DEPARTMENT OF HEALTH AND HUMAN SERVICES
FOOD AND DRUG ADMINISTRATION (FDA)
CENTER FOR DRUG EVALUATION AND RESEARCH (CDER)
OFFICE OF NEW DRUGS (OND)

+ + + + +

EVALUATING THE NEGATIVE SYMPTOMS OF SCHIZOPHRENIA IN CLINICAL TRIALS

+ + + + +

PUBLIC MEETING

+ + + + +

FRIDAY
AUGUST 16, 2024

+ + + + +

The public meeting was convened at FDA White Oak Campus, Building 31 Great Room, 10903 New Hampshire Avenue, Silver Spring, Maryland, at 9:00 a.m., Dr. Bernard Fischer, Deputy Director, Division of Psychiatry, Office of Neuroscience, CDER, presiding.

PRESENT
BERNARD FISCHER, MD, Deputy Director, Division of Psychiatry (DP), Office of Neuroscience (ON), OND, CDER, FDA
ANTHONY O. "TONY" AHMED, PhD, HSP, Associate Professor of Psychology in Clinical Psychiatry, Vice Chair for Psychology, New York-Presbyterian/Westchester Weill Cornell Medicine
RACHAEL BLACKMAN, MD, PhD, Clinical Reviewer, DP, ON, OND, CDER, FDA
JACK J. BLANCHARD, PhD, Associate Provost for Enterprise Resource Planning and Professor, Department of Psychology, University of Maryland

STEPHEN BRANNAN, MD, Chief Medical Officer, Karuna Therapeutics
TERESA BURACCHIO, MD, Director, ON, OND, CDER
MICHELLE CAMPBELL, PhD, Associate Director of Stakeholder Engagement, ON, OND, CDER, FDA
CHRISTOPH CORRELL, MD, Professor of Psychiatry and Molecular Medicine, Feinstein Institute for Medical Research
TIFFANY R. FARCHIONE, MD, Director, DP, ON, OND, CDER, FDA
WILLIAM P. "BILL" HORAN, PhD, Professor, UC Los Angeles, Chief: Psychosis Section, VA Greater Los Angeles Healthcare System
ERIC JARVIS, MD, Associate Professor of Psychiatry, McGill University
BONNIE KAISER, PhD, MPH, Associate Professor, Department of Anthropology; Global Health Program, UC San Diego*
RICHARD S.E. KEEFE, PhD, Professor Emeritus in Psychiatry and Behavioral Sciences, Duke University School of Medicine
DEANNA KELLY, PharmD, BCPP, Acting Director, Maryland Psychiatric Research Center, University of Maryland School of Medicine
BRIAN KIRKPATRICK, MD, Professor, Psychiatric Research Institute, University of Arkansas for Medical Science
STEPHEN R. MARDER, MD, Professor, David Geffen School of Medicine, UCLA
MARK G. OPLER, PhD, MPH, Chief Research Officer, WCG, Inc.
MATTHEW RACHER, CRPS, Certified Peer Specialist, CureSZ Foundation
ROBERTA RASETTI, MD, PhD, Clinical Reviewer, DP, ON, OND, CDER, FDA
DAVID REASNER, PhD, Director, Division of Clinical Outcome Assessment, CDER, FDA
MICHAEL SAND, PhD, CEO, S2 Consulting, LLC
NINA SCHOOLER, PhD, Professor of Psychiatry and Behavioral Sciences, SUNY Downstate Health Sciences University
BRANDON STAGLIN, President, One Mind
GREGORY STRAUSS, PhD, Franklin Professor of Psychology, University of Georgia
LAURA SWETT, PhD, Reviewer, Division of Clinical Outcome Assessment, CDER, FDA

SOPHIA VINOGRADOV, MD, Department Head and Professor, Department of Psychiatry & Behavioral Science, University of Minnesota Medical School
HEIDI WEHRING, PharmD, BCPP, Clinical Reviewer, DP, ON, OND, CDER, FDA
PEILING YANG, PhD, Office of Biostatistics, CDER, FDA

*via videoconference

CONTENTS

P-R-O-C-E-E-D-I-N-G-S

(9:01 a.m.)

DR. CAMPBELL: Good morning, everyone, and thank you for coming. My name is Michelle Campbell and I am the Associate Director of Stakeholder Engagement and Clinical Outcomes in the Office of Neuroscience. On behalf of my FDA colleagues, I would like to thank you and welcome you to our public meeting on evaluating the negative symptoms of schizophrenia in clinical trials.

We are really looking forward to today's great discussion and learning from you. We'd like to quickly go over some housekeeping items and then we'll get started. You should all have a copy of the agenda. If not, it is located on the FDA's website, under CDER, under meeting and events.

We'll be spending the first half of our morning hearing about the lived experience as well as talking about the circuitry in relationship with the brain. We'll then be

breaking for lunch around 11:40 and we will have an hour for that. The second half of the day we're going to be focusing on clinical meaningfulness, outcomes, and trial design.

Just as a reminder, this meeting is live and is being recorded. To our people who are online, thank you for joining virtually. We will be having someone who will be monitoring the question and answer portion online and we'll do our very best to make sure to incorporate your questions into the live discussion.

Also, for our attendees online, we will be taking back your questions if they're not asked, so we can learn from them and see how we can incorporate your questions and thoughts into our everyday work. Additionally, to our audience that is live and in the great room with me today, if we do not get to your questions and you really have some comments and thoughts, you may send an email to O&D public support at FDA.hhs.gov, and in the subject line

please put, negative symptoms of schizophrenia, in there and that mailbox will filter those emails to us in the Office of Neuroscience.

Real quickly for some more practical logistics, the restroom is out of the hall, down the hall and to your right. There is a kiosk available for lunch and coffee.

So, at this time I would like to invite my colleague and Director of the Office of Neuroscience, Dr. Teresa Buracchio.

DR. BURACCHIO: Welcome, everyone. So good to have those of you here in person, and I understand there's quite a contingent of people online as well, so welcome to you all for our workshop titled, Evaluating the Negative Symptoms of Schizophrenia in Clinical Trials. I'm sure many of you are experts on this already, but I'll just briefly provide a little overview of the day, that we're going to start off with a discussion of negative symptoms, which are flattened affect, poverty of speech, lack of motivation, anhedonia, and

social withdrawal.

And we are well aware at the FDA that there is a great unmet need for therapies to treat these symptoms, we understand that these symptoms cause a substantial impact on patients with schizophrenia, their ability to function in daily life, and in their quality of life, and current antipsychotic therapies are maybe effective for positive symptoms and can treat some factors that contributed to negative symptoms, but negative symptoms can persist and are disabling even in patients who are adequately treated with antipsychotics for their positive symptoms.

We also recognize that there is a great challenge to developing new therapies and conducting clinical trials in these populations. Some of the challenges, we're going to touch on many of them today, but to highlight a few are the use of concomitant therapies with antipsychotics, either at the time of a clinical trial or past history of the

use of these drugs which can impact negative symptoms and potentially blunt treatment effects of therapies, appropriately defining a population to be enrolled in clinical trials that could be anticipated to be responsive to treatments, and developing clinical outcome assessments that are sensitive to change and are capable of measuring clinically meaningful effects on negative symptoms.

The goal of our meeting today is to have an open dialogue about these challenges and identify areas where we can advance drug development and regulatory science in this space. We really have a fantastic agenda for today, I will say that Dr. Bernie Fischer who will be giving our opening comments really was a lead on planning this and he just did a fantastic job, so thank you Bernie. So initially we'll have some introductory overview from Dr. Fischer, who will go through the background on negative symptoms, and this will provide some clinical context and stage-setting

for the sessions that will follow. We will also hear from Mr. Brandon Staglin with One Mind, who will present the perspective of lived experience with schizophrenia and the impact of negative symptoms.

Session one will then present an overview of the current science on neurotransmitter systems and brain circuits related to negative symptoms and overlap with cognition, session two will focus on challenges in designing studies to assess the effectiveness of negative symptoms, session three will focus on the cultural considerations of assessing negative symptoms and how to establish a clinically meaningful change, and then session four will focus specifically on clinical outcome measures for negative symptoms of schizophrenia.

We have many outstanding speakers and panelists today, I really think that we have a fantastic day set for you, and so now I would like to turn to Dr. Bernie Fischer, who

is our Deputy Director of the Division of Psychiatry in CDER, and he will begin with our initial session on providing an overview of negative symptoms. Thank you.

DR. FISCHER: Okay, and then to advance the slides. So, I am going to just do a quick introduction to negative symptoms just to make sure that we're all on the same page. I know many of you are experts in schizophrenia or experts in negative symptoms, but maybe not everyone. So, I do have a number of references at the end of the slides, and they'll probably be best viewed when the slides are posted online after the meeting.

Let's see. All right. So just as an outline of what I'm going to talk about today, I'm going to talk about some early descriptions of negative symptoms, the origin of the terminology, what are the negative symptoms of schizophrenia, why they're important to public health, and why FDA is interested in this, and then I'm going to talk

a little bit about how you describe populations of people with negative symptoms for possible enrollment in a clinical trial. And for that I'm going to focus on the three Ps, predominant negative symptoms, primary negative symptoms, and persistent negative symptoms.

So, in the early 1900s, late 1800s, we had some astute clinicians that were trying to make some sense of mental illness, and they were describing the symptomatology and prognosis to define certain disorders. And the negative symptoms were a fundamental part of early descriptions of schizophrenia, which was at the time called dementia praecox. So, if you look at Kraepelin and what he had said, he said that schizophrenia included a weakening of emotional activities that formed the wellspring of volition, which is a very poetic way of saying that people had problems expressing affect and problems with motivation. Bleuler also noticed the same thing, saying when affect is lacking, there's a lack of drive. So,

people were noticing that diminished emotional expression and diminished motivation.

In 1974, Strauss, Carpenter, and Bartko published a series of landmark articles that were informed by the WHO's international pilot study of schizophrenia, and they looked at the phenomenology of schizophrenia and drilled down on symptoms. They borrowed some terminology from some English neurologists from the 1800s, John Russell Reynolds and Hughlings Jackson, who had used the terms positive symptoms and negative symptoms to talk about brain pathology. When Strauss, Carpenter, and Bartko looked at symptoms, they noticed that some had the appearance of being an active process, like hallucinations and delusions, and they referred to those as positive symptoms. Then there were other symptoms that seemed to involve an absence of normal function, and they called those negative symptoms.

So, what are the negative symptoms? There's been some, a little bit of change over

the course of decades, but for the most part there's five symptom domains that encompass negative symptoms. The first one is blunted affect, which can take the form of unchanging facial expression, decreased spontaneous movement, lack of expressive gestures, the affective non-response can take the form of -- in a clinical interview you tell somebody a joke and they don't crack a smile, they don't laugh, they're just kind of flat. There can be poor eye contact or lack of vocal inflection. Then there's alogia which is poverty of speech, people don't talk spontaneously very much.

There's the avolition and apathy domain, where people may have poor grooming and hygiene, they may have difficulty keeping up with those tasks, they may have physical inactivity, they may stay at home on the couch watching TV all day, they may stay at home all day on the couch with the TV on and not even watch the TV. There's also difficulty seeking employment, keeping employment, succeeding in

school. Then there's anhedonia, where people can demonstrate few interests or hobbies, there's maybe decreased sexual interest, and it's not because of not necessarily finding an appropriate partner. Then there's the asociality domain, where people seem to have few close relationships even among family members, they seem to have few friends, and in social situations they may appear isolated.

So, when you think about negative symptoms, are they best viewed as a single construct, as these are the negative symptoms, they all kind of move together? Well actually, there's a variable presentation that people have with negative symptoms. Some people have a lot of negative symptoms, some people have few negative symptoms, and even within people that have a lot of negative symptoms, they may have different patterns, some people may have more difficulties with motivation, other people may have more difficulties with emotional expression. So, some work has identified two

distinct factors that the negative symptoms can map onto, and they seem to be emotional expression, which would include the domains of blunted affect and alogia, and motivation and pleasure, which would include anhedonia, asociality, and avolition.

More recent work has found that the best fit might actually be thinking about those negative symptom domains as each one representing a separate factor, so that would be the blunted affect, the alogia, anhedonia, asociality, and avolition. And maybe the best way to think about these concepts is that these domains represent a primary order of the factors, where they may map onto those in a hierarchal way, those other two factors of emotional expression and motivation and pleasure.

So, why is it important to think about this? One reason is because the different negative symptoms domains may have different underlying neurobiology, they may

represent different treatment targets. And we know from some studies that they seem to have different impacts on prognosis and course of disease. For example, the avolition factor seems very much related to poor functioning in school and work.

So, why are negative symptoms important for public health? Why is FDA interested in negative symptoms? Well, one reason is because when you look at people with schizophrenia and you see the poor functional outcome that many people experience, it's more closely related to the negative symptoms than it is the positive symptoms. You see people have trouble persisting and work and at school and trouble managing a household including the household finances and keeping the house in good repair, and that may be tied—in some studies it has been tied—to the amotivation symptom domain of the motivation and pleasure factor.

When you look at recovery from

schizophrenia, which has been defined in the literature in various ways, but when I talk about recovery I mean something along the lines of no major symptoms or hospitalizations for the past 12 months, people have some school or part-time work that they're doing, and there is some social engagement with people. When you look a recovery from that point of view, people with high levels of negative symptoms have low rates of recovery. And then when you ask people with schizophrenia what matters to them, people with negative symptoms have poor quality of life, so people are expressing that negative symptoms impact their quality of life. So that's why it's an important treatment target.

So, now I'm going to shift gears a little bit and talk about how you might think of populations that have negative symptoms for inclusion in a clinical study, and there's been, over the course of years there have been lots of different ways to think about people with negative symptoms, but I'm going to

summarize that by talking about the three Ps: Predominant negative symptoms, primary negative symptoms, and persistent negative symptoms.

So, to take them one at a time, first to talk about predominant negative symptoms. When you think about the diagnosis of schizophrenia, positive symptoms are required for the diagnosis. In order to get diagnosed with schizophrenia, there are positive symptoms, but the positive symptoms tend to wax and wane over the course of the illness. There are periods where the symptoms are exacerbated, and then periods where the symptoms are maybe a little quieter. Negative symptoms on the other hand, they tend to be, they're independent of the positive symptoms, they tend to occur earlier in the course of illness, people that have negative symptoms—about 70% of them have the negative symptoms before they demonstrate positive symptoms—and they tend to be kind of stable over the course of the illness.

The concept of predominant negative symptoms means negative symptoms that are greater in severity than the positive symptoms, and that definition requires two things. You have to consider the baseline negative symptom severity, then you also have to consider the severity of the positive symptoms at the moment. So, to illustrate this, if you have a graph with symptoms severity on the Y axis and course of illness on the X axis, when you plot out the negative symptoms, you see that they tend to start in the prodrome or clinical high-risk state, and then they tend to go on during the course of illness and be somewhat stable. When you overlay the positive symptoms, you can see that there's the first episode of positive symptoms and the diagnosis of schizophrenia, but then you see over the course of the illness, they tend to wax and wane, there's some variability there. The period of predominant negative symptoms would be that period where the positive symptoms are less the

focus of treatment than the negative symptoms, or the negative symptoms are more severe.

Now I'm going to talk a little bit about primary negative symptoms. Primary negative symptoms are negative symptoms that are due to the neurobiology of schizophrenia; the schizophrenia is directly causing the negative symptoms. Secondary negative symptoms are caused by something else, but they have the appearance of negative symptoms. Some examples of causes of secondary negative symptoms are, one, positive symptoms. Positive symptoms can cause secondary negative symptoms, and an example of that is someone who has high levels of paranoia. They may have social withdrawal, they may have poor eye contact or rapport with an interviewer, but it may not be because of primary negative symptoms, it may be due to the paranoia. Antipsychotic effects can present secondary negative symptoms. And this could be something like Parkinsonism, where you have a masked face or decreased spontaneous movement,

but you can also have sedation, causing somebody to maybe have less drive or less social engagement.

Other mental illness can cause secondary negative symptoms. Examples of that are, depression causing anhedonia or a lack of motivation, or PTSD causing social avoidance. Environmental factors can be a big cause of secondary negative symptoms. If you live in a resource-poor environment, you may not be able to develop a hobby or engage with people, it may not be safe to leave your house and go walking in the neighborhood and make friends. Then there's stigma. People with mental illness face a great deal of stigma, and people with schizophrenia probably more than most. Some people with schizophrenia can internalize that stigma and start to believe what other people say of them, and that can affect them, they could withdraw socially.

So, I want to drill down a moment on the picture of negative symptoms, secondary

negative symptoms, and antipsychotics. Because as we talked about in the last slide, antipsychotics may be a cause of secondary negative symptoms, but antipsychotics may also treat some secondary negative symptoms. For example, antipsychotics may improve depression, they may improve someone's paranoia. If you have experience with doing clinical trials in people with schizophrenia, then you know that if you have a group of people that have an exacerbation of positive symptoms, and you enroll them in a study, when you do clinical ratings, over the course of the study you will see their negative symptoms reduced, but it's the antipsychotic effectively treating the positive symptoms and other secondary causes. We know that after effective antipsychotic treatment people still demonstrate negative symptoms.

