbidities in the clinical study population?

Dr. Watson: Thank you. Okay, Panel. Question A, does the clinical study provide sufficient information to understand whether or not the device is detecting OUD specifically or risk of OUD in addition to other comorbidities? Please discuss. Dr. Dunn and then Dr. Compton.

Dr. Dunn: No because they really didn't do a good job of collecting information about comorbidities. I don't think we can conclude anything from the relationship between OUD and comorbidities here.

Dr. Watson: Dr. Compton?

Dr. Compton: Well, I'm not quite as negative. I think it could have been improved, but I appreciate the way they stratified by the measures they had of comorbidity and documented within those measures, that they didn't find a major a bias. I did point out earlier in discussion the lack of Tobacco Use Disorder, which is a known risk factor for Opioid Use Disorder, or at least has been purported in numerous studies now, is a gap that I'm sort of surprised they missed, since they collected it on so many people.

Dr. Watson: So, in general, are you concerned that they're specifically detecting OUD or maybe other comorbidities? What do you think?

Dr. Compton: I was convinced based on their analysis that they were detecting OUD.

Dr. Watson: Okay. Thank you. Dr. Farrar?

Dr. Farrar: I agree with that. I think they did a reasonable job, but that would echo the issue that the comorbidities were a real problem, leading to my concern about bias in the whole study.

Dr. Watson: Dr. Bierut?

Dr. Bierut: So, overall, I agree that they demonstrated that they're detecting OUD, but I think they were underpowered for us to really be confident that it's only OUD. And what I know is there's a genetic correlation between the substance use disorders.

Dr. Watson: Thank you. Any other comments? So, what I'm hearing is that there is a little bit of disagreement there. Overall, I think most people agree that they were detecting OUD, but because of problems of collecting comorbidities in things, there is less confidence. Is that fair? Okay. All right. Then question B: does the information collected following initial study completion in form three clarify whether the device may be detecting comorbidities in the clinical study population? Kind of a similar question. Please discuss.

Dr. Ruha: This is Mrs. Ruha. I felt pretty comfortable that the device was not detecting comorbidities, just my general impression from what they presented.

Dr. Gallagher: This is Dr. Gallagher. I also think that they were able to identify OUD instead of, and not just, the comorbidities.

Dr. Watson: Thank you. Any other comments? Okay. Here's what I heard regarding A: that, in general, we think that the study is detecting OUD, but because they didn't do a great job of collecting comorbidities, there is some concern that that might not be the case. But overall, we

think they did detect OUD. And for B, we feel pretty comfortable that the device was not just collecting comorbidities, but rather OUD. Is that fair?

Dr. Farrar: This is John Farrar. I think with the first one, the additional thing that was expressed that is very important is that the study was too small to be sure that there weren't differences between the morbidity components as a subgroup.

Dr. Watson: Thank you. That was important. Okay. For A, I say that in general you feel confident that they were detecting a risk for OUD rather than others, but they didn't do a great job of collecting comorbidities, and the study was underpowered because it was small. And for B, we feel pretty comfortable that the device was not just collecting comorbidities. Is that fair? Okay. Hearing no dissent. Dr. Kelm, did we answer your questions?

Dr. Kelm: I do want to clarify. I think in some ways I was thinking ticking what everybody was saying and thinking a little bit from the other side though. If we actually start with the fact that they have not powered the study with sufficient number of comorbidities to actually make that distinction… I mean, I acknowledge that there was a sense that they had enough OUD, but would the Panel say you can still determine that given the lack of numbers in the study?

Dr. Watson: That's a good question. Panel?

Dr. Farrar: Right. This is John Farrar. I think that the point was they showed data to suggest that the sensitivity and specificity were nearly the same in the various comorbidities that they had. But because the study was relatively small and the subgroups, therefore, were not large enough to really be able to detect differences, we would argue that it would need to be repeated in a larger study to be confident. Is that fair?

Dr. Kelm: Sure. So, the small numbers means the confidence interval's not significant. But obviously, the trend shows that this might be true if we had a larger data set.

Dr. Watson: Promising but needs confirmation.

Dr. Kelm: Okay. Thank you. I think that helps. Thanks.

Dr. Watson: Then can we go on to the third question?

FDA: Three. The reported sensitivity and specificity of the AvertD test, when tested in the clinical study population, is 82.76% and 79.23%, respectively. The negative likelihood ratio is 0.22 and the positive likelihood ratio is 3.98. A) Does the reported device performance in the clinical study population represent the probable performance of the device in the intended use population? B) Please discuss the clinical significance of the study results including sensitivity, specificity, positive and negative likelihood ratios. C) With the consideration that genetics is only one contributor to the overall risk of developing OUD, please describe the level of sensitivity and specificity that would be clinically acceptable for a genetic risk test for helping to identify individuals at increased risk of developing OUD.

Dr. Watson: Thank you. Okay, for this question, AvertD test in this clinical study population has 82.76 and 79.23% for sensitivity and specificity, respectively, the negative likelihood ratio is 0.22, and the positive likelihood ratio is 3.98. Does this reported device performance in this clinical study population represent the probable performance of the device in the intended use population? Please discuss.

Dr. Goldstein: I think it's impossible to know. I think the trial that was done doesn't represent what the intended to use population is going to look like in the long run. So, this is

going to be used in hospitals and private practice and all over the place if it has deep market penetration, which is what the sponsor hopes for. So, I think there's no way to know here.

Dr. Watson: Thank you. Dr. Farrar?

Dr. Farrar: I think I've said probably more than people are willing to hear that the problem here is in the positive predictive value, and that this level of sensitivity and specificity for a disease entity that we're trying to detect that is on the order of 1 to perhaps 10%, is simply inadequate. I think it is possible that the likelihood ratios are appropriately defined here, but with those likelihood ratios, the ability to predict in a way that is going to be potentially useful clinically, I think is simply not adequate. And that clearly, they need to do better.

Dr. Walker: I just want to make sure I'm following the argument correctly, which I still may not be. Is the argument that, because our prevalence is so low, you're going to be wrong a lot of the time. But at 82%, even if you were at 99.99, wouldn't that argument still hold?

Dr. Farrar: The answer is yes. With a very rare disease, even if you were at 99%, you would still not do terribly well, but you would do a whole lot better than you do here. There are substantial experience with looking at HIV, for instance, and using a high eye sensitivity screen first, but not telling the person they're positive until they've done a very, very highly specific test. The issue is that although genetics, a single gene, can be very, very specific, right? Huntington's, chorea, you're very specific about that. A mix of genes like this is another diagnostic test with all the problems of diagnostic tests in needing to be very specific.

Dr. Walker: Well, I would say your argument is obviously a good one, but I think saying that that 82% is not good enough when 99% will not be good enough, I think we're setting

the bar at an unattainable goal. And so, I just wonder if that's the right way to be looking at this particular decision.

Dr. Farrar: I'm sorry, maybe I misspoke. This is John Farrar again. 99% would not be a terrible process for this. But the whole point is that it probably requires more than one step in order to be useful. And what was said during the presentation was that it would be used in clinical decision making and consultation, and then they showed us a slide that showed that all of the other measures we have don't do any better. And so, it is a high bar, but I think that the problem is that we need to be able to explore this in a way that will hopefully understand it more broadly. And, like Dr. Wang has said with regards to the Warfarin method, there are ways to expand the probability of being correct by correlating these with other findings to make the best diagnosis possible.

Dr. Walker: All right. May we have Dr. Bateman?

Dr. Bateman: Yeah. Could I just ask Dr. Farrar some clarifying questions? I take your points. If this was a diagnostic test to diagnose cancer or some pathology, I think the bar you set makes a lot of sense. But if we're envisioning this as something that's used in risk stratification as one component of a multifaceted kind of approach to risk stratification, would you have the same interpretation of what the threshold should be?

Dr. Farrar: So, I guess I should answer this. The issue is that it is being used as a screening test. And a screening test has to meet substantially higher bar than a diagnostic test. And it's not just the sensitivity, just to be clear, but the specificity also is small. And if you were to increase both sensitivity and specificity to 99, you would achieve an 84% prediction. If you went to 95 for both of those, then you would be at 50/50. So, it's a hard bar, but I think it's an

important one to keep in mind when trying to decide how to label people as likely to have an Opioid Use Disorder, or in the future.

Dr. Walker: Okay. Dr. Wang.