And that leads me to the third P, persistent negative symptoms. So, negative symptoms often persist after you treat the

causes of secondary negative symptoms that you can identify. When somebody presents to you with negative symptoms, it can sometimes be difficult to tease out whether those are primary negative symptoms or whether those are secondary negative symptoms that just haven't responded to treatment. And this can be especially the case when you have a short screening visit to enroll in a clinical trial, it may be difficult to sort that out, but maybe these persisting negative symptoms, after you've tried to treat secondary negative symptoms, maybe that is the treatment target.

You can operationalize that population by saying people with persistent negative symptoms have some degree of negative symptoms, with low levels of positive symptoms, low levels of co-occurring mental illness, like depression, low levels of Parkinsonism on rating scales, and some clinical stability prior to enrollment in the clinical trial, so something along the lines of no

hospitalizations or changes in medication recently.

So, I'm going to bring up a few questions that I'd like you to keep in mind as we discuss things throughout the course of the day. First question is, what is our target patient population? Should clinical trials enroll people based on predominant negative symptoms, primary negative symptoms, persistent negative symptoms, or is there some better way of defining a population of interest? How should clinical trials ensure that we optimize treatment of secondary negative symptoms before enrolling in the clinical trial? Another important question is, how should development programs for drugs account for real-world antipsychotic use when designing clinical trials for negative symptoms? Thinking about this afternoon's talks, where we talk about clinical outcome measures, what's the best way to measure improvement? We have a number of scales out there, which one might be the best

choice? Are there several that might fit what we're looking for as far as an outcome measure? Should clinical trials measure negative symptoms as that single construct of negative symptoms? Or should we start to look at various factors and symptoms domains separately?

Should we account for cultural differences when we look at negative symptoms? I'm sure it's occurred to you as we were going through the negative symptoms that things like eye contact can be very culturally bound. People can have poor eye contact because that's how they were raised or that's part of their culture, not necessarily because of negative symptoms. So, when we have development programs that look internationally, how do we ensure that we account or those differences, and make sure that the results of those clinical trials are relevant to the United States population?

This is an important concept here,

how do we determine what amount of change is meaningful to a patient? So, we can see a statistical difference on a clinical rating scale on negative symptoms, but does that mean something to somebody with schizophrenia? Do we need to see some kind of a co-occurring functional improvement to put the results of that scale into a context? And then finally, how do we incorporate new technology into assessing our endpoints? Is it possible to use technology as a primary outcome measure, rather than a clinical rating scale, or is technology best used to inform the clinical rating scale, and we can decide what's relevant and how much change is meaningful by incorporating this technology?

So those are some questions to keep in mind. I'm going to quickly show the reference slides, but again, those reference slides are probably best looked at when the talk is posted online, and I'd like to now take a moment to introduce Mr. Staglin who's going

to talk to us about lived experience. Mr. Staglin is the president of One Mind, and he will be talking to us a little bit about his journey and why negative symptoms are important as a treatment target.

MR. STAGLIN: Hello. And thank you, everyone, for being here at this important meeting and for the important work that you do on behalf of people with schizophrenia. I'm Brandon Staglin, and as co-founder and Chief Advocacy and Engagement Officer for One Mind, I'm here to talk to you today about negative symptoms of schizophrenia.

So, I believe all people facing psychotic illness deserve chances to thrive. Why isn't that the common outcome? Like about 24 million of us worldwide, I live with schizophrenia, and I'm very grateful to have recovered from the darkness and debilitation and devastation that entailed for several years, however about two thirds of us don't get that privilege, two thirds of us who live with

the condition. This shouldn't be the case. Many of us live lives in limbo, recovery limbo, unable to engage with the world, unable to work, unable to engage socially with people around them, these can be lonely lives that end far too young. We need to address this. As a person with both professional and personal experience with schizophrenia care, I believe negative symptoms are the primary impediment to recovery for many, many, many people. Today I'll share how negative symptoms have impacted my life and the people I know, and also talk about ways to address negative symptoms that are derived from the experiences that I've had, that others have had, and also research.

I survived two schizophrenia episodes over a six-year span throughout my life. The first episode was essentially a vortex of psychotic dread. The positive symptoms were predominant in that episode. I experienced the conviction that if I made any moral mistakes over the course of any day

throughout the first six months of my psychosis, that demons would jump out of the shadows and drag me kicking and screaming into the abyss to spend eternity in misery and damnation. Needless to say, this provoked a lot of terror within me and worry and constant hypervigilance not to make any mistakes, so to speak, like stepping on a crack or eating too much food at a meal, but I'm very thankful to have recovered from that first episode thanks to my family's loving support and early science-based medical care like you all are delivering and improving, and staying involved with the community to rebuild a sense of agency and purpose. That was the first episode.

My second episode took place six years later when I was working as an engineer in Silicon Valley, and I had got into graduate school and I went off my medication in order to sleep less, as the medication had that as a side effect. And this episode was a devastating setback, and it took much longer to

recover from than the first episode, primarily due to negative symptoms. Although I controlled my psychosis pretty well, the negative symptoms dug in.

For three years I mostly played video games at home and drove around aimlessly throughout the countryside, not working, not motivated to be social. And while I was content with this for a while, because I wasn't embroiled in the turmoil of psychosis anymore, it was dawning on me gradually that there could be more to my life, there could be more that I could accomplish and achieve and experience. I began to fear this might be a dead end for me in my life, this limbo, these doldrums that I was in. I would give up easily on pursuing complex goals due to self-doubts and rumination. So, sometimes these self-doubts still haunt me today. Research indicates that avolition may be the result of dysfunctional reward anticipation. What if enhancing reward anticipation and calming rumination could

revive volition? Are there ways to do this? Something to consider. We'll come back to this.

Later I had the opportunity to use a new medication, aripiprazole, a partial dopamine agonist, that re-ignited my volition big time, it made me very ambitious, it made me want to be more physically active, and want to grow my social status, and I'm sad to say that I made some reckless decisions and comments to people that I loved, during the early period of my time with aripiprazole, mainly because I didn't have the cognitive control needed, the executive function and the attention to manage my volition effectively and to govern that for healthy relationships and productivity. I still regret some of the comments that I made to family members, telling them that they were terrible people, when I was in the throes of these symptoms and the medication's effects. So, it was very, very important to improve cognition while you improve the volition of

people dealing with these symptoms. Can enhancing cognition help to guide volition? Another question to consider as we progress through our day today.

Later, I had the ability to participate in a study conducted by our very own Dr. Sophia Vinogradov, seated there on the front, of a neuroplasticity-based treatment called cognitive training. This improved my cognition dramatically. Cognitive training is a treatment that uses the brain's ability to remold its neural pathways in response to targeted, gradually increasing in intensity, challenges that can improve people's ability to, say, pay attention, to focus better in conversation, remember what's being said. I benefitted from this treatment by doing it for about two months, and by the time I was done I was enjoying time with friends again and working again, it was a turning point in my recovery. This improvement in cognition dramatically unlocked my sociality, can this be

something that happens for people in the general population? Something to consider.

What else can enhance sociality and cognition? So, the studies of Dr. Nina Kraus at Northwestern University have shown that musical training can act very much like cognitive training in enhancing cognition, and in essence it can enhance the ability of the brain to be plastic, to change, to mold itself, and improve people's ability to do various cognitive tasks. But I've been through this dramatically in my recovery later on. When I was about 35, becoming more social, I decided I wanted to have a new hobby that would help me be more social, so I took up guitar. And another member of our audience, a speaker today, is Matthew Racher, also plays guitar. I had the opportunity to compose a song and perform it in front of an audience several times, actually, in 2017, about schizophrenia recovery. And when people sang along to the final chorus in my performance, it was like

such a thrill, and it gave me such a sense of satisfaction that my creativity could inspire joy and hope in other people.

And this was a huge motivating factor, huge boost of motivation, put it that way. It enabled me to do much more in my career, that boost in motivation enabled me to get a Master's of Science in healthcare administration from UCSF and took a yearlong program to do that, and become president of One Mind shortly thereafter. I'm very proud of what I did during that six years tenure as president of One Mind.

So, I believe music enjoyment is very important, because it's mediated by dopamine function, important for schizophrenia. Because it's mediated by dopamine function, and especially in the mesolimbic pathway, and by enhancing this dopamine function repetitively through musical enjoyment, I believe we can actually help people with schizophrenia neuroplastically to reduce their anhedonia and

improve their ability to enjoy things again as well. So, not only has musical training helped my sociality, it's also helped my motivation. Can musical training treat multiple negative symptoms? Question to think about.

Practicing sociality has also made me feel whole, so after the benefits I received from cognitive training and musical training, I became much more social. My adult social life and my ability to interact with people socially as an adult began to grow with my relationship with my grandmother, who you see pictured at the top right. Her name was Darlene, she was my close friend and confidant while I was recovering from my second episode of schizophrenia, and she was a great support for me during that time. As she got older and more frail, her needs began to outweigh mine, and I began to care for her, helping her relocate to a new assisted living facility near where my family and I lived, helping her to organize her time and her belongings at home, et cetera. I

learned from this relationship to care for people in general, to take care of others. This was a huge benefit for my growing sociality as I began to develop more relationships. Over time, sociality can develop into intimacy, where people understand and accept one another without having to have too many words used in communication for meaningful enjoyment mutually, and I'm happy to say, and grateful that I've attained such a state with my wife and our animals. My wife and I, who you see pictured at the lower right, have been happily married for 15 years, it's better every year.

So, I believe to strengthen sociality, it helps to practice it, once you have the tools needed to start. So, some take-home strategies for us to remember today are to strengthen volition, increase reward anticipation, and decrease rumination, okay? How can you do this? Well, there are some treatments that address this directly. For

example, one thing Dr. Vinogradov is working on in her EPI-MINN program, the Minnesota arm of the EPINET program for early psychosis care and research, is an app called PRIME, personalized real-time intervention for motivation enhancement. It addresses this directly, she can tell you more. Musical training can also help, as I mentioned earlier. Secondly, to improve volition successfully, also strengthen cognition, keeping it managed, keeping the volition in healthy channels. Cognitive training can help with this, as can new medications coming down the pipeline, from large pharma companies as well as smaller biotechs, like One Mind helps through its accelerator program.

And then thirdly, to improve sociality, enhance cognition, and follow up with opportunities to socialize with family members in ever-widening circles. How can you do this? Well, community clubhouses are one way that people can engage if they live with

serious mental illness with others, like the Fashion House, for example, in New York, coordinated specialty care programs often involve group therapy sessions and family focused therapy that can help with this, another group called Students with Psychosis is a group that connects young people with psychosis with each other and communities and for advocacy. Also, another upcoming One Mind at Work program may soon address this for young people in the workplace.

So, we're faced today with a shadow crisis of negative symptoms, I know so many families whose sons or daughters are trapped in recovery limbo due to their negative symptoms, not working, not being social, spending all their time at home, and this is a great source of consternation and turmoil for these families. I know one family whose son has overcome the avolition but still has the blunted effect of alogia, and very sadly, although he's very talented and skilled and

motivated to work, when he goes through the interview stages of his job applications, he fails because the employers can't see past that blunted affect to the great person that he actually is, and he doesn't get the jobs. So, loneliness as well due to negative symptoms is a huge problem. Not only is it corrosive to health, it's also a source of torture for so many people, just being alone so much. And I experienced this myself while I was recovering, but also Jeremy Novell, a colleague and friend, believes that this could be a cause of the early death of so many people with schizophrenia, the loneliness and the corrosive health effects.

As clinical scientists, you must deal with many people who have these symptoms and are in these situations. What can we do? As a scientist, you have the tools to discover solutions. You have the caring, the compassion, and the know-how to make a difference in millions of people's lives. This

is our task today, to tackle negative symptoms using the inspiration that I've delivered, that others will deliver, and the idea that many that will speak today will offer. Our job today is to figure out how to address these key obstacles to recovery at scale and to develop and employ these solutions to meet the needs of the community.

If we can keep young people healthy through our preventative and early care, leveraging innovations inspired by what I've suggested today, I believe we can save lots of lives and transform lives for the better. This is essential for the future of so many people as well as for our society. Thank you. Are you with me?

I do have a minute left. May I take questions if there are any? Any questions? Yes. I can hear you, although I'm not sure the mic is on.

DR. KIRKPATRICK: This has implications for clinical trials.

Can I whisper now?

MR. STAGLIN: I can hear you.

DR. KIRKPATRICK: Okay great. This has implications for clinical trials, if we think broadly, and that is there are a lot of reasons why even if someone had a medication that was effective for negative symptoms, there are going to be so many obstacles for that becoming obvious, and for people to have good outcome, to have good level of function. Some people have these problems begin in childhood and adolescence, and they may not have acquired the skills that they need, and certainly if someone has had this for ten years, I think it's going to be very normal for many of them not to have much confidence to go out and do these things.

So, if we have a medication, even with a small effect size, that makes them able to respond to the psychosocial treatments that they really need, social skills training, cognitive rehabilitation, et cetera, that may

be a really big deal, and we should keep that in mind. Also, it's just going to take a long time for people to show these results, to improve their function. Thank you for the talk, it was great.

MR. STAGLIN: Great comment, thank you, very true. Thank you. Okay, we're out of time, but --

DR. CAMPBELL: Brandon? We're here. I've got some web thoughts. So first of all, someone says, 'we're all with you, thank you for your remarks.'

MR. STAGLIN: Thank you, thank you.

DR. CAMPBELL: I have two quick questions if you're okay with answering them, if you feel okay. Is that okay Bernie? Okay. So, the first question is, did you find any benefit of musical treatment to your ability to discriminate and process sounds, and do you think that helped with cognition as well?

MR. STAGNIN: I found the musical treatments, primarily, yes, they helped my

ability to appreciate the nuances and complexity of music, so yes, I think as an auditory phenomenon, music is something that I was able to appreciate a lot better thanks to musical training, yes.

DR. CAMPBELL: I have one last question if you're okay. All right. You did not mention the lack of awareness of negative symptoms we see in many patients. Can you speak to this behavior barrier to treatment?

MR. STAGNIN: Yes, I can. So during my second episode, I can't remember if I mentioned this, but I was fairly content to live a life of not engaging in work or not engaging in social activities, and I didn't really care that much about it, I wasn't aware that I was missing anything at that time, I didn't feel like I was missing anything, but it took a while, I began to think that there could be more to life again, there could be more that I could accomplish, remembering how I used to be, basically, and the achiever that I had

been. And so yes, for quite a long time I wasn't really aware the negative symptoms were affecting me, but eventually I did become aware. Thank you.

DR. RASETTI: Good morning, everyone. Can you hear me, can you see me? Welcome to Session 1. My name is Roberta Rasetti. I am a psychiatry and a clinical reviewer for the Division of Psychiatry in FDA, and I will be moderating session one, session one is on the brain circuits and relationship to cognition.

This session will be a brief overview of the current science on the neurotransmitter system in the brain circuits related to negative symptoms. This session will last 30 minutes, we will have the first 20 minutes with the presentation, by the speaker, and then there will be 10 minutes of a Q&A. After this session, we will have the first coffee break that will last an hour and ten minutes.

Okay, so now it is my pleasure to introduce our speaker, Dr. Sophia Vinogradov. Dr. Vinogradov is Professor and the Department Head of the University of Minnesota Department of Psychiatry and Behavioral Science, and she leads the Translational clinical neuroscience lab focused on the cognitive dysfunction in psychosis, and also she leads a network of early psychotic clinics in the State of Minnesota. The title of her presentation is Negative Symptoms, Cognitive and Neural System Features. Thank you.

DR. VINOGRADOV: Good morning, what an incredible gathering of individuals, and you may not believe this, but Brandon and I actually did not talk before today, and we somehow had some mind meld going on, because Brandon, set me up perfectly for my topic. The other thing I want to say is that the topic of the interplay between negative symptoms and cognition and underlying neural circuitry could be a weeklong workshop in itself at the end of

which we would come to no conclusions.

So for those of you who are not researchers in this field, and I have really so many incredible colleagues, I feel like I'm standing on the shoulders of giants, but for those of you who do research in this area, please forgive me, I beg your indulgence, because I'm really simplifying, I'm going to simplify for the sake of 20 minutes today a really complex set of topics. For those of you who are new to this area, I'm going to beg your indulgence, because it's going to appear much, much simpler than it really is, but perhaps this is a starting point just for kind of shaping some of our kind of thinking as a group going forward, and as you're going to see, Brandon's personal experiences and the conclusions he has drawn from his lived experience really do fit well with some of these general themes that I will provide in an overview to you.

So let's see. Okay, did I go in the

wrong direction? Here we go, okay, sorry about that. My disclosures. We've already heard from Bernie about the five negative symptom domains, we've heard about the two factors and I'm just refreshing your memory here, because this is going to be kind of very germane to where the research has evolved recently in thinking about these interplays which I was describing. Bernie already described to us the importance of negative symptoms in terms of functional outcomes, and has already mentioned that we've got this consistent association between negative symptoms and poor outcomes, and what is very interesting and important is that we see this association, that negative symptoms are predicting some additional variants and functional outcomes, even when we start to account for general cognition and functional capacity in individuals.

More recent work, Tony Ahmed and his colleagues, who are here today, have done some very interesting, I would say, kind of external

validator analyses, and have shown that when you drill down into some of the specific domains, those five domains that we described, we see some strong and consistent domain-specific associations with functional outcome, particularly with the avolition anhedonia factor as you've already heard, and to a lesser extent with blunted affect. And again, this point about these five domains that's important to keep in mind is that, when you do this kind of external validator work, each of these different domains is showing some specific associations with these range of external validators that aren't accounted for by the two-factor approach, and that includes functional outcome, as I mentioned, psychological measures, such as defeatist beliefs, which Brandon has alluded to, cognitive function, and neural system findings.

So another kind of, in a sense, accepted [concept], now, with decades of research, is that we see a consistent

association between negative symptoms and disrupted cognitive functioning, but as you look historically at the research and then kind of fast-forward to where we are today, there's some inconsistencies, and this is because in the earlier body of studies, there were aspects of measurement overlap and clinical compounds that again, you've already heard partially alluded to. So for example, in some of the earlier rating scales of negative symptoms, cognitive observations were a fundamental part of how negative symptoms were rated, and so then of course we start seeing these relationships between more severe negative symptoms and more severe cognitive functioning.

There was a recent systematic review of about 3,000 individuals with negative symptoms, this was first episode psychosis, and the interesting thing about this sample, of course, is that this is going to be individuals who have not yet necessarily had a long amount of chronic exposure to some of the impoverished

environmental resources, the accumulating effects of medications and so forth, and in this sample, there was a consistent association between negative symptoms rated more generally, using a range of different ratings scales, with lower executive functioning and poorer theory of mind. So what's interesting about that? Executive functioning, right, generally cognitive capacities, certainly strongly pre-frontally mediated, poorer theory of mind social cognition capacities. And then in Ahmed et al.'s study which I mentioned to you, avolition, across three different samples, was showing the strongest and most consistent association with disrupted cognition. There were also some associations seen in a couple of the samples with anhedonia, blunted affect, and alogia. So again, sort of the take-home messages, yes, we do see this association with negative symptoms and cognitive functioning, there is some of these confounds with measurement and other environmental exposures

likely contributing to that, avolition seems to certainly be playing a central role.

One of the other consistent findings in this field is that when we look at samples of individuals with more persistent negative symptoms, we see in those samples a fairly consistent association with structural and functional changes in prefrontal and temporal cortex. So, for example, again, looking at early psychosis samples, when we see individuals with these patterns of prefrontal disruption, such as progressive cortical thinning in prefrontal cortex, in these individuals we see a more severe course of illness, worse functional outcomes, and increasing negative symptoms over time.

So there certainly seems to be this kind of picture that emerges, that when you look at persistent negative symptoms, perhaps persistent and more predominant negative symptoms in individuals, over time we see this relationship with disruptions in prefrontal and

temporal cortex, which of course is the central executive network, right. This is the network in the brain that is responsible for essentially integrating information, coming from all different, kind of sensory perceptual memory inputs and so forth, integrating them in order to be able to carry out abstraction, problem solving, anticipating the future, creating value representations of future behavior.

More recent research, though, and this has been sort of a really exciting explosion in the field, is that as more recent research has drilled down into some of these specific domains, what has been sort of kind of emerging, again, in a fairly robust way, is the relationship between anhedonia and neural system findings. I'm sorry, avolition and anhedonia. So again, that sort of often, it's often considered that motivation factor, and very specifically again with avolition, avolition has been associated with lower

glutamate and GABA concentrations in anterior cingulate cortex. Avolition has also been associated in a very recent study with amplitude of low-frequency fluctuations, which represents intrinsic neural activity across multiple cortical regions, prefrontal, even more posterior, anterior cingulate, and again, avolition as a domain showed this association in a more specific and enveloping pattern than some of the other domains did, although some of the other domains were showing, again, drilling down, some very interesting specific associations.

And then looking at the motivation factor for avolition and anhedonia, the patter that is, again, consistently emerging is that it's associated with reduced ventral striatal activity and disrupted connectivity between the ventral striatum and other regions of the brain. And so now we're talking about the reward processing circuit, right? So ventral striatal, mPFC, again, new to some of you,

medial prefrontal cortex, very important in valuation, reward processing, social cognition as well, orbital frontal cortex, anterior cingulate cortex, and this network, again, we kind of, in a sense, call the reward processing network.

I want to come back to the idea of functional outcome, because there's been some interesting analyses, path analyses, structural equation modeling, kind of, that give us some of these clues about the relationship of neurocognition to negative symptoms, and you're going to see some of the themes that Brandon was alluding to emerging here.

We know that neurocognition has a direct relationship to functional outcome, but we also know that it is partially mediated through the effects that neurocognition has (fundamental neurocognitive capacities have) on the brain's ability to carry out social cognitive operations. Social cognitive operations are, even at the most simplistic

level, are things like eye gaze detection, emotion recognition, vocal emotion recognition, theory of mind, in terms of a higher order social cognition capacity. So, we know about this relationship, that's been well established in the literature for several decades now.

What more recent work has shown, out of Giordana et al.'s lab, is that if you kind of elaborate upon this analysis now with measures of negative symptoms, we also see that the effect of neurocognition on functional outcome is being partially mediated through negative symptoms, both the kind of domain of motivation, anhedonia, avolition, as well as the domains related to blunted affect and alogia.

However, when you look at the relationship of social cognition to functional outcome, it is partially mediating -- the relationship of social cognition to functional outcome is being partially mediated particularly by this domain of avolition and

anhedonia. But there's a final kind of piece to this puzzle, which, again, shows you a little bit about its complexity and about how we still have not fully resolved kind of where all the pieces in the puzzle fit together, and that's the fact that, and again, in a number of different studies, if you just look at the right hand part of these paths, this relationship of neurocognition and social cognition to functional outcome is strongly mediated also then by their effects on motivation.

And we know that, again we see this over and over in a number of studies, that if we have -- that we can both target these as treatment targets, and we've done that in my lab around targeting social cognition, targeting neurocognition, we can see improvements in motivation, and distal improvements in functional outcome, as Brandon alluded to. And yet at the same time, we know that some of these relationships are being

partially mediated through the effect of negative symptoms, and yet negative symptoms contain within them this motivation factor, right, of avolition and anhedonia.

So, I want to say a few more thoughts about motivation and avolition, and kind of, again, sort of summarize some thoughts about it. So, now we're talking about motivation as a concept, and maybe talking a little more generally than we had about just avolition, but motivation as a general concept, as it's been increasingly studied in cognitive neuroscience for the last 10, 15 years. Again, it is really understood as arising from the interaction of two major neural systems, and I underline the word "interaction."

You know we as researchers, as people trying to sort of put together complex ideas, we like to pull things apart and say oh, there's a module here, the central executive module, there's a module here, the reward processing module. That's not, of course, at

all how the brain works, but in terms of the circuitry, again, that central executive network, dorsolateral prefrontal cortex, dorsal caudate, - I'm simplifying a great deal of work here - but, essentially, again, as part of what central executive functions do, they're encoding the relationship between actions and potential outcomes, that contingency, it can be representing the expected value of an action, if I do X, this will happen, and it has this value or meaning to me, and then of course there's really strong functional overlap here with cognitive control mechanisms and intentional control mechanisms, which we know are disruptive in the illness.

And if you want to simplify a great deal, you can think of these central executive networks as being really critical for just our capacity to do these higher order functions that allow us to move adaptively through the world, problem solving, a distraction, et cetera.

The other major system, as I've already alluded to, is the reward processing network, ventral striatum, and its related protections and connections. These of course are much more kind of taking on the operations that are related to anticipating reward and then valuation of a reward, as one is anticipating it, and as one is receiving it. The representation of stimulus-reward associations as they're happening, and of course this system them talks to the prefrontal cortex and said, "huh, there was an important value based, you know, reward-based association happening here, learn about it, I want you to learn this," right? Those are those dopamine projections to prefrontal cortex. And this circuitry has a high degree of functional overlap with value-based, reward-based decision making, and social cognition.

And I'm going to do a footnote here, because again I think it was embedded in your talk, Brandon, which is that our social

cognition capacities, our ability to make sense of the social world in which we move is critically important to our survival, right, we're social mammals, and it's very intimately connected to the reward processing system, and in fact, social stimuli are known to be in and of themselves, most of the times, rewarding, right? To see a face, a face that you know, to see someone smiling at you, to smile back at them, to have that reciprocal interaction, that is innately rewarding for the brain, and these are the neural systems that are recruited during those processes.

So, again, keeping things super-simple, we have now an important interplay happening between cortical systems, cortical circuitry, and subcortical systems. Thinking about models of schizophrenia etiogenesis, right, this is the interplay between cortical excitation and inhibition balance, glutamate-GABA, and subcortical and cortical dopamine modulations, so dopamine - that's originating

in the midbrain -, both in terms of what's, kind of the circuitry in terms of the ventral striatal and dorsal caudate, but also the projections of this midbrain dopamine into the frontal regions.

And, of course, as I said, in terms of etiopathogenesis of schizophrenia, we have theories related to models related to excitatory inhibitory imbalance in the cortex, glutamate-GABA imbalance, and of course we have this notion of subcortical hyperdopaminergia in the dorsal caudate, and that these two systems are out of balance, and that's giving rise to the different symptoms that we see.

Again, people simplify and say well, it's an EI (excitatory-inhibitory) imbalance at the level of the prefrontal cortex that's giving rise to cognitive and negative symptoms, and it's this hyperdopaminergia at the level of subcortical systems that is playing a role in the development of positive symptoms of psychosis. We know it's much more complicated

than that.

For example, as I think we are all aware, and we probably will hear a little bit more about today, we know for example, this is just one kind of additional, in a sense, level of complexities, we know that muscarinic modulation effects these symptoms both cortically and subcortically, we also know that, again, as I've mentioned, the dopamine projections, the D-1, D-2 projections, the effective D-3 receptors on what's happening in terms of dopamine modulation of prefrontal cortex can also be targets that are affecting, essentially, the GABA-glutamate interactions, the capacity of the prefrontal cortex to learn, to engage in its higher order cognitive control operations. So, if nothing else, I want to just emphasize that these are useful heuristics for understanding what some of the neural systems and neural transmitter systems are, that play a role in negative symptoms, but by no means should we think that the story stops

here.

I'm going to end with two points I'd like to make, and this is sort of my perhaps solipsistic thoughts about where we might be headed in terms of understanding more about negative symptoms in general, and its relationship with cognition.

Our lab, which by no means we are not the only lab, there are other major labs around the world that have been really looking at reward-based trial by trial behavior in individuals and doing computational analyses of this behavior. And what's really interesting about this approach is you can come up with tasks that can be done across species and look at these variations in trial by trial behavior across individuals within a species, in humans who might have illness, who don't have illness, who are on a medication, or not on a medication, and you can start to pick up some of these interesting aspects of "how is reward sensitivity effecting behavior of the

organism?" And what we have seen is that when we do these trial by trial analyses, that impairments in reward sensitivity, and that means how much is your decision on a given trial affected by the reward you just saw on a previous trial, and we see a linear relationship between impaired reward sensitivity and avolition ratings in humans, but we can have animals do these same tasks, and we can manipulate the animals.

In the monkeys, we can give ketamine to the monkeys and disrupt the GABA-glutamate balance, we can do genetic manipulations of the mice and see how these different genetic models change this reward sensitivity, you can do pharmacologic manipulations, and so on. So, this is a growing area that allows us to bridge the gap between something which is so subjective as motivation or lack of motivation, and actual animal behavior. And we're going to be just seeing lots of exciting results emerge from that kind of work.

And I want to conclude with just throwing up some data, because it sort of ties together nicely with what Brandon presented to us. It's really important to remember that manipulation of these systems can happen through behavioral means, not just through pharmacologic manipulations, and we have shown just recently in a trial that was carried out entirely remotely across the world that when we offer intensive social cognitive trainings, we're really training intensively the circuits that have to reliably and efficiently and quickly pick up socially relevant information, again, eye gaze, spatial emotion, et cetera. And we pair that with an app that deliberately creates goal-setting and social network support for goal-setting for individuals with early psychosis. That combination drives changes in motivation measures -these are motivation measures done by blind raters - and these are correlating with changes in defeatist beliefs, which are in turn associated with changes in

functioning six months later. So that's just to give you a little bit of a flavor of the different ways one can do behavioral interventions, cognitive interventions. I should also say, we've also seen just doing straightforward cognitive training - just to improve some of these aspects of higher order cognition, as well as lower level kind of perceptual information - we see some associations with improved motivation measures, particularly when we add in some social cognition training.

All right, in sum, negative and cognitive symptoms are considered separate domains of psychopathology, but we see these consistent associations, these shared features. Negative symptoms in general, but particular avolition and amotivation, seem to be specifically mediating aspects of this relationship. We have the frontoparietal central executive network, which seemed to be more implicated in persistent negative symptoms

or worse outcomes, and of course play a key role in general cognitive capacity.

We see the ventral striatal reward network strongly implicated in avolition and anhedonia symptoms. We see the promise of computational modeling of reward-based kind of learning and behavior, as I think this incredible new tool has evolved. And in our work and others, well-designed cognitive training strategies can improve motivation measures and functioning, and, the most important piece, echoing Brandon, is that cognitive and meta-cognitive factors affect the expression and impact and the evolution of negative symptoms over time. Thank you.

DR. RASETTI: Thank you very much for that wonderful presentation. So, we do have like 10 minutes for Q&A, and I think we can start. We can open it up to the floor, if there are any questions, and then if there are any, we can open it up to online questions.

PARTICIPANT: Hi. Thanks, Sophia,

for the great talk. I'm curious in your studies of cognitive training, have you ever looked at what percentage of those patients would meet criteria for having predominant or prominent negative symptoms?

DR. VINOGRADOV: You know, that's a wonderful question, and, preparing for this talk today, I realized a big blind spot that I've had as a researcher is lacking a focus on negative symptoms. I would say, overall, and we've focused so much on cognition as our outcome measure of interest, and we've only looked at sort of symptoms secondarily, that it's like a big lacuna, or has been for me, but now I want to sort of go back and see if we can retrospectively take a closer look.

So, I could not tell you kind of the percentage or the composition of our kind of samples, our participant samples, their baseline in terms of negative symptoms. I can tell you that we have seen, in one study with first episode patients, that when we do a six

month follow up after the cognitive training, we see an improvement in positive symptoms that reaches a statistical significance and a trend towards improvement in negative symptoms. So, we've seen that relationship, but I couldn't tell you at the outset what the relative kind of composition of the sample was.

PARTICIPANT: Yeah, I'm curious, and I'll look forward to hearing that. When I was working more closely with Rich Keefe and Phil Harvey, we were looking at some very big data sets, CIAS trials, and I'm saying this because I think this issue is going to come up a lot today. The overlap between patients in CIAS drug trials who met one of these predominant or even prominent negative symptom definitions was typically south of 20%, so it was generally pretty low.

DR. VINOGRADOV: And what's your sort of thinking or your interpretation?

PARTICIPANT: It's going to be really hard to do trials of patients who have

pure predominant negative symptoms, to get them done in a reasonable amount of time.

DR. VINOGRADOV: Yeah, yeah.

DR. HORAN: So, I was struck by the overlapping neurobiological basis of both cognitive impairment and negative symptoms. So, I have a sort of psycho-pharm question that also bleeds into regulatory to some point, right, so I'll pick somebody else's drugs.

You can imagine a kappa-opioid receptor drug, right, could regulate striatal dopamine activity, improve motivation, reduce anhedonia, and have direct effects on negative symptoms, that have corresponding effects on cognitive performance, right? Either directly or indirectly through engagement, task engagement, right, because we ultimately measure cognition through cognitive measures that require you to be engaged. I think Phil talked about engagement.

So, you can imagine getting it that way, but you could also look to something like

luvadaxistat, right, a DAAO inhibitor that may improve executive function, right? So, in a sense, directly improves cognitive function with a downstream effect on negative symptoms. And so how important is it to disentangle these effects?

You know, to Bill's point, you don't have these overlaps that you might expect, although there's correlation there, but you give them this very difficult decision making process of "are you a pro-cognitive agent, are you a negative symptoms agent, are you both," and how do you really design effective trials and establish efficacy when, realistically, at least from a neurobiological basis standpoint, these functions are so overlapping? There's no answer here, I just wanted to stand up here and spout about this for a second.

DR. VINDOGRADOV: You know, and I have to leave the pragmatics of that to my drug development colleagues, I think as a researcher, the geeky side starts to go "That's

so interesting, which is causal of which, right?" And "Oh, you know, both are going to be useful targets," and "One is going to, as you say, may be driving -- you know, it is a primary factor to the improvement in executive functioning and therefore the ability to anticipate reward action outcomes, and now with stronger executive functioning maybe you're not having as many ruminations, and now it's easier for you to engage, blah, blah, blah," right?

So, I know as a researcher that becomes super exciting and interesting. I think from a pragmatic point of view, in terms of trial design and what are your primary endpoints and so on, I don't want to touch that one, I'll have to let colleagues talk about that later, yeah.

DR. HORAN: Thank you.

DR. HARVEY: I just wanted to raise a point about the unitary nature of motivation and engagement. We finished a large-scale

ecological momentary assessment study, and about 20% of our patients were home and alone during 85 out of 90 surveys that they answered. One would interpret that as being social amotivated, but they answered 85 out of 90 EMA surveys when they were home and alone.