Dr. Wang: Thank you. This is Ping Wang. So, I think this could be approached in two different ways, right? One way is that, in the clinical toxicology world, I think the usual way we screen is to screen with something with very high sensitivity, but maybe modest specificity, and then follow that with another test, with a second tier test that has both high sensitivity and high specificity. But the first level of test, which is a screening test, and which is high speed, will probably give the provider, or whoever needs the information, for example, emergency room physician, to act right away. But that has to be supplemented by the second tier, which will either confirm or not confirm the screening test. So, this is one way to approach that. And the other way is possibly, like the Warfarin case, maybe genetic test by itself, even if it's 99% sensitive, is still not going to generate a high enough positive predictive value, then you really need to find a way to supplement that with a clinical and demographic information so that you derive a composite method, rather than just to rely on a genetic score itself to generate a high enough performance test.

Dr. Watson: Thank you, Dr. Wang. I hear Dr. Wang and Dr. Farrar basically saying the same thing, which is really, in general, just general principles about how we treat screening tests. Not specific to this, but in general, how we treat screening tests. All right. Dr. Compton, do you have a thought?

Dr. Compton: Well, I want to have a slightly opposing view here. I don't see how this isn't a screening test where we have a gold standard, because we're trying to predict the future for

these individuals, and we won't know it in some cases for 10 years. So, I don't see how we can apply those same standards in this case. We'll never make progress. I also think that one way around this conundrum may be in the language of how the results are described. We have this discussion around changing the word low risk, normative risk. The word high risk is awfully generic as well, and seems to imply really something awful, where if you tell me you got about a 15% chance of an outcome, I'm going, “Oh, okay, well, I have 85% chance not…” It immediately provides a more nuanced outcome. I would suggest part of our solution here is modifying the description of the outcomes of the test if it moves forward.

Dr. Watson: If I can reign this back into the actual question she asked, does the reported device performance in the clinical study population represent the probable performance of the device in the intended use population? What do you think about that, Dr. Compton?

Dr. Compton: I think it probably does. I also have a question that I'm hopeful that, in their last comments the sponsor will address, which is: they powered this to detect something, if I remember right, something in the 50 to 60% range for sensitivity and specificity. And they chose those on purpose, and they exceeded them. But what made them pick that? Because that's nowhere near what an ideal test would look like.

Dr. Watson: Of course. Yes. Mr. Swink, is it appropriate for the sponsor to weigh in right now?

James Swink: Yes. We can give him a few minutes to address it.

Dr. Watson: Oh, okay. Thank you. Sponsor. Please respond.

Dr. Donaldson: I would thank you for the opportunity to talk about this. In fact, Dr. Compton's interpretation is one that we agree with. This is not a screening test. If you're interpreting this like an HIV screening test where you have a 99.9% negative predictive value, this type of test will not meet that metric. So, I just want to make sure that you understand what you're judging against. This is a risk assessment test, not a screening test.

The other comment, before I answer the question on the performance goal setting, because it feeds into these other two very distinct issues, is that Dr. Wang's comments in terms of a preferential development of warfarin from when the genotype first developed over time, to where now warfarin genotypes are used in a complex model that can give dosing, took years, right? So, the genotype, identifying people that may be at increased bleed risk started, and then, over time, that genotype was incorporated into predictive modeling with other pieces of information. So, the portion of that that I want to make sure everybody on the committee understands is, right now, what we're talking about is that predictive phenotyping of the genetic test and, over time, how this interacts in the clinical models or clinical care deliveries needs work. So, it's not necessarily a fair assessment to say warfarin, although it is a genetic test that has a similar performance in terms of post-test probability or positive predictive value, you have to understand that it takes time to incorporate as medical practice changes this type of information into a complex complete evaluation.

The last thing that I wanted to say here in terms of setting the performance goals, this was a dual performance goal. So, and the dual performance goal was set at the intended, and I'll show it right now — the dual performance goal was set at threshold based upon earlier studies, but it also was purposefully set to give a good combination between the sensitivity and specificity so

that we could derive a clinical difference in populations where it's not around today. So, the thresholds were based upon earlier studies that we were performed. There were no previous methods available to set a comparable performance, which is something that you would normally do. So, we set the performance goals at preliminary study algorithm estimates of sensitivity of around 76% in specificity, around 72%. We then minus what a normal interaction with regulatory agencies would do to set a performance threshold goal. And that got down to the specified pre-study performance goals of 59.5% and 55.2%. Of note, the likelihood of meeting both of these goals, even though they're low when you do it in combination as it was done in this study design, ensures the overall quality of the study.

The other thing I would say is the observed performance was much higher than the pre-specified performance goal. So, I thank you for your time. I appreciate it.

Dr. Watson: Thank you. Okay, Dr. Farrar and then Dr. Dunn.

Dr. Farrar: Just to make a comment, whether we consider it a diagnostic or a screening test is a matter of how broadly it's going to be used. My sense was that the company would very much like it to be used in every acute surgical procedure, in which case it really becomes a screening test. But we can talk that through at some point. I think the other point to make is that if you can improve the pretest probability, for instance, ancestry, other addiction related activities, and get that to the point where it is sort of closer to 20%, then a test that has 95% or 90% sensitivity/specificity will perform much better, but the current levels will have difficulty even at that level.

Dr. Kelm: This is Kellie Kelm at FDA, and I just wanted to weigh in, just to be careful because it isn't a screening test and it isn't a test that's for diagnosis. It is a risk prediction tool.

Obviously, the idea is to try to identify patients either at increased risk, or we can argue at not increased risk, of OUD if they were exposed to opioids. So, it's sort of a flavor of its own. And I think you can still look at the math, how you want to look, and how well 90%, rounding up, or 80%, sensitivity and specificity with this prevalence makes sense. But obviously, when something is a risk predictor and not screening, it's not going to have a follow up test, and they aren't wording it that it is a diagnostic test, then, I guess one of the things to continue to think about is, how would you guys think about that as we move along with these questions, with that sort of question built in about how someone should be using it. Hopefully that helps.

Dr. Farrar: If I could just make a very quick comment. The point is that I have a significant concern that it will be used as a screening/diagnostic test. You can call it anything you like, but if you increase the chances of the risk with a socially very hot topic of whether you're likely to become opioid use dependent or whether you are, I think it just will require a higher bar, even though you're right that just a risk assessment by itself could meet a much lower bar. I just worry very much in terms of your second question about clinical applicability is that it will be used as either a screening or diagnostic test.

Dr. Kelm: Well, maybe we can get to this sorry Kellie Kem now or later Karol, Dr. Watson. But one of the things I guess we could ask is, if that is a concern, do you have thoughts on how to present this information so that, if it is to be marketed, how that should be clear so that we can do the best with a risk prediction tool and keeping it in the right place as best we can.

Dr. Watson: Okay. Let's take a little side detour and address that question. How can we best present this information so that it's used appropriately as a risk prediction tool rather than a

screening test or a diagnostic tool? Anyone have thoughts on that? I mean, some of the things we heard was getting rid of the word high risk and things like that. That makes sense.

Dr. Kelm: Dr. Watson, we could even get to that with question six possibly, which talks about mitigations and labeling, but obviously, we could also put it under that tent if people want to weigh in on that.

Dr. Watson: You're right, we should stay in the order we were in, I'm sorry. And we were about to hear from Dr. Dunn and Dr. Joniak-Grant. Dr. Dunn first.

Dr. Dunn: Hi, Walter Dunn. In regards to this conversation about diagnosis versus screening, my very basic understanding is that whether you run the test or not, in a traditional sense, that's not going to change your diagnosis. That's not going to modify your diagnosis. Right? But this, as a risk calculator is going to modify your management and having the result of the test should actually modify your risk after you get the results. So, to the original question about if this is going to operate in the intended use population, absolutely not. Because this current study does nothing. Or you don't see anything about how management is going to be changed. In reality, when this test is run in the clinical population, prescribers are going to modify their behavior, and perhaps patients will also modify their behavior so that they decrease their risk for developing OUD if they test positive. In that sense, we have like, no idea, but I think that's the ultimate goal of, hopefully, this type of test. So, I would say that the conversation about sensitivity and specificity and diagnosis and screening, I don't know if exactly applies to this. It's kind of a risk calculated hybrid type of intervention product. I think that will eventually kind of play out in clinical use. But from that perspective, that's my kind of two cents into how to look at those numbers.

Dr. Watson: Thank you. Dr. Joniak-Grant.