So, what this suggests is that they weren't completely unmotivated, because they were doing something when they were requested to do it. Maybe you could say it was because of the dollar they were getting when they answered the survey, but it suggests that there are layers of motivation. And I think your positing the complexity of the interaction between cognition and motivation may help resolve some of those questions, because one would think that you would view some of these digital phenotyping strategies as being very hard to pull off in people with significant negative symptoms, but that's how we know they have such significant negative symptoms, is because they're answering the digital

phenotyping surveys.

DR. VINOGRADOV: Right, and so two important points there in terms of actual kind of, you know, working in the real world, which is because of some of these motivational deficits, it's challenging to have participants want to engage in these intensive treatments without some sort of very concrete reward attached to it.

Abstract reward is difficult. So we, in our first episode work, we find that offering sort of cognitive training and use of the motivation enhancing app, just offering it, the uptake is low. If we offer it along with, "here's a chance to earn a few dollars every time you use it," the uptake is much better, right? it's almost like we have to hijack the motivation system initially to kind of get it jumpstarted in order then for the individuals to engage in these more plasticity-based treatments, right, that eventually will allow for the symptoms to be much more self-

perpetuating. So that's point number one.

Point number two, and I think I was just hearing from Dr. Strauss about really this strong relationship between the kind of motivation levels, and forgive me if I'm not getting it right, that people are experiencing over the course of their days and their weeks, and where they are in terms of an environment, whether the environment is enriched, not enriched. You know, so it becomes almost a social deafferentiation situation, or a sort of general stimulation deafferentiation, in that there's this again, interplay between environmental impoverishment and motivational, the kind of motivational systems in the brain. Did I get that right? Pretty much?

DR. SCHOOLER: So I just wanted to get back to the issue of pharmacologic trials, and from what I'm hearing you say, one of the questions that becomes important is actually, I can refer to the movie Groundhog Day, which I think everybody remembers, and the fact that by

the end of it, the character can do all sorts of things that he couldn't at the beginning, because he's repeating the same day over and over. So the question becomes, what would the role of cognitive training be in evaluating a medication that was primarily designed for negative symptoms and improving motivation?

And I just wanted to raise one other point, which relates to Bill Horan's comment about the very low percentage of people in their trial, in their study that had predominant negative symptoms, and that's that people with lack of motivation are not motivated to participate in trials. And I just want to link what you've been talking about to what the goal of the session today overall is, and I wonder if you have comments on it?

DR. VINOGRADOV: Just that in real world clinical practice, the patients I worry about the most are the ones that are not motivated to come into their treatment, right? And those are the ones who have the poor

outcomes, and again it becomes that sort of downward spiral. They have predominant negative symptoms, they don't engage in treatment, they don't engage in studies, they don't engage in trials, et cetera.

But in terms of, like, in the ideal world, I mean, I know the focus here is drug development and I don't want to take away from that focus, but we all know that the optimal kind of treatment approach for individuals with psychosis spectrum illnesses is multi-modal, and in my fantasy world we would be combining these motivation enhancing negative symptom addressing medications with kind of evidence based behavioral and psychosocial interventions.

And I honestly think the combination together could be done quite efficiently, like I think we could get significant improvements over two to three months, which is something you can sell to, you know, young first episode patients and their families.

MS. RASETTI: Thank you very much. I think we are running out of time.

We're going to have now a coffee break of 10 minutes, and then we will start with Session 2. Thank you.

(Whereupon, the above-entitled matter went off the record at 10:20 a.m. and resumed at 10:31 a.m.)

DR. BLACKMAN: We will try to stay as close to on time as possible. I know we're already running a little over, so let me just get started with some introductions and logistics as people sit back down or get back to their computer.

I just want to introduce myself. I'm Dr. Rachael Blackman. I'm a clinical reviewer in the Division of Psychiatry here at the FDA. I will be moderating Session 2, which is on study design.

So, in this session we will focus on the challenges of designing studies and clinical trials to assess the effectiveness of

negative symptoms in schizophrenia. Currently, there is no consensus on the best way to design clinical trials for negative symptoms in schizophrenia, so this is important conversation for us to have.

Before we start, I just want to make clear that there are currently no approved medications for negative symptoms associated with schizophrenia. Therefore, any drugs mentioned by the speakers or panelists today are either investigational or being used off-label in any examples you may hear about.

Now let me describe the format of the session. We will hear from two speakers, and then we'll have a panel of esteemed respondents. So I will introduce all the panelists after the speakers have finished and we will hold all questions until the end of the session, during the Q&A part.

And now it's my great pleasure to introduce our first speaker, Dr. Cristoph Correll. Dr. Correll is a Professor of

Psychiatry at the Zucker School of Medicine at Hofstra/Northwell in New York, as well as Professor and Chair in the Department of Child and Adolescent Psychiatry at Charite University Medicine in Berlin, Germany. Dr. Correll completed his medical studies at the Free University in Berlin, Germany, and Dundee University and Medical School in Scotland. His focus has been in areas such as identification and treatment of severe mental illnesses, psychopharmacology, and clinical trials. If you've read any articles in recent years, chances are you've read one of his publications, because he's highly prolific and very well cited. Dr. Correll will be talking about considerations for drugs designed to be adjunctive to antipsychotic medications today. Dr. Correll...

DR. CORRELL: Thanks so much for the kind introduction. We're unfortunately 15 minutes late, but I'll try to stay on time. I've already been reminded to do so, so that

our panelists can really have their two minutes. I was complaining I only have 20. But when I heard you only have two, I'm okay with that.

So there are many great minds in this room; there are many great minds listening. And I think it's wonderful -- kudos to the FDA to convene this meeting -- that we try to put our heads together, as has been done, like, over a decade ago. Steve Marder was leading that effort. And the question is, have we advanced since then? And I would hope that out of this meeting we can also have another paper and consensus come out of this. How to not only raise problems, but maybe even solve them, or make some suggestions what the next avenues should be.

So this is my disclosure information. And this is what I would like to cover. A lot has already been in my first section alluded to, so that makes it easier. So the negative symptom considerations will be

easier and faster, but then treatment considerations and trial design and some recommendations will be the meat of this.

Now, you've already heard that we are talking about predominant, primary, and persistent. But there's also been a question, should it be really predominant? Or how much admixture should we allow in order to be more real-world? And where's the pseudo-specificity coming in?

Also, there's another question about pseudo-specificity of comorbidities, either dimensionally or categorically. Here is a paper that looked at the transdiagnostic presence of some of these symptoms across diseases. And are we excluding and making the sample squeaky clean? Or are we allowing some of these comorbidities in, especially anxiety disorders, so that we even have a pool of patients and can generalize to the patients out in the real world? But, then, what does that mean? Where does the negative symptom come

from?

We've already heard about that there are different components in negative symptoms. And here I just took out three studies. Although the negative symptom component -- or the percentage overall was different across the studies, you can see that the distribution, which of the negative symptom domains was most frequent, was consistently the highest in the social amotivation. So, for each of the studies, that was the highest. So should we enrich for that? If we have more patients with these symptoms, would a drug work particularly well if it captures that aspect of it? Do we need to look much more, not just at the outcome, but at the in-come, what the patients look like? And, particularly, do we want to enrich? Do we want to even approve, or consider going for approval, for medications that have specific effects rather than global effects?

We've already heard about depression

being a potential confounder. And it's interesting that antidepressants, in patients who say they are not depressed and have schizophrenia with negative symptoms, seem to work, with an effect size of about 0.28 overall, 0.37 when the primary outcome is negative symptoms.

So, should we rule out some of the patients that have depression based on an overall depression scale or specific items in depression? So, should we exclude patients who have pessimism or suicidality, because that might be much more related to depression, rather than the blunting that we see otherwise with schizophrenia and negative symptoms?

And already we've also heard about secondary negative symptoms, that we need to rule out potential effects when patients improve on depression, or EPS, or sedation. Do we have a stay or switch design where both can move? Or do you have a drug that has less EPS, or even treats EPS, or less sedation? Does it

help with sleep or reduce pain, which can all also mimic negative symptoms? That's all very relevant. And, obviously, with the predominant versus primary or prominent symptoms of negative type, we need to talk about the positive symptom overlay also.

Now, negative symptom treatment considerations. What has been done in the past? We did an umbrella review of 42 augmentation strategies. And that's my talk. Steve Brannan will talk about monotherapies. And we were quite flabbergasted how many, when you look meta analytically, how many treatments actually are better than placebo. Wow; that's quite a lot here. And they have effect sizes of 0.2, 0.3, 0.8. Really? One, 1.4. This is all our meta-analysis, SNRIs. That's amazing. So we have all these great treatments that we should use in clinical care, but somehow it doesn't seem to work. And they should all be approved. And nothing is approved. So what's going on here? And that's for non-clozapine

treatments.

And here we have clozapine where nothing seemed to work.

And so we thought, wait a second, this is weird, and looked at the quality of the meta-analyses. And, actually, the meta-analyses were done well. Fifteen of the 32 meta-analyses were mine. We weren't doing such a bad job there. But, then, wait a second. Why are they good? Because we have quality metrics. Every meta-analysis now has to follow certain procedures. So they were done well. But, what's going on? There's a disconnect. So we created another quality metric. And that's not for the meta-analyses, but for the studies that enter them, and for the consistency between the individual studies and the overall outcome. And that's where it all fell apart. So, basically, this goes until eight as a quality score. And you see there's nothing higher than five. And all of the significant effects had three or four or two of

the quality. And that is because they were all small. There was a big publication bias. Because if you have a small study and it doesn't work, well, then it's not powered. But if you have a small study and it works, yes, we found something, this is a great lead. So, we have a real problem. In none of these meta-analyses, in none, did the larger study ever confirm what all the individual studies that drove the mean effect actually showed. So we need to get away from these meta-analyses that put everything in and celebrate. But it also tells us that the cliff from phase II to phase III is enormous. Because we have a small study with little expectation, that works. And then everything falls apart when we do the larger study.

And we also need to consider what happened in prior trials. And there are people here in the room that have driven that forward. So this lack of approval of any treatment for negative symptoms is not for lack of trying.

We've tried. So why have failed? Is it the design? Is it the molecules? Is it both? And we're obviously here in order to hone in on the design so that we can make it easier for molecules that could work to actually show this. Now, this is a phase II trial with bitopertin, a glycine transporter 1 inhibitor. A $1 billion endeavor that failed, unfortunately. And the first thing that I've learned from clinical trials, when your phase II study that has little expectation just scrapes the 'p' less than 0.05, forget it. You don't have any buffer. It will get lower, the effect size. And if you then have to do a sub-analysis on per protocol on patients that actually fulfill the protocol, and only then it works, you have even less of a buffer. So we need to learn from phase II, and maybe do another phase II, to see which of the patients do better. But what was also interesting is that our usual linear, higher dose is better than lower dose, doesn't always work for

certain of these mechanisms. So here there seemed to be an inverted U-shaped curve. Do we need to learn from that? But what you can see is that at eight weeks, which is considered relatively short, there was around 6.5 point improvement for drug. And there was about 4.5 for placebo.

That's basically eight weeks here of 6.5 and five points. It's interesting that when you now go for not eight or 16, but 24 weeks, or three times the duration, the effect at eight weeks was much lower. Why was that? I mean, that's weird. Because if at eight weeks you already have so much drug and placebo effect, now people know I have three times the time, so maybe the expectation is it will take longer. And so at the end of 24 weeks there was not much higher, actually, a similar effect for the drug. The drug didn't move. But the placebo effect, obviously, got larger. So that's a problem here. Why does placebo increase and drug doesn't increase? We see

that with positive symptoms or total symptoms. Generally, the placebo effect goes up, but the drug also gains.

And then the most recent effort to look at pimavanserin, a 5-HT2A inverse antagonist. And it showed at Week 26 in phase II, again just a barely significant effect. The effect size was 0.21, 11 versus eight points. Yes, there was a dose effect. In this case, the higher dose was better. It actually had an effect size of 0.34. That's something you can rest on and say, okay, 0.34, that's maybe the minimal effect where it can see something, even at a larger study. So phase III study was begun. But should we have trusted these results? How consistent do they have to be? Well, first thing was, unfortunately, the drug again in phase III had the same effect, about 11 points, but drug placebo caught up, and you can see the effect size dwindled from 0.34 for the 34 milligrams now to 0.07.

But should we have trusted the phase II effect size and effects? Well, do we need to also triangulate and understand different measures measuring different things of the same construct? I mean, we've talked about confounding. But also it could give us some estimate of robustness of finding. So in the phase II study there was the small effect size, 0.21 for the NSA-16 total. And, again, this is not divided by dose. But you can see that it didn't really generalize across different domains in the NSA-16. It only separated with an effect size of 0.26 for one single aspect. So, out of five domains, only one. Is that enough to do a phase III trial? All right, maybe it is, because it all adds up with small effect sizes. But then, does it have any clinical relevance? There was basically no CGI-S effects. So if it has a negative symptom effect, why doesn't it, in predominant negative symptom, move the needle overall? And then, yes, we're not happy with the usual PANSS

negative symptom score because two of them are cognitive. But they should correlate, correct? And there was, again, nothing there. So is that enough then to say, yes, we have our primary outcome and need to move forward? Do we need to look at consistency? Do we want, actually, to measure different aspects and be sure that they move in the same direction, or at least carve out subdomains?

And we talked already about the placebo effect. So, this is a meta-analysis of placebo effects published in 2019 by Fraguas et al. And you can see that, overall, across different agents with different mechanism of action, the effect size was small, 0.2. But there was huge heterogeneity: studies where placebo beats drug and some studies where it's basically a negative finding, and some where there's a big effect size.

So the question, then, is: what drives this?

Wait a second, sorry, I forgot about this.

So this is the effect size of placebo from baseline. Remember, 0.2 is a small effect size, 0.5 is medium, 0.8 is large. Three, the Cohen's d is three. And it goes from basically 0.2, almost nothing in these studies that I just showed with bitopertin, to 14 and 12, if that was calculated correctly. I'm a little bit doubtful that it's that high, but it was calculated and published. But that's enormous, obviously.

How can we mitigate the placebo effect? So, here are some regression analyses. And what you can see is multiple factors in univariate analyses drove the placebo effect. Actually, higher study quality drives more placebo effect. That's confounded, because the good placebo control lab studies that have good metrics actually were negative. Later year of publication, our placebo effect has gone up, both for overall effect but also negative symptoms. Longer duration of the trial, interestingly. But what really drives it is

higher number of arms in the trial. So if you have too many doses you have a problem. Too many sites. These are all the industry-sponsored studies, obviously. Number of countries. Also number of patients. That's because these are phase III studies. Interestingly, also the mean age, lower mean age, but only in the placebo arm. And also severity of the lower severity of the negative symptoms drove also the placebo effect. Industry funding drove it, because these are the larger studies. When you do a multivariate meta-regression analysis all the other factors, some of them that might sound a little strange are basically driven by these three elements: industry funding so these are large studies, with a higher number of arms, and higher number of sites. So maybe you want to have a small study. But the small studies are usually done when you don't know whether it works. When the drug works, everybody has expectation.

So I want to finish up in the last

five and a half minutes by drilling into some of the design issues that are overall relevant for the effect of medications for negative symptoms. And then particularly focus, on the next slide, on what might be relevant for the design of augmentation studies.

Some of it will be discussed also in other sessions here, especially Session 4, which is very important, about which outcome measure do we actually use. So, population, which age should it have. Is it inpatients or outpatients or mixed? What's your recruitment frame? How do you get them? Do you announce: do you have negative symptoms and want to come in? Or do you actually get clinically-defined patients? I think that's crucial. Because I have patients call me, are you Dr. Correll? Yes. Do you have a study for me? Do you do clinical trials? Yes. And then they said, well, do you have a study for me? And I say, well, I don't know, tell me what's your diagnosis. And the answer was, what do you

need?

(Laughter.)

DR. CORRELL: And I say, wait a second. That's not how it goes. So what's your diagnosis? Schizophrenia? I say, well, I don't think this is going to work. So, I don't think so. And he said, wait a second, wait a second. Do you need a healthy control? I said, wait, what? You just said that you have schizophrenia. But I was once a healthy control.

So we have to be careful with announcing our studies. If I had said, oh, yeah, I have this and this and this studies, he would have said, yes, maybe I fit that one. We have to be, I think, more in the clinical validation of these patients.

Cultural differences will be addressed. Who wants to be part of the study? Illness stage. Do we want younger patients, early in the illness, less affected by dopamine blockade or the illness? Do we want more

generalizability? What are comorbidities that could also respond to the medication? What of substance use that is maybe ruled out at baseline, but then still has a background effect?

Prior treatment. Which antipsychotics? For how long? Will our augmentation studies be different once we have non-post-synaptic dopamine blockers approved and in the mix? Will that allow us to do different studies with different agents that were before dampened by dopamine blockade?

How much washout do you want? Can you washout? I mean, that's the monotherapy question that you will have. How do we enrich? Is it the severity? Is it the type of negative symptoms? Will it then have to depend on the mechanism of action of the drug?

How long for the stability or the trial duration be? Assessment, I think, will have to be dealt with in Section 4. But I guess we want the persistence. Maybe digital

markers are much more important than retrospective recall, how did I feel in the last month?