Dr. Joniak-Grant: So, regarding the specificity and sensitivity, I do have some concerns about how different the numbers are for people over 65, particularly women, where specificity drops down to 45%. There's some large confidence intervals going on there, too, with various age groups and, and sects where the lower ranges go as low as 41.9%. And there was a really wide confidence interval for non-white specificity. So, to say how this would play out in the intended population is very difficult for me, at least to determine based on what seems to be a lot of variation happening with small numbers on top of it, right? So, I can't really judge some of the things like p value based on the small numbers.

Dr. Watson: Thank you. Yes. And Dr. Farrar.

Dr. Farrar: Yeah, Dr. Donaldson made a comment that it took 10 years to make the Warfarin test usable. And while I'm sympathetic to the concept that it takes time to do that, I would argue very strongly that at least a movement in that direction is absolutely necessary. They said this will be used in combination with clinical assessment, right? But there was no comment about or thinking about how that would actually be implemented into the question from the FDA. Those things need to be established and tested before I would think about approving this. At least that's my opinion.

Dr. Watson: Okay. Thank you. FDA, can we put back up your slides so we can have the question? Great. So, for A, here's what I'm hearing, and tell me if I got this right. Does the reported device performance in the clinical study population represent the probable performance in the intended use population? I think the answer I'm mostly hearing is we don't know. There's no way to know, because the population was selected, and we think that we just don't know how

it's going to be used in real life. We don't know how it's going to make patients or providers modify behavior. We just don't know. Is that fair, Panel?

Dr. Farrar: Yeah. Makes sense.

Dr. Watson: Dr. Kelm, is that acceptable?

Dr. Kelm: Yeah, that's what I heard. Thank you.

Dr. Watson: Okay. Question B. We were talking about, discuss clinical significance of the study results, including sensitivity and specificity, positive and negative likelihood ratios. We had a lot of discussion about that, and discussion kind of went all over the place from with a screening test, these numbers are not nearly good enough to make it really useful; with others saying, well, this is not a screening test. This is a decision aid, a risk assessment test, so you can't really use sensitivity and specificity. Others noted concern about the varying confidence intervals in certain subgroups, like women over 65 or non-whites. So, I'm sorry that that's so all over the place, but I kind of heard comments all over the place. Any thoughts about how we can encapsulate that, Panel? Oh, I see Dr. Gordon, do you have a thought?

Dr. Gordon: Yeah, this hasn't come up yet. We're talking a lot about the population, but I'm actually interested about the providers. Many of the providers in this study actually had acumen with regards to addiction, were actively treating OUD or had some experience treating OUD. In the intended use population, you're not going to have those providers around necessarily. They're going to be in the preoperative clinics. They're going to be in primary care doing preoperative evaluations. And I worry a little bit about the acumen of any of these sensitivity/specificity issues because you may not have a gold standard to go back to when

someone's trying to make an assessment whether that person potentially could have OUD. So, I just want to point that out. Thanks.

Dr. Watson: Thank you. And Dr. Farrar?

Dr. Farrar: Yeah. I want to talk about the clinical significance from the following perspective, which is that we currently already should be taking care of every patient by using the minimum amount of analgesics necessary, primarily non-opioid, and, when necessary, opioid at a minimum amount. And careful monitoring of those on opioids as we move forward, not only in the acute postop period, but beyond. The question I have is how will this improve that? And it seems to me that, if you get a positive increase in risk, you ought to treat patients by giving them non-opioid and giving them opioid only if they really need it. And if you get a negative one, there's a 20% chance, at least, that you're wrong. So, you probably ought to give them another drug and an opioid only if they need it. I have trouble seeing how this is going to impact care in a substantial way. What I'm concerned about is that people will get a positive diagnosis and then be refused opioid even when they desperately move it.

Dr. Watson: Thank you. So, discussing the clinical significance of the study, including sensitivity and specificity, positive and negative likelihood ratios. So, Dr. Kelm, we're all over the place and I didn't really know of a good way to give you one simple one liner. It seems like there's just a lot of differing opinions on this. I actually really like the question of, it's not a question you asked us, but, if we should already be using opioid-sparing procedures, how will this test change, or how should it change behavior? I don't know. But for B, should we try to get you a one-liner, or are you okay with what we have?

Dr. Kelm: You know what, we'll take the expert opinion of the Panel, you know, acknowledging that you know, at times there may be differences of opinion. So, thank you very much.

Dr. Watson: Great, thank you. Because I was going to have a hard time finding a one-liner, so thank you for that. Okay, let's go to C. With consideration that genetics is only one contributor to the overall risk of developing OUD, please discuss a level of sensitivity and specificity that would be clinically acceptable for genetic risk test to help to identify individuals at increased risk of developing OUD. That kind of goes along with a lot of the discussions we've already had. So, I don't know. You guys come up with some new thoughts. Dr. Joniak-Grant?

Dr. Joniak-Grant: As a patient, and some of my patient community members, I think we'd like the predictive power to be more than a coin toss. That would be much more useful, I think, in kind of understanding what that means and what that would mean in terms of making really truly informed decisions.

Dr. Watson: Thank you. Dr. Bierut?

Dr. Bierut: This is Laura Bierut. I struggle with this issue about sensitivity and specificity. I think of it as, do we understand who's at greater risk and lower risk, and prevalence of the disorder comes into this. And I think it's very hard to give a specific level of what this risk is. What I do think is coming and, the science backs it, is we can identify people who are at higher risk and at lower risk. And people want to know if they're at higher risk or lower risk. We also need to think that there's patient behavior that may change with this. So, I think this is coming, and we're struggling with what, what these numbers are. But I like that we're struggling with it and that we're thinking about how it's going to be implemented at some point.

Dr. Watson: Thank you. Dr. Ness.

Dr. Ness: A lot of what our issues are is because we're establishing a categorical high risk, low risk, as opposed to simply reporting a number and saying what this literature would support as a number that says, “This is a higher risk; when you cross a certain threshold, it's a lower risk.” I mean, we can mitigate all of our concerns. Sensitivity and specificity is going to be a difference of where you set that number in terms of related to their algorithm and things like this. I just harp back to my original sort of thing, which is, I don't like to call someone high risk/low risk. The purpose is to give information to the people. Then you're giving them the 15 SNPs and you're giving them a value with associated literature that would support or not support at what levels that value is important. See, that, to me, is when you get these numbers, sensitivity/specificity is all about, where you set your level for categorical distributions.

Dr. Watson: Thank you. Dr. Zaafran.

Dr. Zaafran: Yeah. Even though it may not be a direct correlation, but when we look at cardiac risk factors and we assign a number to it, or when you look at the stop bang number for obstructive sleep apnea, I would look at this as one component of risk stratification. If you put it that way and you concentrate on making sure that it's one of many factors, hopefully that'll prevent it being the only factor. Because I worry that if you make it too much of an objective number, that it'll become the only factor out there, when it really shouldn't be. It's just the risk stratification thing. So, if you really want to make sure that it's honed in that way, then you put it as one of many and showcase that it should be scored as one of many, as you're scoring what the risk stratification is.

Dr. Watson: Thank you. Dr. Gordon.

Dr. Gordon: Yeah. This is similar to my prior question/comment on B. I real quibble with the first part of this question about that it's only one contributor. I think you could get a lot of addictionologists researchers in the room, they would argue a lot about different contribution factors with regards to development of OUD, let alone you're then now asking a generalist or primary care provider, et cetera, to actually make that universal assessment about the risks. I really fear very similar to the prior comment that if, if there's a genetic test, it's a little bit more thought of, more objective than the subjective other stratification tools that are available out there. And I worry about that that will not be a contributing one to a milieu of different instruments. It actually will be the instrument that will be pressed on a patient. Thank you.

Dr. Watson: You mean that it will be given inappropriate weight because it's a genetic test?

Dr. Gordon: Correct. That's one thing, especially with the naive providers that we're dealing with out there.

Dr. Watson: Thank you. And Ms. Higgins.

Dr. Higgins: Yes, I concur with both of the last two statements. It seems like it might be more expedient as well. It might get greater validity in people's minds as practitioners, but also, it's more expedient to just run a test than it is to do a myriad of other kinds of assessments clinically.

Dr. Watson: Thank you. Dr. Farrar.

Dr. Farrar: Yeah, I just wanted to comment that I really like what Dr. Ness just said, which is that if we provided a value, let's say between 0 and 10, and the risk at 10 is very, very

high, and the risk at 0 is, is very, very low. And then, allow people to make a judgment about how they think it will affect, gets away from labeling people as potential OUD risk and gives people the chance to say, "Oh, I should look at other things, or I should talk to the patient." I actually like that potential solution for sort of figuring out how to convey the risk in a way that will prevent people from being labeled either high or low, or yes or no, or whatever, and may encourage what we want from clinicians, which is to talk to the patient and figure out how best to handle it.