Pseudo-specificity, we talked about that. Also clinical stability. Additional outcomes. They can enrich and enhance trial burden, and maybe also the placebo effect. Because we have tender loving care. There is already an intervention as we get patients into the study. They come often out of rarified environments into environments that are much more enriched. And people want them to come back. Give them coffee and say, please, we like you, come here. And here you also have the reward of remuneration. And then we need, obviously, retention.

Data analysis. MMRM is the name of the game. We're basically biasing toward a completer sample, correct? Because we're imputing the outcome from the patients that are the super-responders and stay until the very end. It's maybe a little less severe as in the

total symptoms studies. So we need to look at missing, not at random, if that's the case.

So, here are some of the highlights for the augmentation. Does the augmentation have to be longer, because the effects are often smaller when you augment rather than monotherapy?

What are comorbidities? Can they dampen or enrich the effect if you have co-treatment of them? Do you want less dopamine blockade? What about washout? What about prominent negative symptoms. Do we need them in to be more generalizable?

Also, what about lead-in? Do we need a double- or triple-blind placebo lead-in? Variable add on beginning in order to not have people be so expecting of an effect. And then maybe exclude patients that have improved beforehand.

There has been a big effort before that already has, I think, guided us. We want severity to be defined. Steve Marder is the

first author a decade ago.

We want to exclude EPS. Antipsychotics, should first and second generation be there? Should we do stratification based on D2 blockade? A question. Should cognition be co-registered? Do we need, basically, a single global score? No, I think we also need to look at sub-scores.

And clozapine is a different kettle of fish because of the treatment resistance. Even in the augmentation meta-analysis it didn't work for clozapine.

This was not discussed in the meeting. But this was discussed. And the only lack of agreement, and we should focus on that, is predominant versus prominent.

Yes, we want to exclude depression, have a functional outcome, but it’s not required as co-primary.

Informants. It's nice to have.

Trial duration. Should there really be a difference between phase II and phase III?

Usually, that can make things a little complicated in the interpretation. We want clinical stability, and also prospectively assess that before the trial for negative symptoms.

So I would say we have problems with the augmentation data that have so far been the name of the game. Monotherapy is only recently coming in, because all of our dopamine blockers made these negative symptoms potentially worse.

Partial dopamine agonists might have some effect. But is that because when you augment you're bumping off the other dopamine antagonists from the receptor? And in a meta-analysis that was an effect, adding aripiprazole to dopamine blockers, even in high quality studies, not only in low quality studies.

So we need to define the underlying treatment very much. So the illness duration, comorbidities. There will be a lower effect size, most likely, than in monotherapy. What

about functional unblinding? Keep placebo effects in check.

And then, also, we talked about combination treatments. Well, should these drugs be used in order to boost? Or would the drug effect be more clear when you do a psychosocial intervention? Because otherwise you have people who are in rarified environments and can't even execute what's now better in their brains. But, on the other hand, you might enhance a placebo effect even more and wash out the drug difference. So we need to consider that.

And symptom enrichment and lead-in options, I think, should be given consideration, and, clearly, how to assess negative symptoms.

Thank you very much.

DR. BLACKMAN: Thank you so much for that wonderful talk. For our second speaker of this session we will be hearing from Dr. Stephen Brannan. Dr. Brannan is the former CMO

of Karuna Therapeutics and a neuroscience drug development expert who has held senior positions overseeing both clinical development and medical affairs for more than 15 years in industry. Dr. Brannan's experience includes drug development registration, medical affairs, launch and lifecycle management in areas of anxiety, depression, epilepsy, schizophrenia, the list continues. He trained in psychiatry at the University of Texas Health Science Center at San Antonio and holds an MD from the University of Texas Health Science Center at Dallas Southwestern Medical School.

(Off-microphone comment.)

DR. BLACKMAN: So, today he will be talking about the considerations for drugs being designed to be monotherapy. Dr. Brannan...

DR. BRANNAN: Thank you.

And disclosures.

So I'm going to talk about some general issues. Dr. Correll covered a lot of

this, and some of the previous speakers. He and I even talked a little bit so we could parse them out a little bit. So, some of these I'm not going to spend a lot of time talking about. The duration of the trial. Negative symptoms purportedly take longer to respond. There's some evidence for that, but it's certainly not universal. And what happens if you have new mechanisms? Is it the same? An issue: how stable do subjects need to be regarding the negative and positive symptoms? So, there's some arguments about that. Certainly, stable and persistent for the negative symptoms. How many recent hospitalizations or symptom changes. So, how OC do we want to be about all these things? There's an issue about relapses and rescues. I'll talk a little bit more about that. And then enrichment, what are the appropriate thresholds for negative symptoms and relative to positive symptoms? So, these are issues that I think that are not quite resolved and

probably deserve some discussion. Dr. Correll specifically mentioned, you know, some of the exclusions with depression or EPS. This also comes from ISCTM recommendations that have come out, like, over ten years ago, and then just a couple of years ago. So they've been very consistent. Where do study subjects come from? What is the role of the site and regional differences? And this becomes more important when you get to these large trials. Probably increases variance a great deal when you go too large, by the way. Do you also assess for cognition? Age range? So these are all things that are covered. I'm not going to dive into. I'll say a little bit about placebo issues, because we felt it was important for both of us to kind of chat a little bit.

So, let me go to enrichment and relapse issues. General consensus is that you need to exclude the actively psychotic individuals, such as those seen in most of the acute schizophrenia studies, and excluding

subjects who lack stable symptoms. Again, this has to do with reasons of variance. Now, there's a question. Should one exclude any subject scoring above a certain threshold for an individual item on the PANSS that's a positive symptom item? So this, I think, is, some people say yes, some people say I'm not sure. Study only stable patients with predominant negative symptoms. Again, how much do we go in this direction? But I think, generally, people are thinking this is the way we want to lean toward. In the studies particularly I'm going to be talking more about on monotherapy, are European sites' patients favored? And there's some reasons for this. In the U.S., a lot of the patients coming into the industry trials are not well-known to the PIs, whereas that is a little different in other parts of the world, such as Europe. And do you add a standardized, here I have vocational rehab and social opportunity, but Sophia and others were mentioning some of the

other things that one can do. And is this important to do either prior to the trial or throughout the trial? Would it be synergistic? As Dr. Correll mentioned just in parting there, there are pluses and minuses to this because this could also exacerbate placebo issues, which we know are a difficulty. Let me talk just a second about relapse. So, particularly if you're looking at, and I'll get more into this in a few slides, monotherapy and placebo, you need to have, really, even if you're doing it for all medications, what rules trigger when a subject should be withdrawn? How much fluctuation can one tolerate? So, there's always going to be fluctuation in the trials, especially across the entire population. So you need to start thinking about, okay, sometimes there's fluctuations. We can handle that. And the sites are pretty good at this for the most part. But, in the interest of safety, and even the integrity of the trial, when do you need to start withdrawing subjects?

And how do you set that up? One interesting issue and we'll talk more about this, like with the Minerva trial is, is there a subgroup of patients less prone to relapse? And can they be used in these types of trials? And who are they in the general population? What should be the role of a support network and informants? So these are also issues. When I do my trials, I consider myself very pragmatic. So, there's lots of wonderful things one can do, but you shouldn't do it in a trial when you're trying to kind of get stuff done. Keep it simple. And, of course, there's a role need for a DSMB safety board. And this is true for many, many trials. But what is that role here? Is it changing any for negative symptoms in particular?

I want to show you just one slide here. This is really from our ISE, not our, excuse me, ex-Karuna now BMS. And it has to do with the KarXT stuff. And it's only looking at the Marder negative symptom scale. On the

left-hand side, the full sample, you can see that it looks like there's a drop in negative symptoms, which is pretty common when you're treating the general study population for these acute trials. Then, in a sub-analysis that Bill Horan and others were doing, they looked for predominant negative symptoms, which is a little hard to do. So, it's a small subset within that whole group. And you see it's a little bit more back and forth on the placebo, although it ends up in about the same place. But, you see that group seems to respond even better. And this may just be, you know, you're really finding people with a lot of negative symptoms, so you're seeing more of a drop. But it's an interesting thought, if you're kind of trying to look and see how can we get subjects that are particularly good, or might be good to look at, if you think you might have something useful for negative symptoms.

So I'm going to primarily leave the outcome measure issues alone for Session I

think it's 3 or 4. I can't remember. So, there's a few things that are still important for the general study design. A relative consensus that functional co-primary is not needed. But should a global scale be included, such as CGI, specific for schizophrenia? And people have already touched on this. Does one look for negative symptoms as entire totality or dimensions? And which dimensions do you look at? I think there is a relative consensus building for depression and EPS scales, mainly to rule out confounds. And then there's the issue of accurate and stable ratings. I want to concentrate on accurate here, one of the bugaboos I have. Is there a need for informants? Do they actually help the accuracy? In the schizophrenia trials, that's not necessarily the case in the U.S. I'll just throw that out. The other thing I was very impressed at, in talking to a number of people while trying to put together this talk, is many of the scales for negative symptoms are not

easy to rate consistently. So, based on previous studies, there may be some things you want to do to try to shore that up, and also to know that you add multiple ones of these that can't be done well, that may not be a good idea, either. And then, finally, the use and role of non-rater measurements. So direct speech and facial measurement and the role for ecological momentary assessment, even actigraphy. Now, they're new, they're exciting, but I think, in the next five to ten years, they might be more important as people further refine some of these things. I'm just going to show you some very preliminary kind of stuff that I've seen that I think is sort of interesting.

One is, as one looks over the course of six months in just a non-controlled safety study, unproductive versus productive activities. I won't go into the details of how that's all determined. But one can see a very interesting and promising thing here, that it

looks like, over time, from the EMA-type information that one is getting, that both the unproductive activities drop and the productive activities increase during this safety study. So it looks like, there are things here that might be of interest in the future.

And then this is just very simple daily steps. And so, over the course of six months, again, we don't know how many steps are, important steps or anything like yet. That's a couple of years away I think. But you can see pretty clearly and easily that people are walking more. There's more steps being taken, more activity.

So, placebo issues. I am going to concentrate on the middle of the three, the need for rater surveillance. So, I just mentioned a few minutes ago that some of these ratings are not that easy to do, particularly consistency over the course of a trial. And so I'm a big believer that rater surveillance is probably needed. Why have I put it under

placebo? It's a lot of non-specific stuff. But, from my experience, this is one of the areas you really need to be careful about rater drift and other sorts of things. And the idea that things are being surveilled, I think, really helps. You sort of get this effect of being watched. And the sites and the raters, in particular, I think, behave differently over time if you have this. Particularly for negative symptom patients, the issue of staff attention is important. If you think of the nature of it, these are people who may not get a lot of attention. And so when you bring them into a place and they're getting a lot of attention, that has some non-specific effects that we also call placebo. And this final one and there hasn't been a lot of stuff published on this, but I think increasingly and I think there's going to be a session coming up at ISCTM in about a year, blinded data analytics. So, there's things that one can watch that can be important to, again, help people understand

how to stay...I'll just give a brief discussion of something that I got a call last night from somebody trying to put together a study. And they're like, well, you know, these rescue medications, you know, how much is too much? It had to do with benzodiazepines, I think, in mania. But, it depends on what you're doing, how long, and so on and so forth. But one of the things I said is, well, you should keep watch by site how this is used. Because I can tell you from our schizophrenia studies, some sites never gave a rescue medicine to anyone; other sites gave a rescue medicine to everyone. So, and there's everything in between. So there are things you can watch and talk to your sites during the trial, without perhaps being overly pushy about certain things. But, reminding folks. I did for one site we had, had everybody over the age of 55 for their first three people. And so I called up and I asked how long the geriatric convention was going on in town. And the next three subjects

were all under the age of 55.

Issues pertaining to monotherapy trials. I think a couple of things here. The comparator arm, are you doing placebo? Are you going to do that against other antipsychotics, in this case, other comparators? And then it also touched on the role of psychosocial interventions. So, you can do a number of things here. And then the idea would be, particularly with the monotherapy, and if you're doing it, is this going to be an additive thing if you're using this? Or is it going to be synergistic? Such that if you actually have something that helps negative symptoms and you have the training or augmentation that is non-pharmaceutical-related, maybe that's the best way to see, actually a signal. So I think all this is not well-determined. If you choose against placebo, there are safeguards. And I'll talk about this in the next slide. For longer durations, you just need to and even if it's

not against placebo, define symptom worsening well. You need to have an active safety board. And you will need study subjects who are non-relapsers, for lack of a better term. So, stable, no history of relapse, and probably they also have a good support network. And, again, the number of patients who do that and also are willing to come in to the trial may be small. Against the comparator, you want to have a drug with minimal EPS, or anything that would sort of confound while it was going on. Pragmatically, one cannot use all antipsychotics. So, which one, or ones, do you use are sort of issues. We may not decide today. But important one.

So, I'm going to start off here with a placebo-controlled trial. So, there's always a screening period and washout, prior to randomization. So this is important for all the things. Now, I don't have any open label extension on these, but you can clearly use that further on. And all these boxes are going

to be very similar. So it's really the stuff in this box here that we'll be talking about that are the differences. So, there is a notable relapse risk. I'll show you some of that. At least if you look at the general population. There's a high potential for symptom fluctuation, and it can complicate treatment benefit. And, of course, this is probably true for most of them. There can be a reluctance to participate, particularly if they know they're going to be on placebo for a long period of time. And this could extend your timelines to enroll.

Now, I'm going to talk a little bit about the Minerva trial in a second, about an example of this. Maybe not exhaustive. But, again, we do know that the patients on placebo had an increased risk for relapse. And then you can see again here, in some nice meta-analyses, that if you're looking at this, it does show that these people do relapse. So it's not that you can't find a subgroup. But,

overall, this appears to be the data the we know motivates and it's important for why we use antipsychotics in particular. There's ethical consideration really every way you do this. The Minerva trial, 12 weeks duration with a 24-month open label extension. Age ranges, there's 234 subjects, so pretty good. Negative symptom stable for three months. So, how long people need to be stable, people argue about. I found it interesting that BMI over 35 you were not allowed to be part of the trial. But I've done mainly U.S. trials and this is primarily European sites. So it probably all fits.

You can see actually that in the Figure 2 is really the overall PANSS total. And a low group and a high group. And there's some differentiation. And then, of course, for the PANSS negative subscale here on the right-hand side, you can also see a change starting two weeks or four weeks, depending on how you want to look at it, and then carrying on

through. So the other thing that this slide doesn't show is they did have a they had relatively few dropouts compared to what one typically sees in such trials. So it's proof that you actually can carry on such a trial.

Another one is, was running a superiority trial versus a comparator. And here, the comparator should be an approved treatment for positive symptoms that is neutral, which I believe is probably all of them. Some them may be a little negative. And then you do have some risk of functional unblinding with approved treatments, from sedation, weight gain, particularly if you have a long-term trial. Only a single drug comparison. If you have only one drug, it could lead to some label issues. But I'm not going to talk about that in front of the FDA. I'll leave that to you guys. And, again, large study meta-analyses suggest that most of the antipsychotics are equivalent in their lack of effects on negative symptoms, although not

necessarily in small trials, like Dr. Correll showed.

So the cariprazine trial is the one I'm going to sort of use in example here. The fundamental assumption is your comparator adequately controls for positive symptoms without impacting the negative symptoms. And in a positive trial, is it due to both positive effects or is the active comparator decreasing negative effects? There's some things there. But it's probably a very good way to do it. It may be more feasible, practical to run, than a trial just on placebo. And there's some arguments about, at least for the total population, if there's an ethical positive as well benefit.

Here, you can see the trial that cariprazine ran. And again it showed some separation. It was run a European sites again. So I think this is both with stability and some other things. Might be an advantage.

And it didn't appear...I probably

should have changed this slide, I just took theirs. It said no pseudo-specificity. It looks like there's not a lot here. So it does look like the comparator arm was able to work.

In the interest of time, I'm going to go through this a little quickly. This is more stay/switch design. So if you're focused on first at a single approved antipsychotic, or you could do it on multiple antipsychotics. So, half your patients would not be at risk to develop clinical instability on a new agent. They would increase your attrition at least for that group. But it probably benefits the stay group because they're used to it. You do increase the potential for unblinding. And when people are looking for new investigational products they're already working against enough problems, they're probably not wanting to benefit the stay group.

And then the second one is, you could also not just do it against one, but do it against standard of care. There is an issue

here that you would need to talk to our regulatory colleagues about, about sub-analyses of the different groups, because some of the groups would be much smaller than others.

So, in summary, there are important issues that relate to both the adjunctive and monotherapy trials. We've tried to highlight some of these considerations. For monotherapy in particular, the choice of comparison, I think, is a big issue that one needs to determine and then decide how you're going to deal with it. One can also envision the need or the use of psychosocial intervention within the trial, prior to randomization, perhaps as a run-in period or even throughout the trial to help augment or sharpen the differences. There's a variety of choices one can take depending on the specific aims of the trial in question and what the mechanism is of that particular NDA. Regardless of choice, concerns about practicality and the threat of high placebo responses should be important

considerations when you're designing your trials.

And I believe that's it.

DR. BLACKMAN: So now I'll ask the respondents to come up. And as they do I will just do brief introductions for them in no particular order.

I do have an unfortunate announcement, which is that Dr. Buchanan is not able to participate in the panel today. And he sends his apologies for that.