Dr. Watson: Thank you. Dr. Ness.

Dr. Ness: Just to follow up on that, one of the nice things about having just general statements and supportive literature related to these things is it's easy for that to change over time. They don't have to keep coming back and asking for an approval to now call it positive/negative at a different thing, but they can provide additional information that, actually, it's really 0.37 that is really the meaningful level, or something like that. That's a much more evolving process that then can also incorporate other clinical data into that kind of a report.

Dr. Watson: Thank you. So, I think we're having problems sort of coming up with a simple one line answer to this for the same reason we were having a problem with the others. But I think you've gotten some excellent feedback, and I hope this has been helpful to the FDA. Is this adequate, Dr. Kelm?

Dr. Kelm: Well, as before, obviously differences of opinion and input, and we'll take that back and use it as we continue the review. Thank you.

Dr. Watson: Great. Thank you. All right. We can move to question four now.

FDA: Four. Please discuss the benefits and risks of genetic testing as an aid in assessing the risk of developing OUD following exposure to prescription oral opioids for acute pain. Thank you.

Dr. Watson: We've kind of gone around this a lot, this question, please discuss the benefits and risks of genetic testing as an aid in assessing the risk of developing OUD after exposure to prescription oral opioids. So, we've had a lot of discussion. Does anyone want to kind of encapsulate the thoughts about this question?

Dr. Joniak-Grant: May I ask a quick question?

Dr. Watson: Yes, please.

Dr. Joniak-Grant: This is Elizabeth Joniak-Grant. I wanted to ask everyone on the Panel: the sponsor kept talking about 50%, it contributes 50%, but then there were some articles cited that gave the range of 20 to 60%. And so, I just was hoping for a little feedback of, is this still very highly contested of how much it contributes, what it's really telling us? Or is there some agreement among the community who studies this?

Dr. Watson: Anyone in the Panel want to weigh in?

Dr. Bierut: This is Laura Bierut. I could weigh in on this. When you think about the 50%, you think about that as a population level. So, at a population level, we think that genes and environment are important, and half of the variance is explained by genes. And half of the variance is explained by environment. But when you talk about an individual, we really don't know what that number is at an individual level. So, you could be at very high risk genetically, but if you never take opioids, your risk is zero. So, you always have to think of this interplay

between genes and environment. We know everyone who's exposed to opioids doesn't develop Opiate Use Disorder, and we know that some people get a very little exposure, and their opiate use disorder kind of runs off.

Dr. Watson: Thank you. Dr. Farrar?

Dr. Farrar: I just want to be sure that, in summarizing this question, it's very clear that genetics is the word of the decade. And that when people hear a genetic test, they will ascribe a much higher potential significance to this testing than may be appropriate for this particular milieu of SNPs. And trying to convince both patients and physicians and other caregivers that that is not the case, I think, is going to be important. The issue about the risk is that it'll be overly interpreted. The clear benefit is that, as was said, 50% of this is determined by genetics. Now, I would argue that if you look at almost everything in our beings, 50% is about right, other than, as I've said before, specific genes for specific disorders like Huntington's or the BRAC gene or things that are much more specific.

Dr. Watson: Thank you, Dr. Bateman.

Dr. Bateman: I just want to make the point that, even if the epidemiologic studies have shown that 50% of the population level is explained by genetics, that's not to say that 15 SNPs can explain 50% of the inter-individual variability.

Dr. Watson: And that was my exact same comment.

Dr. Bateman: Yeah. The largest GWAS study, I mentioned this before, they found that it only 3.8% of the inter-individual variability was explained by SNPs that were measured in a genome-wide way.

Dr. Watson: Thank you. Dr. Gallagher.

Dr. Gallagher: Yes, this is Colleen Gallagher. I think that even removing it from the idea of this specific test, I think with the continual growth of precision medicine using genetics to do so much in medicine, that there's a great benefit to having that information for assessing risk and for doing lots of other things. But I think it also comes with some very high risks in terms of people accepting what's being said to them and sometimes overusing or relying on that information to make personal decisions. So, I think there has to be some great care given and good conversation with healthcare providers around what genetic information is learned.

Dr. Watson: Thank you. Dr. Beirut.

Dr. Bierut: Yes. I'm going to answer this question, what I think the summary is generically. So, when we think of genetic testing, the benefit is that the physician could do a better job treating the patient. The patient could have more knowledge about their own risk, and knowledge is empowering. The drawback for this will be that some individuals will be identified at high risk that may be denied treatment. They may be identified at high risk and limit their own behavior to a level that they don't need to. And other people may be identified as normative risk and think that they have a health certificate and use all the opioids that they want. So those are the generic risks and benefits with this. In the field, people know that they have risk of substance use disorder running in their family. They often hear about that. And so, the better we are at actually identifying what is transmitted biologically will be empowering. And yes, it may be misinterpreted, as all tests are potentially misinterpreted. And one of our goals will have to be, how do we improve communication, improve knowledge, at both the provider and at the patient level? But this is coming no matter what.

Dr. Watson: Very nicely put, Dr. Bierut. Dr. Dunn.

Dr. Dunn: Just quick comments. It was mentioned during the sponsor and then also during the Panel Discussions about this is a step towards precision medicine. I would argue that pain providers are doing this already in terms of taking individual histories, looking at different clinical factors, and making a specific pain management approach for every patient. So, this is not something that you need to do to apply or practice precision medicine. I think that's an important thing, and I think everybody here probably appreciates that. But to the patient, that that might not be entirely clear. They may think that, "I need a genetic test, I need an objective measure that looks at my genome to receive precision medicine care." So, I think that's kind of one thing to point out.

The second thing about potential risks: the intended use is fairly narrow, right? This is an acute pain situation from the first prescription, but this result is going to stay in your medical record forever. Technically, once someone's been exposed to the opiate, any use of this data afterwards for subsequent episodes requiring pain management is off-label use. And we have no idea how that applies in those situations. So that would be another concern I would have, that, if at all possible, that once someone's exposed to an opiate, this thing comes off their medical record because it can no longer be used on label. But I suspect, physicians will be using this off-label, in perpetuity, is one concern.

Dr. Watson: Thank you. Dr. Compton.

Dr. Compton: I really appreciated Dr. Beirut's summary of risks and benefits. I would want to particularly emphasize the potential risks in prescribing and patient behavior at the normal or lower risk range, where that idea of that possibly being a carte blanche is a real

concern to me that we don't know the answer to, because no one's done this yet. It suggests to me that, at a minimum, if this does go forward, that some post-marketing evaluation of how prescribing patterns and how patient behavior changes based on the implementation of this study would be incredibly important.

Dr. Watson: Thank you, Dr. Joniak-Grant.

Dr. Joniak-Grant: Thanks. I think there could be really great benefit if the results are quite accurate, right? It would be. I do agree that good information, contextualized information is empowering. I don't think we're necessarily there yet, but I do think, just based on what we have, there are real risks for patients for under-prescribing. What kind of doors would this open up for legal liability? I have a patient whose result is that they're high risk. I still gave them an opioid, and now they develop Opioid Use Disorder. We already know and see that legal liability is at the forefront of a lot of clinicians' opioid prescribing decisions. We're seeing that throughout the nation. I'm also concerned about delayed treatment if this has to happen before treatment. A lot of people were talking about postoperatively, but the sponsor themselves said this is for acute. So, if you have to take the swab, send it overnight, it then takes, they said, 24 to 48 hours for the lab to get the results done. Do you have a patient in an acute situation waiting 48 to 72 hours to get approval to, in a sense, to see, oh, can they have opioids or not? So, I think that could be a real risk.

Also, I think there's a real risk of stigma if you're identified as a potential Opioid Use Disorder candidate. We have to be very mindful that stigma; stigma is rampant right now related to Opioid Use Disorder. And that could impact future interactions with clinicians. Would behavior be interpreted as more drug seeking, for example, because there's this bias now of this

patient having a relationship where you have always have to be on the lookout with this patient. I think there could be benefits there too, of course, in the right hands of maybe more mindful watching, for example. But we really have to be mindful of this.

And then, also, institutional responses. I know it was mentioned that insurance companies don't have access to it, but how would hospital administration want to handle these results? What would professional organizations and societies recommend? Also, what would legislatures — we see more and more state legislatures getting involved in opioid prescribing and what should happen there. So, I think we have to be mindful of how these things may play out in the current characteristics we have in our society about opioid prescribing.

Dr. Watson: Thank you. Dr. Bateman.

Dr. Bateman: I was just going to note, we've heard conflicting information about whether or not, or the degree to which the algorithm might be biased by genetic admixture. If, in fact, it is, as you know, some of the experts who have provided input to the Committee suggest, then, a risk of putting this out into practice is it really could exacerbate disparities in the way pain treatments are used. And we've seen examples of racially biased risk calculators around kidney disease and in other settings that have really had had negative adverse outcomes and disadvantaged certain populations. So, I do think there's a need to be attentive to that. And I would encourage the FDA, if they're going to move forward with this, that they get input from statistical geneticists that have real expertise in correcting for genetic admixture and look at the algorithm and the way it performs and make sure that this isn't an issue.

Dr. Watson: Thank you. I don't see any other hands up. That was a fantastic discussion of the risk and the benefits I think of this. Does anyone have any final things to add? Dr. Kelm? I thought that was a great discussion and I hope —

Dr. Kelm: Yes, that was excellent. Thank you.

Dr. Watson: Okay, great. So, let's go on to number five.

FDA: Number five. Taking into consideration the current methods for assessing the risk of developing OUD after exposure to prescription oral opioids for acute pain, please discuss the clinical validity of AvertD.

Dr. Watson: That's a great question. It's kind of bringing it all together. So, who wants to start off? The clinical validity of AvertD. Dr. Bierut?

Dr. Bierut: Yeah, I'll start off with it. I think I have no idea about the clinical validity of AvertD. I continue to have concerns about how this test is really doing and these 15 SNPs and if they really are accurately differentiating people who have developed Opiate Use Disorder and not Opiate Use Disorder, I feel I need more data to give me confidence in that. And then the other issue with it is, there's the assumption after you give this test that it changes the behavior of the physician and changes the behavior of the patient. And I don't think we know how it's doing with that.

Dr. Watson: Thank you very much. Other thoughts? Clinical validity of AvertD. Panel?

Dr. Dunn: Actually, could I ask a question to the Panel members? I mean, what are the current methods for assessing risk of developing OUD after the prescription of opiates? Is there a

standard, like I follow up with you in three days to see what your use is like, I'm tracking your medication use over a month…? Is there kind of a standard practice there?

Dr. Farrar: This is John Farrar. Perhaps I can answer that a bit, and then others should jump in as well. There are guidelines that have been published, but I have to say that the guidelines vary dramatically from institution to institution and organization to organization. One of the slides presented by the company gave an indication of the positive/negative predictive value of a variety of different patient reported outcome measures. What is not clear there is whether using all of those together actually gives you a better understanding. And I would argue that it does and that that process is right for further study and would be a key piece to understanding what the answer to this question by itself. I agree completely that I'm not sure what the clinical validity of this is in combination with other things. And with testing in a more uniform, or in a more representative population, there could be data that would suggest that it works. Let me leave it at that.

Dr. Watson: Thank you. Dr. Compton.

Dr. Compton: Well, you asked how is this kind of work done, and it's not an easy project. Because what we're talking about is, how do we track people from first prescription to an outcome as distal as Opioid Use Disorder? So, a retrospective study, like the one done here that we're trying to evaluate, is one of the approaches, because it's so difficult to do those long-term outcome studies, and there's so few of them. I would point out that an intermediate outcome that has been studied quite a bit and is used by many policymakers and some of the state-based regulators and so forth would be persistent opioid use, as one intermediate outcome on the way from first exposure to you use it frequently, and then use disorder being even further down the

road. But there are some studies, they're typically observational longitudinal studies. And I would point out the Monitoring the Future study out of the University of Michigan is one that has looked at, in this case, adolescent, which, it's kind of struck me today that, although we're focusing on the population over 18, a number of the public comments were of people exposed much earlier in life, and opioids are exposed to adolescents for third molar extraction, often before you hit age 18, or for sports injuries, or for other acute care needs. And though it has been shown in the Monitoring the Future study, for example, that early exposure medically is a risk factor for later problematic use, they don't have great measures of Opioid Use Disorder. So, I can't say it's gone that far.

Dr. Gordon: Yeah. This is Adam Gordon. I'll just add to what Dr. Compton said. I will say in the VA, there's very, very good predictive tools, Storm Dashboard, et cetera, that we use, but not for opiate use disorders, mainly for opioid related adverse events, whether that be an overdose event or problematic opiate use or opiate misuse. And many of the scales and instruments that we use, both clinically as well as preoperatively, et cetera, are really looking at those shorter term outcomes. So, problematic use and/or overdose related events. And so, one thing that is very unique about this is that there’s not, not to my knowledge in the addiction world, is there any predictive tool that we have for opiate use disorder. And I would love to hear any of the panelists who could contradict me, but usually it's other opioid-related morbidity.

Dr. Farrar: So, if I might jump in, just one quick comment to Dr. Compton's comment about following chronic opioid use. Having taken care of chronic opioid patients for quite a long time, there are patients who are just better on opioids than not on opioids, and we need to be careful not to throw the baby out with the bath water. The second issue is that the definition of

Opioid Use Disorder is not infrequently confounded with chronic use. And I can tell you, in my own practice with my own faculty, I go out of town for two weeks, and all of a sudden, my patient now has an OUD diagnosis on their chart, because they called while I was away because of a flare up and required additional opioid. So, I think there is a real issue about how we define it. And if it's not carefully defined, then that raises issues about how this works. Now they tried very hard to define it, but I'm not completely convinced that it was exactly the definition we might have used. Thank you.

Dr. Watson: Dr. Zaafran.

Dr. Zaafran: Yeah, Sharif Zaafran again. I can't emphasize the importance of what was said earlier. There's a lot of confounding factors with just labeling one entity or one test, how much clinical validity it has along with it. In the chronic pain arena, which I know we're talking about acute pain here, but even when you have positive, not doing the right thing, one of the things that we're trying to emphasize with physicians is that's not an excuse to stop treating them. That may mean that you need to continue treating them, but they may need other help for Opioid Use Disorder alongside with that. I really, really worry about this being looked at as one main factor to under-prescribe or to under-treat people who are susceptible. And the unintended consequence of that is that you're going to push people from the regulated community to the unregulated community. They're going to find these medications if they're in pain in places that we don't want them to, where they're going to have overdose deaths. From the standpoint of acute pain, I mean, we haven't even talked about the fact that people have acute pain on top of chronic pain. And how do you manage that process? We already know that these people are probably susceptible as chronic pain patients, and they've been on opioids for a long period of

time, and now, are they more susceptible because they have acute pain on top of that? So, we just have to be very careful. I'd like to go back to what we said earlier about risk stratification, and, just talking about clinical validity from this standpoint, only being at an increased risk or a decreased risk, but not using it as one clinical tool to assess where that risk stratification is for these patients.

Dr. Watson: Thank you. I don't see any other hands. I think we've heard very consistently the answer of nobody knows the validity, and you heard a lot of different reasons why. Dr. Kelm, have you gotten what you needed on this question?

Dr. Kelm: We thank the Panel for everybody's input on this question, and I think we can move to six.

Dr. Watson: Okay. So, six. If you believe that additional information in the labeling, like warnings limitations, would be appropriate to mitigate some risk for this task, please describe the specific risks and the labeling mitigations that should be included to minimize those risks associated with the use of the device. Are there other mitigations to consider to minimize risks associated with the use of the device? Lots of people answered. Jennifer Higgins.

Dr. Higgins: Thank you, Jennifer Higgins. I came to the conclusion that it should be multi-pronged, obviously, in a way to communicate to both prescribers and patients about that this is one tool. It's one tool to assess risk. It's not the only thing, and I guess I land on, like, REMS. There's some challenges with REMS. I understand it needs greater teeth sometimes. And the PDMP, which is what the sponsor referred to using, they're at various stages of development across the 50 states that some have longer histories than others. So, I don't know whether that would be a foolproof method. And then labeling, I think that needs to clearly indicate that this is

only one tool in assessing risk of Opioid Use Disorder. And then patient education, obviously, and I don't really know what it would look like, but those are the four things that I think I would recommend in terms of warnings and limitations.

Dr. Kelm: And this is Kellie Kelm. For devices, we don't have a REMS program or a REMS-like program. Also, I guess it's unclear to me about whether the PDMP applies and includes information about testing or is just about prescription drugs. But, if you want to clarify, Dr. Higgins?

Dr. Higgins: How do I think it relates? Is that what you mean?

Dr. Kelm: Mm-hmm.