So, for the respondents we have Dr. Farchione, who is the Director of the Division of Psychiatry here at FDA. We have Dr. Yang. She's a supervisory mathematical statistician for the Division of Biometrics I here at FDA. Dr. Michael Sand currently serves as a consultant to a number of pharmaceutical companies and the National Institute of Mental Health. Dr. Richard Keefe is a professor emeritus of psychiatry and psychology and neuroscience at Duke University Medical Center.

And Dr. Nina Schooler, who's a professor of psychiatry and behavioral sciences at State University of New York Downstate Health Science Center.

So now I'll just give each of the panelists one to two minutes to respond to the talks today. So maybe we could just go down the line and start with Dr. Farchione.

DR. FARCHIONE: So I'll just make a couple of quick comments, one general thing and two more specific things. Because I know some, I'm assuming some, of the questions will probably be more pointed. But, the general thing is that it's clear to me that despite my crusade over the last several years to try to banish the word pseudo-specificity from our collective vocabulary, I have failed on that point. I'll let it slide for today. But, you know, two quick things that sort of stood out to me in the presentations. One in Christoph's talk about the phase II studies and how hard it is to replicate things in phase III. And one

thing that we see a lot is when Sponsors come in to us and have these meetings and ask us questions, and they'll try to do things in phase II like enriching for younger subjects or fewer prior treatments and things like that to try to enhance signal detection. And it strikes me and we're always like, oh okay, yeah, you know, if you're really, you know, you're trying to identify a drug that might potentially work. But maybe, that's not the best strategy in the end because maybe what you've done is you've detected a signal in that smaller subset of the population and have actually now made it harder to win in phase III where we expect the results to be more generalizable. So, you know, it worries me that now maybe some of the advice we've been giving where we said, yeah, you know, that's fine, go ahead and do that, maybe that wasn't great advice. And I think that will be useful for folks to think about and talk about with the panel. Because now I'm concerned that

we've been giving bad advice.

But then, Steve, in your talk, your talk was supposed to be about monotherapy. And it struck me that when you described the impact of potentially adding on social and cognitive training and things like that, that you described it as potentially enhancing the placebo effect. But I want to reframe it, because then it's actually you kind of strayed from the script a bit and we're talking about adjunctive therapy instead of just monotherapy in your talk. But I think it's important when we're talking about adding a standardized therapy, one lesson that we have come across very recently is that that standardized therapy should also be standard of care and evidence-based, so that we have some frame of reference when it comes to labeling and things like that. And in that case, when you're designing your study, you need to account for the fact that you expect some improvement in the placebo group because of that intervention. And so

maybe that requires a larger study to be able to detect what will ultimately be, unless it's really working synergistically, and then you actually would see greater separation. That's what you would hope for. I mean, that's kind of the whole point in adding it on there, right? So, I don't know. Those were the two sort of specific things that I had to comment on. For the sake of time I can pass it over to Nina, who always has really insightful comments.

DR. SCHOOLER: Well, we shall see what we shall see. So, the thing that struck me most strongly was that both Christoph and Steve raised more questions than they answered. And I was hoping for more answers. So, given that, it feels only fair to raise further questions.

(Laughter.)

DR. SCHOOLER: And the issue for me that's really paramount in terms of all of our studies of negative symptoms is that one of the

definitions of being someone with negative symptoms is lack of willingness to do things. And participating in a clinical trial involves doing things. We heard a fair amount of discussion about the issue of once you're in the clinical trial it can be very rewarding. But the fact of the matter is that these people have a lot of trouble coming in to a trial. And my experience in consenting patients, and I regard myself as very good at this because I'm really positive and energetic, has been that there are some negative symptom patients who I dream of who would not enter the trial under any circumstance whatsoever. And so I think one of the really important things is going to be to try to design strategies that work for the consent process, as opposed to, and then we can go on to what we have later. But having a broader population of patients with true negative symptoms may be very important. And I'll just give an example of something that can be used in that regard. The first is the use

of some kind of motivational interviewing strategies, which have been shown to be really important in a lot of contexts, as part of an informed consent procedure, which can perhaps enhance the breadth of the patient population that enters in. And the second is that there was really interesting discussion earlier this morning by Sophia which addressed the question of reward, and the idea of the kinds of rewards that you can pay people who participate in trials can be done in different ways. An immediate reward is always going to be better than delayed reward. So the kind of reinforcement that immediate reward can provide may be very helpful to keeping people in studies as well.

And I'll stop there.

DR. KEEFE: Thanks, Nina. So I thought that...and first of all, I'm not exactly sure what the product of this meeting is going to be and I'm very interested in seeing what that will be. I felt as though

both of you raised all of the important points about the methodology of various types of negative symptom trials. My concern, having lived through, day-to-day, the aftermath of the MATRICS consensus process which took place about 25 years ago, we still don't have a drug for cognition and schizophrenia. And that's not the MATRICS process's fault. I think it stimulated drug development. It got a lot of companies trying to do something about this. And I think that the MATRICS recommendations were, in many ways, spectacular about laying down the ways that the proper methodology for clinical trials. However, as everybody knows, I think that industry folks take these recommendations so rigidly and they adhere to every single word. And so what has happened, I think, through the course of that process is that innovation has been stifled. And so I think what we don't want from these recommendations is to leave with a set of incredibly rigid recommendations that don't

allow people to take advantage of new technologies, and I hope Greg Strauss and others this afternoon are going to talk about digital assessments for negative symptoms, because my reading of this is they are just so much superior. And we don't want it to be that the only way we can innovate in this space is when all of us up here on the panel have retired or are dead.

Thank you.

DR. SAND: So as one of the former rigid industry people, I'd like to make just a couple of comments. Other than tremendous presentations, and you really did a yeoman's job, both of you, of identifying the issues, one point I would raise is we haven't talked about, and I think we need to hold our FDA colleagues' feet to the fire on this, is how some of these things are going to affect labelling. Because, at the end of the day, we're in this to commercialize a compound or a device. There are now devices being

investigated for negative symptom treatment. But we have to bear in mind that every one of these things that we put into our trials may have a consequence to the label. And that's something that needs to be thought of, and the academicians don't think that way, but we, as drug developers, must. We have to be aware of what does this mean for our label. So if we're looking for the right left-handed, six-foot, redheaded, blue-eyed people, that's great because that's where we have our greatest effect, but unfortunately then our label looks like that and we don't have a drug that we can commercialize. That's one point.

Second point is I started my career in human sexuality and I was telling Rich when I walked in here to this building this morning, I was having PTSD from FDA interactions over that over many years, trying to explain female sexual desire to male urologists here was an interesting experience. For some reason, a centrally acting drug for a DSM condition was

being reviewed by urology. So that was fun. But one of the things we talked about and needed to make clear to people is when you look at the phenomenon of desire, what's the right level of desire? Well of course, it's very individual. What might be an appropriate level for myself might be completely different for somebody else. Can someone tell me, what is the right level of avolition? Can somebody tell me, what is the right level of social engagement? I have some in-laws I would hope became less socially engaged. So I think we need to be aware at the end of the day that this is a very personal thing. And I think for all of our brilliant work that so many people are doing to parse out the neurobiology of this...Brandon started out by mentioning the word thrive. Okay, and there's a big difference between an improvement on a scale and thriving. And so I don't think we should lose track of what is this meaning to an individual person in terms of feeling better

and doing better about in their own lives. Because it is very individual, and what might be an important thing on a mean basis might not be to an individual. And, I think that's worth thinking about.

And the last thing I want to toss out, which was not mentioned, is adherence in clinical trials. We know that pill counts are worse than useless, but that's sort of the standard. I conducted a trial and in my last role looking at relapse in preventing relapse in schizophrenia, and used a digital app to assess relapse over the six months. And we found that only about 50 percent of people took at least 80 percent of their drugs. Okay. And in overall trial failed, but in that 50 percent who reached at least 80 percent, there was a 90 percent reduction in relapse over six months. I remember Bob Temple, being at a meeting where Bob said "I don't care what happens to people who don't take your drugs. I care what happens to people who do take your drugs." And I think

there's a little bit of an elephant in the room when it comes to schizophrenia trials that we don't really know very well what level of drugs people are taking. ISCTM looked at this, AstraZeneca gave the working group data from five failed phase II studies, where they had PK data, and 20 percent of people in those trials had zero detectable drug levels when they were assessed for PK. So if we have almost any signal of effect in the negative symptom trial, it probably means we're onto something good, because a whole lot of those people aren't even taking the drug. And I think we need to think about how we can improve on that in trials.

Thank you.

DR. YANG: Hello. I am a statistician, so I am going to comment from a statistical perspective. Speaking of the study design, there's a question about what factors could affect the sample size calculation when designing augmentations and monotherapy studies. The sample size calculation many

depends on the distribution of the efficacy outcomes in each treatment group. And we often calculate sample size based on the assumptions of treatment effect and the standard deviation of the distribution. And the treatment effect can be affected by patient population, the countries where the study is conducted, placebo response, and other factors that are suggested in Dr. Correll's and Dr. Brannan's talks. In addition, the sample size calculation is also affected by the dropouts. And we know that in the augmentation studies, the dropout rates tend to be lower than the monotherapy studies. In recent years, the implementation of the estimand framework has somewhat affected the statistical analysis of efficacy outcomes. The estimand framework consists of several attributes. One of them is the intercurrent event, which is an event that occurs after treatment starts and may affect the interpretation or existence of the outcome data. Examples include treatment

discontinuation and changing the background therapy in the augmentation studies. And so we have been asking Sponsors to include the estimand framework in the protocol, particularly listing intercurrent events with the strategies to handle each of the intercurrent events. These strategies essentially deal with observations, whether observed or missing after the intercurrent events. As a result, these strategies may affect the treatment effect depending on how you handle the missing data or the data even though observed as intercurrent events. So I think Sponsors may consider these factors in their sample size calculations now. Having said that, we need to have good data for this exercise. And especially for endpoint that we have less experience with. And I second what Dr. Farchione just said in the beginning. And the phase II studies, sometimes we see positive results in phase II studies, but a lot of times, my impression is that the result of the

phase III studies are not as good as in the phase II studies. And that happens quite frequently. So if you intend to use the results from the phase II studies to design the phase III studies, you want to keep in mind of the variation. Just for example, when you underestimate the treatment effect based on the phase II studies, you want to allow for more variability. So you want to make sure you have really sufficient study power for phase III studies.

Thank you.

DR. BLACKMAN: Thank you. So now we'll open it up to Q&A before we take any questions. So if people have questions in the room, feel free to come up to the mic. Just in the interest of time, I had just spoken in random order about our presenters, but if anyone needs it, from the podium over is Dr. Correll, Dr. Brannan, Dr. Farchione, Dr. Schooler, Dr. Keefe, Dr. Sand, and Dr. Yang. That's the order people are sitting in today.

Similarly, I will ask, if you have a question, if you could introduce yourself and then ask your question. Thank you.

MR. MARTIN: Yes. Hello. Steve Martin from UCLA. This is taking off from Tiffany Farchione's talk. I thought that one of the profound things about Brandon's talk was that when he, and I think it's correct, that although you had negative symptoms, you felt a loss and a desire to address that. And what I think one of the problems in negative symptom trials is that many patients don't have that. They don't experience suffering. And I wonder if the people who are indifferent to improvement, whether or not we should be studying them in these trials. Because you can give people dollars in order to participate, but if they don't have a motivation to get better and to reengage in life, as many patients don't, maybe they should be excluded from these trials, particularly early on, when we're looking at signal detection.

DR. FARCHIONE: I mean, I don't know if that was so much a question as a comment, but I think it also piggybacks on what Nina was saying in terms of, you may not need to exclude them if you aren't able to recruit them anyway; so.

MS. PANI: Luca Pani, University of Miami. Steve, probably I'll just throw this past you, and I'm not a statistician, but when you did the practical analysis and the prominent negative symptoms, the number of patients went from 300 to about 30, something like that, so it's only 10 percent, but the impact was higher. My question is why did the two arms became unbalanced? You had prominent symptoms. You had 31 in one and 22 in the other. Is anybody or maybe Dr. Yang, offer me an explanation of why by doing this you imbalance completely the two arms?

DR. BRANNAN: I don't know that we can answer that. This is all retrospective, of course. There may be something there. Again,

what I've learned over time is with relatively small numbers, to be somewhat tolerant of the facts of it's not as equally balanced as I would have liked, but it's not too unreasonable...

MS. PANI: It's like we're losing the randomization principle somehow. We should...

DR. BRANNAN: Yeah. The comment Dr. Correll was saying, it's not stratified. So we didn't stratify subjects...

MS. PANI: I see.

DR. BRANNAN: ...on predominant, we weren't even thinking about it at the time. We were just doing schizophrenia. So that probably is a good, they weren't stratified, so they weren't necessarily equal numbers to begin with, and then all sorts of things.

MS. PANI: So it means the other way around. So you cannot infer from some of those analysis to the generalizability of the general population, or no?

DR. BRANNAN: Well, I think if you're looking at that data, you have to be careful because of all the things that you're bringing up. Although to my eye, it looked like you saw a much stronger signal...

MS. PANI: Oh, no doubt. Yes.

DR. BRANNAN: ...in what already is a strong signal.

MS. PANI: Right.

DR. BRANNAN: So if it was really, again, this gets back a little bit, I think, to what Dr. Correll was saying earlier. When you have really small things, you have to be really careful. And Michael probably will agree with me that industry people aren't always careful. But especially going phase II to phase III. But I'm not, by the way, don't get too worried about, letting people see signal early on, and I'll get back to that.

MR. STAGLIN: Hi. Brandon Staglin. I spoke earlier, One Mind. So questions for, I think, two of you up there. First is for Dr.

Sand, partially a comment, partially a question. I applaud how you called out the need to understand what are meaningful outcomes for people who are taking these treatments. What's the right level of sociology? What's the right level of motivation? It's very personal, as you say. And then also about the adherence question to using a drug. Both these can be addressed by talking with people with lived experience directly, obviously. And so, I know that many pharma companies include patient advocacy and patient consultation functions within what they do. But I just want to advocate for more of that, and to ask how can pharma companies and other research entities take advantage of the fact that there are many people with lived experience who would like to help develop new treatments for people like us, like them, and contribute to the process of making lives better for our population, in consultation with treatment developers? So, it may not be something you

can answer right now, but how can we make sure that happens on a larger scale?

One quick comment. One Mind has a lived experience council, which we just formed this year. It currently involves four individuals, and we're seeking to grow it. But I'd like to see more of these councils grow, and the applicability of them grow, and the trust in them grow. So that's the first question. Any thoughts on that, Michael or anyone?

DR. SAND: So I'll comment, having utilized these for years.

MR. STAGLIN: Mm-hmm.

DR. SAND: I have found it invaluable in our clinical trials to involve organizations such as yourselves or NAMI. Invaluable to have worked with them and have both patient and caregiver panels. We have them review our protocols. We have them review our patient-facing materials. And all of it has a language. I mean, there's so many things

that people with lived experience and their caregivers can provide us as drug developers that's valuable. I think anyone who isn't doing that, because again, as you well know, I mean, these organizations are more than happy to help. And I think everyone doing this kind of work should consider, engaging with them.

MR. STAGLIN: Thank you.

DR. FARCHIONE: And if I can just jump in as well, I know I'm going to call out Steve again, but, you mentioned at one point in your talk that there's a relative consensus about maybe not needing a functional outcome. And, you know, maybe.

DR. BRANNAN: Co-primary.

DR. FARCHIONE: Well, and that is a co-primary. Yeah, that's true. But I think that we do, to your point, we still need to focus on clinically meaningful change. Like what matters to the person sitting in front of you, you know? If you're not making a change that matters to the person who's being treated,

then what's the point of what you're doing? And I think that, if we have a session on clinical outcome assessments, and I think that'll probably come up during that session pretty extensively because one of the things that we really want to do is actually talk to patients and find out what matters to them, how much change. And if we're developing endpoints, are you measuring the things that matter? If you're not measuring what matters, you're not going to see an outcome, even if one exists, right? So...

MR. STAGLIN: Very true. And I think that can be a key to improving the adherence to these medications or other treatments. Because if normally, if they meet the needs and interests of people who are using them, then people are more likely to use them, right? And also, simply knowing that they are co-designed with the lived experience community is a way to kind of get in the door with people who might consider taking them, because they

know they their interests are being represented.

DR. KEEFE: And I think, Brandon, this is for you as well as for the general group. But I think we need to make a distinction between insight about somebody's symptoms and insight about somebody's functioning, and just how they're doing, and whether they're satisfied with their treatment. And the example I'm thinking of is I just started wearing this WHOOP band to sleep, right? And so I'll get up and think oh, it's a pretty good night's sleep I had. And my digital data said uh-huh, no you didn't. You didn't sleep well at all. And so, you know, here are the data. And actually, the variability in the WHOOP is significant. The variability in my perception of my sleep is significant. And they actually don't correlate very well. So my insight...

DR. FARCHIONE: That's not that unusual for sleep measures.

DR. KEEFE: Yeah, I know. I know. But my insight into my symptoms is very minimal. However, I think my insight into how my day went the next day, whether I was functioning, whether I was able to pay attention and so forth, those things do matter. And so I think it's insight to symptoms versus insight to what your functioning is like, and whether it's satisfying to you, those are the things that probably really matter to you, not insight about symptoms.