Dr. Higgins: I feel like if drug monitoring is going to be there, there has to be greater communication, I think, between prescribers and clinicians and even patients. And I think that, if we're talking about use of this test in ERs and — I guess I would just want there to be tightened up PDMP monitoring, that's all.

Dr. Watson: Thank you, Dr. Compton.

Dr. Compton: Well, thank you. We've raised a number of potential possibilities here in terms of labeling and related issues. I would say the labels of low risk and high risk should be reconsidered, if this moves forward, to some other approach that's more specifically tied to the nuances of what is meant to hopefully minimize the risks of overreliance on this as a single test to give you a yes/no answer of whether you're going to have a bad outcome or not. It's clearly a mistake. This is a risk assessment, not a definitive outcome one way or the other. That's probably my main point here.

Related to that might be the recommendation for consideration of post-marketing studies to understand the impact on prescribing patterns. Because ultimately, that's prescribing patterns in patient behavior. So, ultimately, that's the goal, as well as some of the real gaps in data for subpopulations that you just couldn't do because it's a sample of 380 people. But if it does go forward, to require additional studies in the post-marketing studies for subgroup analysis. Those are at least what occurred to me, having been listening today.

Dr. Watson: Thank you. Dr. Ruha.

Dr. Ruha: One other thing. I agree. I think there should be warnings, but one other thing that I would just recommend is that it just be very clear with the labeling that opioid sparing strategies are recommended for all patients. And the results of this test should not be used to withhold opioids from patients when they otherwise would be considered that the opioids are indicated. I think if we put it right out there for physicians to not make a determination of administering opioids based alone on the results of this test in the label.

Dr. Watson: Thank you, Dr. Joniak-Grant.

Dr. Joniak-Grant: Yes, I agree with Dr. Ruha about putting it right out there. I think it's really important to use plain language to not get lost just in the weeds of the statistics. Clinicians are busy. They're not statisticians. And to really say things directly, like, "Do not use this test alone. Do not use it to withhold treatment." Put things into what the predictive power is, maybe, as part of the results, versus just a strict yes no. For example, I, I looked at the limitations paragraph in the insert and it starts with, "Genetics and lifestyle play a significant role in a person's risk for OUD," and then you go down a whole paragraph, and then it says it shouldn't be the sole determinant, and there's other factors, non-genetic factors at play. So, it's just buried in

there among pages and pages of things. And I think it really needs to be emphasized, bulleted, and maybe even have a brief fact sheet of how it's supposed to be used. Something that can be skimmed really easily.

Dr. Watson: Thank you. Dr. Ness.

Dr. Ness: I won't keep reiterating my, don't like high risk/low risk sorts of things, except to say, as a clinician, I would feel hesitant about ordering a test like this if it felt it was going to have to lock me into a pattern of treatment related to the patient. Because "Gosh, they're low risk. How can I not treat them that way?" My personal feeling is that anything that's like a polygenic assessment can never really be answered with a categorical yes/no answer. My opinion, I've already said it before, is it's great. I want this information, but I don't want to be told that I have to do things in a certain way. And our medical legal world is, such as people have said, is if it says high risk, a lot of people, they'll never give them opioids.

Dr. Watson: Thank you. Dr. Walker.

Dr. Walker: Yeah, I would like to draw an analogy here. Some of you may have been familiar with the Oncotype DX test kit that's done for breast cancer patients, where they do the genetic profiling and then they tell you what is your likelihood to benefit from chemotherapy. And if you're one end of the scale, they'll tell you you're very likely to benefit. If at the other end of the scale, you're very unlikely to benefit, and if you're somewhere in the middle, you're not going to have a lot of guidance on that. But I think the reason that's a valuable thing for us to consider in this particular discussion is, what is the likelihood that an individual is going to have all 15 things in a certain order or not in a certain order? And so, I'm guessing that, in fact, there will be a spectrum, and there will be, as Dr. Ness was saying, not a yes or no answer, but

somewhere in the middle. And so, I just wonder if it's worth considering what that algorithm would look like if it gave you the whole spectrum and if the risk really does vary depending on where you are on one end of it or the other.

Dr. Walker: Thank you. Dr. Zaafran.

Dr. Zaafran: I took a stab at writing a sentence. "This test is one of many possible risk factors for Opioid Use Disorder that must be used with clinical correlation, and it's only for aiding in managing patients in acute pain." I guess that's the only way I would put it out there, just from the standpoint of, it's one of many risk factors. And I think it just needs to be emphasized at the very top and not buried down at the very bottom.

Dr. Walker: Thank you, Dr. Farrar.

Dr. Farrar: Dr. Ness was very nice. Not to say that he was going to push his point more, but I am a strong believer in the devil's in the details. And I very much, if this is going to go forward, which there are some concerns about without additional studies, from my point, it needs to be a continuous measure between 0 and 10, or something else that Dr. Walker was just describing, or others have, so that it's not a yes no. Whereas Dr. Ness rightly said, we're basically backed into having to do something based on it. The idea would be that, if you're in the higher end, you need to do other things to ascertain what the risk is. And then there's a whole bunch of things. And even if you're at the lower end, it should say you still should use minimum opioid necessary, but treat the pain. That small difference, which is a continuous versus a yes/no, it's a huge difference. And I think that will be at the top of any label, which is, it's not yes/no, it's a scale, not unlike others, and that it requires patients and physicians to think about what they're

doing because of the way it's presented. If it's presented yes/no, I have great fears that it's going to be misused and back a lot of us who take care of patients into corners we can't get out of.

Dr. Kelm: This is Kellie Kelm with FDA. I appreciate, especially those of you who probably have a lot of experience with genetic tests and, and assessing information from multiple, for example genes and, and in a, in a process like that. Obviously, the device we have in front of us is one that provides either — right now it is categorical in that it provides you a number that they want to say is a higher risk of developing OUD. And one that we can talk about is either lower risk or at current normal levels, if you will. I just do request that, as you're providing your assessments for us, it would be helpful if you can keep to the device that we have in front of us, because unfortunately, we have to make a decision on this device and not one, for example, that might provide something outside of this categorical result. So I thank you.

Dr. Farrar: I will have to say that if it's going to be categorical, I am strongly opposed to it. I honestly would have them go back and redo it as to provide a continuous scale. And I know we're not here to set policy, but that would be a strong concern for me.

Dr. Walker: Okay. Dr. Bierut?

Dr. Bierut: I want to bring up this issue about testing people under the age of 18. When we heard all the family discussions, there was a lot of, “What about my child who was 16 going to the dentist,” all of that. I'm not sure how to mitigate that risk, but I think that there will be desire for the testing in adolescence, especially as they're going and getting their wisdom teeth out. And I don't know how we mitigate this.

Dr. Walker: Thank you, Dr. Gallagher. My concern was the same. The public comments from parents were very passionate, of course, and recognizing their concern and their experience

and having empathy for them, I would want there to be some kind of limitation or something listed so that it could be managed, because I think you will find parents going to physicians and saying, I want this test done on my child. But again, a lot of places won't do genetic testing on someone younger, but it just, to kind of work through those issues. But something that would mark that difference, because this was designed for the test market of people 18 or older. So, some way to protect that.

Dr. Walker: Thank you. Okay. I think I've seen all the hands. So, I think what I've heard is there are a number of recommendations for how you might mitigate some of the risks. A lot of them dealt with changing from categorical to continuous, which as you very well point out Dr. Kelm, that's not this device. This device is a categorical device. I hope you got a lot of good comments. They really were some good comments in that.

Dr. Kelm: I appreciate the comments from the Panel. And no, thank you very much.

Dr. Walker: Great. So, if we've received all the feedback you need from us, Dr. Kelm, then we can go on to summations. Are you okay with that?

Dr. Kelm: Yes. I think we're good to move on. Thank you.

Dr. Walker: Great. Thank you so much. At this time, the Panel will hear summations, comments, or clarifications from the FDA. And you have 10 minutes, FDA.

Dr. Kelm: Hi, this is Kellie Kelm. Mine is going to be short and sweet. I just really want to thank the Panel today. The discussion has been excellent and illuminating, and I think it's going to be very helpful for us. I want to thank everybody, as well as the sponsor, the public

comments, and, of course, thank my review team here, as well, for great meeting today. Thank you very much for your time. Appreciate it.

Dr. Watson: Thanks so much, Dr. Kelm. Now we will hear summations, comments, or clarifications from the sponsor. Sponsor, you have 10 minutes.