DR. SCHOOLER: One more comment. So this actually gets us back to the point that Steve Marder raised about who should be in these trials, and should we not be including people who don't see the problem? And what I would argue is that in many schizophrenia studies, we include many people who don't see what we see as symptoms worthy of treatment as symptoms. In other words, somebody who's genuinely having a paranoid experience doesn't think of themselves as having a paranoid

experience. What they're thinking is the FBI is after me, or whatever the experience is. And so I feel that to be a kind of tricky thing to think about in terms of who should be in studies. But I clearly understand the problem in getting people to enter into anything that they don't think benefits them in some way. And so the issue is, what is the hook that you can use that's a legitimate hook that will still engage people who don't necessarily consider what your target of interest is to be their target of interest?

DR. BLACKMAN: Thank you. Apologies. I think we probably only have time for maybe one more question. So I'll take the next question in the room here.

PARTICIPANT: I was just going to make a quick comment sort of looping all these things together, because the intersection between what Steve said and what Brandon said in his talk suggests that there's not only awareness of the presence of your negative

symptoms, but there has to be a motivation to overcome them, which also would imply sensitivity to treatment effects. And we know a tremendous amount about lack of awareness of cognitive deficits, functional deficits, and things like that. But we don't know much of anything about how people with schizophrenia who benefit from treatment, such as cognitive remediation, are aware of the extent to which they improve. And Rich published a paper that involved a successful digital device intervention for major depression, and the people in the active treatment arm who objectively improved in their cognitive functioning didn't report that they were functioning any differently than the people who didn't improve. And so there was this disconnect between objectively measurable improvement on neuropsych tests and the group as a whole's ability to report that that had any impact on them. They found the same thing in a TBI study that was done by Henry Mahncke

with BrainHQ. So there's a number of streams suggesting not only unawareness on the front side may be important, but the ability to perceive a gain. How are you going to deploy your new skills if you can't perceive them? And I think the more motivated someone is to benefit from treatment, that may carry through. Can we expect people to develop that over the course of treatment if they're actually getting better? In our view, right? As Nina defined carefully, that we're treating a lot of people who don't think they have symptoms, but we know how to measure what they don't think they have, and we're not asking them to translate that into functioning.

DR. BLACKMAN: Thank you for your comment. For just the last couple minutes, I just did want to reserve some time, since we have Dr. Farchione on the stage, and I'm sure a lot of people have this question in their minds, how can they best engage with us here at the FDA in terms of the regulatory process and

the drug development process?

DR. FARCHIONE: I think that everybody that works in drug development knows about our meeting request and everything like that. So the idea that, you know, you should talk to us early and often, I'm sure you've all heard me say that 65 million times, so I probably don't need to say it again, although I did just say it again. But I think, ultimately, what we've heard so far today, and what I'm sure we'll continue to hear is that, addressing negative symptoms is very complex. And it isn't a matter of, I don't think that there's a single approach that encompasses everything that we're talking about. So ultimately, the way that you design your program really needs to be hypothesis-driven in terms of, what aspect of negative symptoms are you trying to address, are you trying to go from monotherapy, adjunctive therapy? You know, what is your treatment paradigm going to be? All of those things are going to drive the

type of study designs that are going to be needed in order to demonstrate an effect, and also to help us to inform labelling. Because I think, you've already mentioned that at the end of the day, we have to be able to write a label that tells people: Who the patients are who are going to benefit from this? How do you use it? What happens if the drug doesn't work; do you stop it, do you not? All of those things. We need to be able to write an informative label, so that people can not only use the drug, but use it safely. So depending on what your hypothesis is, what your proposed treatment paradigm is, what your proposed population is, that's going to affect how you would design your study. And those are the kinds of things that are worth coming and talking to us about.

The other thing that I would say is that, again, going back to this idea that it's so hard to see a positive result in phase III, even if you've won in phase II. What I think we see a lot is that folks come in and they've

got this like squeaking by p-value and, like you mentioned, they're just barely getting over the line there. But hey, we saw something, so we're going to go after this. I think what Christoph's talk really showed us is that this idea that, like, okay, we won in phase II, let's just do phase III exactly like we did phase II, and we're going to win again, isn't the best strategy. I think looking at why your phase II study won, who the patients are who benefited, and then kind of trying to refine your design. We always talk about how you need to, phase II is exploratory, you need to take what you learn from phase II and apply that to your design in phase III. But a lot of what we see is just let's do the same thing again. And that may not be the best strategy. And in fact, it sounds like it's actually not the best strategy when it comes to negative symptoms. But that is another place where it's going to be really important to engage with FDA. Because now, if you're going to say, "Like,

look, it looks like we have this specific population that really benefited," we're going to maybe say, "Okay, well, that's interesting, but now you have to replicate that." You have to demonstrate it. Maybe you might want to do another phase to study, to explore that hypothesis, to make sure that it's a real thing before you jump into these massive studies and invest all of this money.

So those are the main points I think I would make. And again, talk to us early, talk to us often. Anything you want to do that's a little nontraditional, it's even more important to talk to us even more often. If you've got questions, we don't really turn down a whole lot of meeting requests, unless we look at it and we say "Oh, that's all review issues." Just send us your protocol. We can't answer it until we see your protocol then sometimes we turn down meeting requests. But mostly, we just want to engage. So that's it.

DR. BRANNAN: If I might just add on

something, one of the things that I think you do that's very helpful is going to meetings, like ISCTM, even ACNP and stuff like that. Because when the FDA's out there you can...

DR. FARCHIONE: We try to go to ACNP.

DR. BRANNAN: Well, yes. There's a lot of reasons why you can't go to certain meetings. I'm somewhat aware of some of that. But for, in case your bosses are listening or whatever, it's important. And that engagement, the willingness to engage, I think, is really useful, and it's been very helpful to a number of us.

DR. BLACKMAN: And Dr. Sand, if you want to...

DR. SAND: Just a provocative question, Tiffany.

DR. FARCHIONE: I would expect nothing less.

DR. SAND: Good. If someone is being seen for predominantly negative symptoms

and they're prescribed olanzapine, do you consider that being prescribed off label? Because I don't think that these D2 blockers have specific indication labels for positive symptoms only. They simply say treatment of schizophrenia. Why, if negative symptoms are a core symptom like the positive symptoms, why wouldn't or should we have or be seeking, as developers, a label for negative symptoms, or should we simply say we're looking for a treatment of schizophrenia?

DR. FARCHIONE: Well, but we already have all of these things that are approved for the treatment of schizophrenia. And there's a general consensus that they just don't help that much with the negative symptoms. But the way that the studies are done, you've got these scales that measure everything. And they do move negative symptoms, they just don't move it enough for it to seem to matter, for people to seem to notice. So...

DR. SAND: So if you turn that

around and say well, what if I'm developing a compound that I think might have effects on cognition, negative symptoms, and positive symptoms? GPR 52, for example, that has at least a biologic reason to think that it could be plausible. Would I need to do phase III trials in each of those domains, or would I simply do take all comers and get a label? You know, you see what I'm getting at.

DR. FARCHIONE: Yeah.

DR. SAND: It's hard to...

DR. FARCHIONE: Yeah. I mean, and I don't want to get into, like, too many specifics because, you know, we're not having a Sponsor meeting up here at the table. But, the idea that there are areas in schizophrenia that we recognize where there's an unmet need, and even if you were to do, if you had a development program where you were able to improve across the spectrum of symptoms, I think it speaks a bit to the limitation of the current labeling, where everything is for

treatment of schizophrenia, and where we have to admit not all symptoms are treated all that well. We would have to come up with a way to still have that broad indication, which is now rightfully earned in a case where you've improved all of those domains, but to also have a description of that improvement in, for instance, section 14 in the clinical studies session that actually addressed it. But I don't think in that case that there would be individual indications.

DR. SAND: Yeah. That's why I...

DR. FARCHIONE: Yeah.

DR. SAND: ...wouldn't think so, but...

DR. FARCHIONE: Yeah.

DR. SAND: I just wanted to know.

DR. FARCHIONE: Yeah.

DR. BLACKMAN: Okay. Thank you. We are going to take maybe just one question from the virtual audience. And if it's quick, maybe we'll be able to get two, and then we'll just

adjust lunch accordingly.

DR. CAMPBELL: I'm going to try to combine a couple questions...

DR. BLACKMAN: Perfect.

DR. CAMPBELL: ...into one question, because we have a lot of questions. I want to thank the virtual audience for being patient with asking their questions.

And so this is going to probably start with Tiffany and Peiling. And while we will be having an entire afternoon session on the outcomes as part of when we think of study design, we have to think about our endpoints and our endpoint hierarchies. And so as we continue our discussions today about trial designs that would work, what are you thinking about the hierarchies in terms of a primary endpoint versus co-primary, or how additional support of secondaries to really inform that overall lived experience? So any thoughts? I want to start with my DP colleagues first. But then if anyone else has a question or a

response about that.

DR. FARCHIONE: I'm not exactly sure what we would say about that. I mean, you know, in terms of, like, how that would be approached in labeling, or how that would be approached in...I don't know. I guess, I don't know. Christoph, you've got your hand up. Maybe I can piggyback off of whatever brilliant thing you're planning to say.

DR. CORRELL: Yeah. I mean, I think the question is whether you go for total negative symptoms or subtypes of it, and where you start. So if you think your drug has a particular affect in an area, and you go for that first, and you know, okay, this might just be subpopulation, but that's my safe bet, and I've enriched for that. And then if I have a positive effect, I do hierarchical testing. But maybe I can get, now, the whole negative symptoms and broaden my indication. So that would be one way of going at it. Or if there's lived experience, you want also, satisfaction

or personal recovery aspect in there as a different lower hierarchical outcome.

DR. FARCHIONE: All right. I see where that's going now. I mean, it strikes me that that kind of an approach might make more sense, like, in phase II, when you're trying to figure out what your drug does and who it works for, maybe. Because then, if you think it's going to work in the subpopulation, you go for that, because that's your win. And then, you know, you start looking more broadly, more broadly.

DR. CORRELL: Right. That's the population. But if you see in this population a particular effect on a subtype of negative symptoms, and that would again be the label, but you could also then, in a second shot at the goal, say well, but maybe we get, actually it's broad enough that we'll also catch other aspects of negative symptoms. That would be on the outcome. But I think group four will go much more into this, I'm sure.

DR. FARCHIONE: Yeah. But then if you don't win on that second shot...

DR. CORRELL: Then you're just one outcome. That's fine.

DR. FARCHIONE: Yeah.

DR. CORRELL: But you have a second shot on a broader goal.

DR. FARCHIONE: Yeah.

DR. CORRELL: Not for the population, but for the outcome.

DR. FARCHIONE: Yeah. Yeah.

DR. BLACKMAN: Was there another question?

DR. CAMPBELL: Yeah. I want to ask, this more a question related to safety. So we've been talking a lot about efficacy during the study design, but obviously safety is a critical thing as well. And so, in the context of requiring patients to have prominent or predominant or stable negative symptoms, is there a risk of relapse? Negative symptoms do not generally lead to hospitalization. But if

stable positive symptoms or predominant, would relapse of a positive be a better measure for safety? And my assumption, this is in a trial of focusing solely on negative symptoms, that's my interpretation of this question, that last piece.

DR. FARCHIONE: Yeah. I mean, that's the elephant in the room, right? I mean, we're always worried about what happens with positive symptoms. And I think in terms of you were talking about monotherapy designs, if you wanted to...

DR. BRANNAN: Yeah. And of course you need to know what the medication is and what it's like. But if you're just talking about relapses, which is the, I think it's the main concern. So we know that there's positive symptoms. We know people off medication in general tend to relapse more if they're on nothing, rather than medication. But there are subgroups, and so it sort of depends on how you're running it and what the subgroups are.

I don't know if that's what they're looking for.

DR. FARCHIONE: But with those subgroups, how do you identify them a priori to say that this is somebody who isn't going to, I mean, because you need the positive symptoms in order to even get the diagnosis in the first place.

DR. SAND: So...

DR. CORRELL: No, go ahead.

DR. SAND: No. I was just going to make the observation that anyone, I mean, if you were concerned about that, you know, a DSMB would be a way of easily handling that. And I think in any case, anytime you had an actual relapse happening where someone was worsening and hospitalized, that would all be captured. So, it's not like it would escape, even if you weren't focusing on it in your trial, it's all being extensively documented and looked at. So I don't think it's a concern, per se.

DR. CORRELL: I think similarly,

exacerbation of symptoms, even below hospitalization or relapse, is coded as a side effect. But it's obviously possible that even if you have the underlying medication on board, that you have a stimulatory effect for the negative symptoms that could also stimulate positive symptoms, and we need to look at that. But going back to the adherence, so even if you have an augmentation trial, if people stop everything and worsen, well, we need to know whether that's the drug or they are not taking the baseline medications. So the question is how much does PK inference our reporting of side effects attributed to the drug, and also the efficacy? Is there a possibility to say that we're doing a sensitivity analysis and exclude people at zero blood levels, even thought it might not be fully randomized anymore, but then you correct for this, whether there are any baseline differences? Do you penalize a drug for the 50 percent of patients, 20 percent that are not taking any medication?

Or could we rescue that trial by using PK to redefine the analysis sample?

DR. FARCHIONE: Yeah. But I think, maybe Peiling could even speak to this, too, but that's part of the estimand framework and how you handle those events and everything. And also, if they stop taking it, why did they stop taking it? Is the drug not tolerable? You know, the...

DR. CORRELL: Well, not only the drug that's experimental, but also the underlying medication.

DR. FARCHIONE: Yeah.

DR. CORRELL: But still, I understand why did they stop it? You could look at that, what the reasons are. Are there predictors of this? But are we doing the service and the patient right service by including people who can't improve because they're on nothing?

DR. YANG: Yeah. This is related to the intercurrent events. And we are nervous if

you want to exclude patients, from the analysis, because of the principle, the randomization principle requires for valuable statistical analysis is lost. But I think it's better to handle this with the intercurrent events framework. Yeah.

DR. BLACKMAN: Thank you. This is a great discussion perhaps we could continue over lunch. I just want to give a hand to our panelists and our speakers. This was a great session.

For anyone on campus here, the kiosk will be open during lunch. If you think you might need a snack later that you want to purchase, purchase it now, because I'm not sure what time it will close. We will reconvene at 12...

DR. CAMPBELL: 12:45.

DR. BLACKMAN: 12:45, promptly. So, quarter of 1:00. Thank you.

(Whereupon, the above-entitled matter went off the record at 12:09 p.m. and

resumed at 12:49 p.m.)

DR. CAMPBELL: I know we are bringing folks in from outside from lunch. I want to welcome everyone back to our afternoon sessions. And really, I think our goal of this panel is to keep everyone from that postprandial slump, post-lunch.

So hopefully, we'll be a very engaging panel discussion on something that is extremely important to us at FDA. It is literally my everyday work, which is a conversation on clinical meaningfulness.

We heard it being highlighted from our earlier sessions today, but this is, you know, our everyday world. And so those of you who come and engage with us on a regular basis, you're very much familiar with this conversation.

This afternoon, we're going to first start on the conversation in clinical meaningfulness, which will include a conversation on cultural adaptation, diversity,

and inclusion, and really understanding the broad spectrum of clinical meaningfulness in the diverse populations, as well as a follow-up conversation, and hopefully we will do a good job to prepare for a very dynamic conversation to end the day, on outcomes. So, how do we take everything we've discussed today and get into the "how do we actually measure this critically important information?"

I'm going to ask my presenters and panelists if they can come up and join us already on the stage because, as we've learned from this morning, we are having dynamic conversations. And I don't want to waste time with people walking. So if those folks will come on up. And I will just highlight how excited I am with this panel. I think you will gain a lot from this session, and really be reflective of why we spend so much time focusing on clinical meaningfulness. All right. We will have two presentations. The first presentation will be from Eric Jarvis from

McGill University in Canada. And then we will have a presentation from my colleague, Laura Swett, who is a reviewer in our division of Clinical Outcome Assessments.

I think it's important to note that any conversation with clinical meaningfulness, I think someone said it earlier today, we could have an entire day, weeks long conversation series on clinical meaningfulness.

And so, we will honestly only be scratching the tip of the surface when it comes to clinical meaningfulness. We will probably not be getting into the quantitative aspects, but really the overall importance of this. So, I'm going to turn it over to Eric for his presentation and get this conversation started.

DR. JARVIS: All right. Thank you, everyone. Thank you for inviting me to this symposium. I've been really interested in the topics. I'm not involved in clinical trials myself in my work, but most of what I do is consultation on cultural and social issues in

mental health, psychosis, schizophrenia. So people will ask me to be part of their grants, or to comment on their programs, or even clinical practices that they're doing. And we -- what we do in our service, the culture consultation service, is we try to identify and find solutions, undo blockages that may be happening because of a person's ethnic or racial or religious or linguistic backgrounds.

I'm never asked about my disclosures. I'm asked to give a positionality statement for most of my work, so this is what I do, here. And just to let you know what I do, the kinds of -- I'm an associate professor at McGill. I'm not a member of an ethnic or racial minority, but I am a member of a religious minority.

This presentation will -- it's a vast topic, culture, psychiatry, mental health. So I kind of brought down the discussion to these three main groups -it may not entirely - papers published in the United States, members

of African American communities, and people with schizophrenia.