Dr. Donaldson: Thank you so much, Dr. Watson. As previously commented from Dr. Kelm, we would like to thank all Committee members for the significant amount of time and effort spent on reviewing a large amount of data and a complex disease state, as well as the significant comments, and particularly under question six, as we will find those extremely helpful as we move through this process. As you heard today, the Panel has had confidence in AvertD test detecting genetic risk for OUD for the first time in a clinical study. We view that as a milestone. AvertD is a genetic risk stratification test, not a screening test, and it's intended to be used in conjunction with a clinical evaluation in acute non-emergent pain. Some of the conversation today was talking about applications of this test outside of the intended use. I just want to make sure the Committee understands that the intended use of this test is for first opioid prescriptions for acute non-emergent pain, not in a chronic pain population.

The other comment that I think it's important that came out of this discussion is, whether it's our test, or any other genetic risk assessment test for OUD, additional work needs to be done in a post-market setting in regard to how the test will be implemented in the clinical practice. This is done in a diagnostic space in a post-market setting with additional studies to evaluate how this test changes behavior. We are willing to do the work here, and, given the need for safe prescribing practices, we think we should start today. As discussed by the Committee under question six. We are confident that the uncertainties discussed today can be addressed through

our controlled rollout process, collection of post-market data, and we are very open to modifying the labeling and description of the test results to ensure appropriate use and avoid stigma. For AvertD, the benefits of providing additional, currently unavailable genetic risk information will enhance the standard of care, and the benefits outweigh the potential risk. Thank you for your time.

Dr. Watson: Thank you very much. Well, we've reached the time now, almost, to get to the vote, but before we do that, I would like to ask our non-voting members to make their final comments. I'll start with Jennifer Higgins, our consumer representative. Ms. Higgins, do you have any final comments?

Dr. Higgins: Just thank you very much. It's been a very enlightening and wonderful conversation and discussion. Thank you to the sponsor and thank you to the FDA for this opportunity.

Dr. Watson: Thank you. Now I'd like to ask Mr. Elijah Wreh, our industry representative. Do you have any comments?

Dr. Wreh: Yes. Sure. Thank you again for this meeting. Very interesting meeting. For the record, my name is Elijah Wreh for industry, and I work for Boston Scientific, but I'm Chair. I just want to thank you, the Panel members, and the FDA, and the sponsor for such great conversation on this product. There are two things I want to share with this group, two key areas that I would like to talk about before we go into our voting section. Really quick, the first point is what I call the CRF, the changing Case Report Form that FDA talks about, and the sponsor also shared on that as well. If you notice in the FDA materials, the FDA raised a concern that the changes in the CRF may have led to inconsistencies defined by study population in the clinical

study. SOLVD Health disagrees with the FDA on that view. And the sponsor has shared with the FDA that, while the form did not change over time, the inclusion in the clinical criteria in the protocol were used to train the sites and enroll patients, as well. So, the CRF from, I believe, from an interest standpoint and standard, and many medical device manufacturers use this form. It's a document used only to capture enrollment data. SOLVD Health further reduced annual rates regarding this out by having all sites complete the CRF form for all patient upon completing the study. And this also yielded at a hundred percent concordance as well. So, I think the result was not surprising at all to the sponsor, as the CRF changes over time and, and the study is very typical for the United States and does not impact the study population.

The second point I would like to share is what I call the enrichment that the FDA talked about as well. If you notice any sponsor materials, an enrichment strategy was used to ensure that an adequate number of OUD positive patients will also include, as well. This type of enrichment is supported by FDA guidance documents for drug trial, but it's also commonly used in test clinical studies. So, I think sharing this perspective is very important for going to the voting process. So, I hope the Panel members will consider those two points as they go to the voting process. Thank you again.

Dr. Walker: Thank you. And finally, I'd like to ask Dr. Joniak-Grant, our patient representative, if she has any final comments.

Dr. Joniak-Grant: Just a brief one. Thank you, FDA, for including patient and consumer voices today, and thank you everyone for the great discussion and for answering my questions. I know I had a few. I just want to say that I really feel for everyone impacted by Opioid Use Disorder and opioid overdose, as well as pain patients. I wish there was a simple test to predict

exactly what was going to happen with a lot of reliability and accuracy. Unfortunately, I don't think this is the one. Maybe it could be, one day. So, for me personally, there's some issues with the clinical design, which we've all discussed, and we've talked about.

So, given that, and given the risks of what will happen in clinical interactions for possibly delayed treatment, under-treatment, the stigma of being identified as high risk, and quite a few times erroneously, kind of outweigh the potential benefit of info that might have, at best, a predictive power of 17%, and likely could be closer to 5%. And I think it's just really important to remember for patients, and we heard a lot about empowering patients, but incomplete info and decontextualized information, especially if it starts to get into the realm of misinformation or accurate info, is not patient empowerment. It can really take away any power that patients have. And so, I would just leave you with that thought. Thank you very much.

Dr. Watson: Thank you. Well, Panel, you have been a fantastic Panel, and we've actually come to the point of voting. So, we're now ready to vote on the Panel's recommendation to the FDA for the de novo request for the AvertD test, sponsored by SOLVD Health. The Panel will vote on one question relating to the risk benefit profile of the device. James Swink will now read two definitions to assist in the voting process. Mr. Swink.

James Swink: The Medical Device Amendments to the Federal Food, Drug and Cosmetic Act as amended by the Safe Medical Devices Act of 1990 allow the Food and Drug Administration to obtain a recommendation from an expert advisory panel on designated medical devices filed with the Agency. The de novo classification request must stand on its own merit, and your recommendation must be supported by safety and effectiveness data in the de novo request, or by applicable publicly available information. To grant a de novo request, the FDA

must determine whether general controls, or a combination of general and special controls, can provide a reasonable assurance of safety and effectiveness. The definitions of safety and effectiveness are as follows.

Safety as defined in 21 CFR section 860.7 – There is reasonable assurance that a device is safe when it can be determined, based upon valid scientific evidence, that the probable benefits to health from use of the device for its intended uses and conditions of use, when accompanied by adequate directions and warnings against unsafe use, outweigh any probable risks.

Effectiveness as defined in 21 CFR section 860.7 – There is reasonable assurance that a device is effective when it can be determined, based upon valid scientific evidence, that in a significant portion of the target population, the use of the device for its intended uses and conditions of use, when accompanied by adequate directions for use and warnings against unsafe use, will provide clinically significant results.

Panel members, we will now begin the voting process. I will read the voting question. Each of the voting members have received an electronic ballot to respond to you. Once I read the question, please vote, and I will tell you the votes and read them into the record.

Do the probable benefits to health from use of the AvertD device outweigh the probable risk for the proposed indications, taking into account the probable risks and benefits of currently available alternative forms of detecting risk of developing OUD? Please, now. Yes, no, or abstain?

James Swink: The votes have been captured. I will now read the official vote into the record. The panel voted 2 “yes” and 11 “no” that the probable benefits to health from use of the AvertD device outweigh the probable risks for the proposed indications. Dr. Watson?

Dr. Watson: Yes. Okay. So, Panel members, we'll go around the table now and I'll ask everyone to discuss their votes. Can we take down the graphics so we can see the entire Panel? I want the Panel members to discuss their votes. If you answered no, please state whether changes to the labeling restrictions on use or other controls would make a difference in your answer. Please state your name and how you voted for each question for the record. Dr. Compton.

Dr. Compton: Yes. Thank you. I'm Dr. Wilson Compton, and I voted in favor of approval, of clearance. As I read, I would've preferred not to be able to answer yes or no, but to give sort of a balance, and I just barely tipped into the 'yes' area. I heard lots of concerns expressed today about potential risks. I expressed a number of those myself in terms of risks. But the part of the question that tipped the balance for me was the second half, where it asked for risk and benefits in comparison to current available technology. And to my mind, I think the genetic tests are likely to add benefit compared to the currently used risk profile tests that clinicians have available at this time. Anyway, that's the gist of why I voted in favor.

Dr. Watson: Thank you, Dr. Compton. Dr. Ness, please discuss your vote.

Dr. Ness: I voted no, and I guess not surprisingly. It's because of the categorical nature of this device in the saying of high risk/low risk. In the sense that I believe a hundred percent of the risk that's associated with having a test is for the false positives and false negatives in these things. Both people being untreated or poorly treated because somehow it came back as a positive result or being given an inappropriate treatment because it said negative on the thing. I can say unequivocally, that if it was just simply reported as continuous data with associated information, I would've said yes.

Dr. Watson: Thank you, Dr. Ness. Dr. Gallagher, please discuss your vote.

Dr. Gallagher: Thank you. This is Colleen Gallagher, and I voted no. My reason is primarily because of the high risk/low risk categorization because that's not understandable to most people. And I think that if you're trying to make this a categorical kind of response, that that confusion to the people who are the patients would be detrimental. I think if it could be something along the lines of, based on your genetic makeup with those 15 alleles, you have this much percentage, versus that much percentage, some other description, that might be more possible. That's my reasoning.

Dr. Watson: Thank you. Dr. Walker, will you please discuss your vote?

Dr. Walker: Yes, I was one of the other two to have voted yes. I came in with reservations about the test validity, particularly based on the one published study, but I felt like there were plenty of asked and answered sessions there, and so I felt better about it. I do completely agree that there could be semantic improvements in terms of how the information is conveyed, but I also felt there was a willingness on the part of the sponsor to make those modifications. And so, in the end, I came down feeling like that there was sufficient benefit to vote yes. Thank you.

Dr. Watson: Thank you. Dr. Farrar, please discuss your vote.

Dr. Farrar: So, obviously, I voted no. Dr. Ness has expressed one of the primary features here, which is that I feel very strongly that the devil's in the details here, and that giving a yes/no is going to end up stigmatizing people and or giving clearance to people who ought to be carefully monitored and taken care of. All of that ought to be standard procedure as well. The other concern I have is that, while this mix of SNPs seems to differentiate to some degree, we do not currently know how much of the genetic component is actually answered by this snip. And

there was very little in the report and or in the presentations and or in the process to talk about the environmental socioeconomic and the environmental structure that need to be accounted for taking into consideration and clearly defined in terms of how they would be appropriately added to this to improve the usefulness.

Dr. Watson: Thank you, Dr. Farrar. Dr. Bateman, please discuss your vote.

Dr. Bateman: I voted no because of continuing uncertainty about the validity of the test for the reasons that I indicated in the discussion. I think there's a need to take this algorithm and this approach and assess its performance in other populations, and there were some suggestions about how that might be done, and a need to address continuing concerns about potential confounding by ancestry, as well as to better understand why these 15 SNPs in the way they're analyzed have better predictive value than what we're seeing in the Genome-Wide Association Studies. I also think that to fully understand the risks and benefits of using such a test there needs to be a study of how it affects prescribing behavior.

Dr. Watson: Thank you. Dr. Dunn, please describe your vote.

Dr. Dunn: Hi, Walter Dunn. I voted no. As I stated before, I had concerns about the design of the trial. I don't believe it answers the question of, can this test or will this test lead to more judicious use of opiates in acute pain control. The one main issue was that, I believe, the retrospective portion of the study, specifically the recount of when the index opiate exposure occurred, enriches for a certain type of patient in a situation that is not representative of the general clinical use, intended use.

And second, this has also been discussed, being able to educate and communicate to physicians to appropriately apply the information commensurate to the level of OUD risk that the

data confers. So, we're not even sure how much risk this information it applies to opiate risk. So now how are we supposed to convey that to the practicing clinician? And, as we've already heard from some of the public commentary, there's already indication that they're going to be over-reliant on this test as a simple, actionable kind of yes/no type of type of tool. To what my colleagues had mentioned, I think the changing the outcomes to some type of dimensional outcome, I think would significantly bolster against something that over-reliance. But again, I think more science has to be done on that.

And then finally, you know, about the ultimate kind of use of this tool. It's about how this would change management of patients, right? The study wasn't designed to do that, right? That wasn't the goal of the study. But I think that's critically important, because that's ultimately what we're interested in. Additionally, this study did not address potential risk. We talk about hypothetically, but this study did not collect that type of information in a systematic manner. Specifically, what's the risk for inadequate pain control if something like this was in place. Thank you.

Dr. Watson: Thank you, Dr. Bierut, please discuss your vote.

Dr. Bierut: I also voted no. My no vote was really based on my concerns about the clinical validity of this test. I am concerned about the validity of these 15 SNPs really being able to differentiate people with Opiate Use Disorder and not Opiate Use Disorder. And I would like to see replication of this finding, which could be done in silico with existing data.

I'm also concerned about different ancestors and how this test will, will really work in different a

Dr. Watson: Thank you. Dr. Ruha, please discuss your vote.

Dr. Ruha: Hi, I'm Michelle Ruha. I voted no, simply because I don't feel confident that it is very safe, because we really don't know how it will be used. We know how the sponsor intends for it to be used, but I think we all have some skepticism that it will actually be used that way. As far as efficacy, there's the limitations with the sensitivity and specificity and then risks. We've all vocalized many potential risks. So, I just felt that the risks outweighed the benefits at this time.

Dr. Watson: Thank you. Dr. Gordon, please discuss your vote.

Dr. Gordon: I vote no for various reasons. I won't go through detail. Number one, the design of the studies as well as the anal analysis, I thought was not adequate to sway me, that it was an effective test. Second, there was no safety evaluation of this, of this, this device, which I thought was a really odd omission. We have no idea what the prescribing patterns would be after the test was implemented. We have no patient level outcomes. We have no assessment of both provider and patients, of how they would approach receiving results of this test. And I think that's a really important point, because, number three, I'm really worried about the false positives. You will have a lot more angst out there in the provider community, as well as a patient community, as you say that you are at high risk of Opiate Use Disorder, and we have no idea how that's going to play out. And then the last point, imagining that this was a perfect test, a perfect device that could predict adequately OUD, I really worry about the OUD population. Currently, right now, they're stigmatized. They're not getting elective surgery. They're not getting adequate pain control after surgery. And I think that could be exacerbated without some safety checks that this device was implemented. Thank you.

Dr. Watson: Thank you, Dr. Zaafran, please discuss your vote.

Dr. Zaafran: I voted no for a lot of the same reasons that were stated out there. I have a lot of concerns about the demographic population that was sampled in the testing and that it's not reflective of the population that would actually be treated out there. I also have a lot of concern about how this is going to be utilized. And it is going to be taken as a categorical, which is what they're intending, when it really should be more of a risk stratification. I think a lot of folks have already talked about the fact that there's going to be overtreatment of certain patients and undertreatment of many patients out there. As I had mentioned earlier with Opioid Use Disorder, that should not be an excuse not to treat, which I think is what's going to happen if this comes out there. It should be an excuse to manage these patients more carefully, but I do worry that it is going to be used in a way to not treat these patients. And I also worry about the fact that physicians who treat patients who have Opioid Use Disorder based on this test may be target targeted by regulatory agencies in an unintended fashion. We've seen that already happen with the CDC guidelines. I can see it happening with something like this, because in their mind, it's looked at as an objective number, when it should not be reflected that way. So anyway, those are my concerns, and that's why I voted no.

Dr. Watson: Thank you. Dr. Wang, please discuss your vote.

Dr. Wang: Thank you. I voted no, and my vote was driven by the concern that the clinical study design, particularly, is not reflective of the target patient population. I also have concerns regarding the safety and how this test will be used in a clinical setting in the real world. I'm a strong believer in that. I think because of the limited proportion that the genetics play in OUD, I think such a test cannot just be provided to the provider in standalone format. I would

like to see clinical information, information on how such genetic test results can be combined with the clinical information to be provided to the provider at the same time. Thank you.

Dr. Watson: Thank you. And Dr. Goldstein, please discuss your vote.

Dr. Goldstein: Yes, thank you. I voted no. Many of the same issues, for me, have been enunciated. My own particular spin on them is: A, I'm very uncomfortable with the binary nature of the readout. B, the clinical trial population, which should be the experimental test of the machine learning algorithm, was not adequately powered for many of the most appropriate racial subgroups in this country. And C, one thing one of the other panelists just reminded me of, we have Genetic Non-Discrimination Act laws, GNDA, for those of you who remember it, Ted Kennedy got it passed before he passed away. And this may cross the line. So, I think we have to be pretty careful in this area.

Dr. Watson: Thank you very much. I want to say a very hardy thanks to the Panel. You guys have been fantastic in the discussion has been wonderful. Thank you to the sponsor and all of the Open Public Hearing speakers for the contributions to today's Panel meeting. And thank you to the FDA. Dr. Kelm, do you have any final remarks?

Dr. Kelm: Nothing additional to add except to thank everybody. Thank you for your time, again, and I hope everybody has a lovely evening. Thanks.

Dr. Watson: And with that, I'm gifting you all back 15 minutes of your life. This meeting at the Clinical Chemistry and Clinical Toxicology Devices panel is now adjourned. Thank you everyone.