We've already reviewed these negative symptoms of schizophrenia. I think you know the general categories. I'm just outlining them here so I can give some commentary on some of the problems I think that can arise in anyone -- whatever rating system you're using to number and categorize people.

So, just a general overview of the problem, negative symptoms of schizophrenia are not so well-studied in ethnic minorities -- not so well-studied in anybody, I guess, for that matter, maybe partly due to beliefs that negative symptoms are kind of brain problems or biologically driven problems, less than sort of social or culturally influenced behaviors, maybe even less than positive symptoms, where there's quite a good literature on social cultural inputs into positive symptoms of psychosis.

We've already talked about this

quite a bit, negative symptoms can be misunderstood as -- can be falsely diagnosed instead of other conditions. We've mentioned many here, depression, PTSD.

Part of the problem for African American communities in North America is our historical stereotypes; the stereotypes that "psychosis, schizophrenia are linked to people of African origins," that "people in the past have deemed people of African origin to be of less intellectual capacity," that "they may be more prone, for these reasons, to have or be deemed to have negative symptoms of schizophrenia, along with members of other cultural minorities."

I just want to read this -- I'll make sure I have the right one -- this is from the APA guidelines from 2004 and were republished in 2010. I'll just read this out to you; "Compared with Caucasians, African Americans, especially men, are less likely to receive a diagnosis of a mood disorder and more

likely to receive a diagnosis of schizophrenia. African Americans with schizophrenia are also less likely to receive a diagnosis of a comorbid affective or anxiety disorder. While it is possible that such differences may reflect actual illness variation among racial ethnic groups, there is growing evidence that cultural differences in symptom and personal presentation, help seeking, interpretation of symptoms, and clinical judgments by usually Caucasian clinicians and treatment referral are likely causing race linked biases in diagnosis, and therefore in treatment." That's sort of a traditional position in American psychiatry.

Now, Gara and his colleagues published a paper in 2012 kind of confirming these impressions, and another one in 2019 -the references are at the end of the presentation, I'd encourage you to look through those if you're interested - so these -- so excessive diagnoses of schizophrenia and psychosis may predispose African American communities,

members of other communities, to be seen as having more negative symptoms if they have kind of doubtful presentations. Cultural mistrust or paranoia, the way we kind of coined these terms about 20 years ago, and they basically talk about healthy suspicion or healthy reticence when interacting with the institutions of society, like, for example, the legal system or the psychiatric mental healthcare system. So a person may come into an evaluation or through the emergency department, and they may appear to be withdrawn or not very responsive or not very engaged for good reason, but maybe misdiagnosed as having schizophrenia with prominent negative symptoms, for example.

We know about the literature from Western and Northern Europe about people of migrant backgrounds, especially from the Caribbean, from West Africa, as having high rates of diagnosed schizophrenia.

And in the United States, wherein in the past the high rates of schizophrenia in

African American communities have been seen as being artifacts of misdiagnosis, like I've been just talking about, there's a new kind of discourse emerging that there may be high rates of schizophrenia in African American communities because of systemic and structural racism.

So, the negative symptom literature in ethno-racial communities is a bit mixed. Generally, I would say that in African Americans and Mexican Americans, people tend to report higher rates of negative symptoms in these groups. Chinese Americans, the literature shows -- the reported literature shows that maybe there is fewer symptoms overall in schizophrenia and psychosis. Native Americans are less studied generally, but, as I say, the findings are mixed, and it's hard to draw conclusions at this stage.

Let's talk a bit about general and more specific rating problems. The first issue is the nature of psychotic symptoms. So with

psychotic symptoms, as most -- as you all know, don't arise purely from brain processes, and then they're sort of displayed to the clinicians, or to the research raters. They're kind of experiences that are filtered and shaped by the context, by the surround. So because of that, it's maybe, under some circumstances, easy to misunderstand what's going on for -- in certain -- in some settings and in some specific populations.

The broad ethno-racial categories that we all use, White, Black, Hispanic, Asian, and so forth, really should probably be discarded, because they're not super helpful in a cultural psychiatry context, and they don't really tell us a lot about the people that we're seeing beyond very superficial physical characteristics. But most studies will fall back and use these categories almost unquestioningly.

And then there's always the issue of the lack of culturally-adapted instruments and

surveys. And I know many people want to adapt instruments and surveys that they're using, but in practice, it can be kind of complicated and expensive.

I'm just going to go through a few of the negative symptoms that we see and raise a few issues that could arise. So the first one has to do with negative or blunted -- or, sorry, flat or blunted affect, unchanging facial expression, poor eye contact.

I think Bernie mentioned earlier today that these kinds of symptoms have to be taken carefully in the context of the person before you. So when a woman from -- a refugee woman doesn't look at me in my evaluation, I know it's not a negative symptom, I know it's out of deference to me as a male -- older male clinician.

Alogia, like, poverty of speech, poverty of content of speech. So if somebody comes from a socially difficult background, maybe a poor educational background, of course

the content and quality of the speech will be quite different. So these kinds of ratings may not be so easy to make, especially in very rapid assessments for some kinds of studies, and even in the emergency department.

What about avolition and apathy, like lack of grooming and hygiene, impersistence at work or school? So some communities may dress very differently. Some may have different cultural norms for dress and behavior. Some may seem to be maybe a little bit lacking in personal hygiene. We have to be careful before assigning that label to them.

There's many, many cases I could talk about, but I'll save just time for one case to review with you, just in a second. Anhedonia and asociality having to do with relations with friends and peers, recreational interests and activities, of course all of these issues can be vastly affected by cultural norms, even family norms and social class norms.

Attention, social attentiveness, and inattentiveness during the mental status testing. Imagine if somebody comes to see you who's feeling fundamentally misunderstood. They know there have been problems between their community and the police. They were brought to the hospital or to see you by the police or by the judicial system. You know that they're not going to be comfortable. They might be scared or apprehensive. Of course they're going to be feeling -- they might be coming across as inattentive or disengaged.

A quick quotation to read to you as well about negative symptoms in cultural context, the overlap between depression and negative symptoms, by Nancy Andreasen sometimes ago: "Just as manics and psychotic depressives are likely to have delusions and hallucinations, so too depressives are likely to have some negative symptoms, such as alogia or affective blunting."

And then the role of social

adversity in the production of pseudo-negative symptoms: "Thus, after experiencing the illness for many years, it is possible that indirect environmental factors, like economy, mass media, politics, government, laws, begin to exert a greater effect on their ability to perform recreational, goal-directed, and social activities that are the foundation of negative symptoms."

So in conclusion, a quick case example. This just was happening on our service in Montreal, and I just wrote my colleague this morning who is continuing the evaluation. So a young, 18-year-old, African-Canadian man referred to us with a possible history of first episode psychosis in 2020. He had auditory, tactile hallucinations, this is all by report -we didn't see any of this - emotional blunting, avolition, paranoia for one year, and then no symptoms for three years. He's taking no current medication. There's not current substance use.

There was a psychiatric evaluation in December of 2023 which diagnosed recurrent depression, but he's not currently depressed. But he went to see a clinic where they began to evaluate him, and let me just give you some quotation from the actual referral, it says "However, though, through multiple subsequent individual psychotherapy sessions, we strongly believe that the patient would benefit from an in depth evaluation to rule out psychosis due to persistent negative cognitive psychotic symptoms."

And they go on and say "Since 2020, he began experiencing restrictive affect, less ability and desire to communicate with others, and less anticipatory pleasure about things that he used to look forward to. As a result, he has found it harder to maintain relationships with others. He's also found it more difficult to feel attraction and romantic interest toward others. We would appreciate your expert assessment to help determine if

there is enough evidence to appropriately consider that he suffered a psychotic period, and that he continued to struggle with the negative symptoms of schizophrenia."

So, kind of an interesting referral and consultation. Not so different from ratings or assessments you might do in research contexts as well. My colleague has continued the evaluation. He's convinced that he has depression, this young, African-Canadian man.

The clinic had been pushing him toward the traditional trajectory, toward a schizophrenia profile. We felt that it was more depression, but he did think he had a psychotic episode after reviewing the history and talking to the young man carefully about three or four years ago. So that's kind of -- it's an example of the conundrum we face in all the work we're doing with negative symptoms and schizophrenia in a cultural context.

So, what can be done? This is always the question. I really am, like many of

you, raising questions. I don't have definitive answers. But we can -- it's a discussion and an ongoing process. We definitely do need to culturally adapt our surveys and instruments. As everyone knows, it's expensive, it can be complicated and time consuming, but critical.

We need to train clinicians, researchers, and raters in what we call cultural humility. Cultural humility just means we don't know everything as professionals. We have limitations to our knowledge, to our understanding. And our patients, our clients, or their family members can teach us a lot about what we need to know to help them.

We need to make sure to include diverse participants in clinical trials and other studies, but even maybe more importantly, we should include members of diverse communities in the research process and as members of our research teams, and then follow

their recommendations and suggestions. This is often not so easy to do in practice.

There is also community outreach and qualitative studies where we can go to hear what people who really have a stake in what's happening to members of their community can tell us about these problems. We need more data, especially ethnographic type data, which is very different from a lot of the quantitative data that we have -- has been gathered, and that is discussed mostly in this field.

Cautionary note, practitioners should always evaluate whether psychotic-like experiences, including negative symptoms, may be better explained via ethno-cultural context.

So my conclusions for this part of this panel, negative symptoms of schizophrenia are understudied in members of minority groups, which everyone knows. Rates of negative symptoms likely vary by ethnic group. But how much of this is due to cultural variation of

illness expression is essentially unknown.

Clinicians and researchers need to adopt a position of cultural humility in their work with minority groups. And members of minority groups need to be part of our research teams, and their recommendations need to be implemented to the degree possible.

And here are the references, which I think are online. You can consult those in case you have doubts about what I've been saying. Thank you so much.

DR. CAMPBELL: Laura, I now turn it over to you to begin your presentation. Thank you.

DR. SWETT: Hi, good afternoon. Thank you for being here. What a privilege it is to be a part of this ongoing discussion related to effective ways to measure negative symptoms of schizophrenia from a regulatory perspective.

Just wanted to capture a little bit of what we've heard so far related to patient-

focused information. We heard about why negative symptoms are clinically important from Bernie. We learned about the lived experiences, how negative symptoms impact people who have been diagnosed with schizophrenia, thanks to Brandon.

Sophia made us aware of relatively new and ongoing conversations regarding the interaction between cognitive and metacognitive factors and their interaction with negative symptoms of schizophrenia. And Eric, thank you for your insights regarding the importance of culture and how it influences our experiences, and therefore our interpretations of different signs and symptoms. We can't assume they'll all be interpreted in the same way by different cultures and subcultures.

Today, I'll be talking to you about regulatory considerations for assessing clinically meaningful within patient change, and Tiffany nicely set this talk up with her comments just before lunch.

This is my disclaimer.

The purpose of my presentation is to set up a framework to discuss clinically meaningful within patient change, and today I'll be discussing three topics.

The first will be the types of clinical trial measures we see when we're conducting a regulatory review, mostly clinician reported outcomes in this context of use, and how the most popular or well-used measures, which are ClinROs, contain an important perspective, which we rely on, of course, for a diagnosis in clinical management, but there is also, of course, an opportunity for a more comprehensive multi-perspective approach. And this is an approach that is generally laid out in our patient focus drug development guidances and applied also to other therapeutic areas.

So I will be discussing also number two, the important concepts from the patient and caregiver perspective. For example, are

some concepts considered to be more important to change than others from a patient and caregiver perspective when we're looking at evaluating clinically meaningful change?

And then lastly, I'll be discussing other types of clinical outcome assessments in terms of looking at clinically meaningful change.

So, this is a snippet here of the clinical outcome assessment compendium, the latest version reflected from June 2021. It captures the schizophrenia disease condition, as you can see, and the second to the right column lays out clinical outcome assessment measures that have been used in clinical trials for approved therapies.

And in the context of schizophrenia, as you can see, historically, ClinROs have been used to assess negative symptoms. But one of the questions we want to ask today is "are clinical rating scales enough?"

We know that, of course, tools have

different uses, and those have been established either through research or clinical practice, may not be sensitive and interpretable in registration trials, even if foundational work on the content has been conducted.

So when a COA is used as an endpoint, we ask, "does it reflect how a patient feels, functions, or survives, which defines treatment benefit?" "Has evidence been supplied to demonstrate the patient experience?"

Through the CARES Act and PDUFA VI, we have an agency commitment to patient-focused drug development, even as exploratory, to collect patient experiences. So for example, moving beyond a diagnostic construct and utilizing other stakeholders, such as patients and caregivers, to understand negative symptoms of schizophrenia, is a part of this commitment.

As we know and have heard, there have been a lot of conversations regarding the most appropriate ways to measure negative

symptoms. We've -- are well familiar with the NIMH MATRICS consensus definition and those five domains.

However, when we are asked to advise on whether a ClinRO has sufficient validity evidence to support its use in the context of negative symptoms of schizophrenia, we find that direct patient and caregiver feedback, or their perspective, were omitted during the instrument development.

We understand that there are challenges in a disease context where self-report is hampered by limited insight, cognitive impairment, or other factors. But potential insights can still be obtained by patients, by caregivers. Our patient-focused drug development guidance 2 lays out some methodologies for collecting that kind of evidence. For example, a focus group with patient and caregivers dyads, or one on one interviews with dyads, or a part -- or an example of one way to collect that information.

Looking into literature is another.

Before I talk about clinically meaningful change, first we need to understand that meaningful concepts that have been identified by patients and caregivers, and whether or not they are incorporated in to a measure. And then we can discuss meaningful change.

So, obtaining these insights can help us to understand - these are listed here on the slide - which concepts of negative symptoms of schizophrenia are important from a patient perspective and caregiver perspective. We heard a lot this morning about increasing drive and decreasing apathy as a point of intervention. What treatment goals are the most important to address in terms of each concept?

And then thirdly, which aspects or attributes of these concepts are relevant from a patient or caregiver perspective? And from this, I mean if we look at these aspects of --

sorry about that -- look at these aspects of these concepts, we're looking at presence or absence, frequency, intensity, or duration. And so we're wondering from a patient perspective - let's just look at avolition or amotivation - what aspect of that domain would be important to a patient, and what would the MOA be targeting?

So for example, would a patient or caregiver consider that duration as important, even if it just moves from a very, like, lack of motivation to mild motivation? Is it important that the intensity of the motivation improves so that, let's just say, the motivation was a zero out of zero and moves in treatment to a six out of ten, is that important? These are the types of information that we find helpful in order to help us understand clinically meaningful within patient change.

So, once these concepts have been identified, are they incorporated into

currently available measures of negative symptoms, or do the measures require modification or supplementation, or does a de novo measure need to be developed?

So in the second part of my talk, I'm going to be discussing aspects of meaningful change that we recommend you consider from a regulatory perspective. And I have these listed just as a series of questions, and understanding, of course, that when you're looking at change, and in a drug development program it's going to depend on your mechanism of action and other factors, but the first consideration is, when observing a change in negative symptoms, can we assume, when change occurs with treatment, that the change is a result of each of the domains moving equally? In other words, do the domains move together? Can we assume the neuromechanism of change impacts all domains equally, or are only one or two particular domains driving the change? We'd like to know

what's driving the change.

Secondly, regarding the NIMH MATRICS consensus domains, can you demonstrate that all concepts are considered to be important to all stakeholders? In other words, do the concepts identified by patients and caregivers align with clinician observations, or are they different?

Thirdly, regarding treatment, when listening to caregivers and patients, which concepts do they consider to be the most important to treat? If I heard Brandon correctly this morning, enhancing reward anticipation would be a treatment goal. Of those most important concepts to treat, which concepts are considered to be the most bothersome?

Fourthly, what might be the most meaningful concept of change from the perspective of a caregiver or from the perspective of a patient, which we've heard may differ? Are some concepts more important and

more bothersome or less important or less bothersome than others?

Fifthly, how does this align with the mechanism of action that the drug is targeting in terms of how a patient feels, functions, and survives? So if a drug is targeting avolition, for example, is this a domain that is meaningful and important to caregivers and patients, or alogia?

Sixthly, when we are looking at meaningful change, when is it that we consider clinically meaningful change at the group level, so inferences are made regarding a population which may be of interest to a health system, versus at the individual level, so establishing that a certain proportion of patients benefited from treatment, which may be of interest to a healthcare or treating physician?

And then the last two questions we're raising for your consideration when you're looking at measurement and regulatory

setting, how much change is considered to be meaningful, improvement or worsening, within each key concept?

Worsening, we find, is as important to the patient, caregiver experience as is improvement, particularly if the treatment is impacting that. We'd like to understand what this looks like in order to be able to interpret meaningful improvement or meaningful worsening in the context of a clinical trial.

And then, lastly, from a regulatory review perspective, we're interested in evidence demonstrating the link between the improvement of negative symptoms and improvement in functioning as a part of the feels, functions, survives focus.

I think the big message to relay is that we are interested in impacts for their own sake as a result of our patient-focused drug development initiative, but also as supportive information for interpreting primary and key secondary endpoints that assess signs and

symptoms, even if those impacts are not going to be mentioned directly in labels. And also, we do document in our reviews PFDD (patient-focused drug development) evidence.

And then lastly, to address the third aspect of this discussion, if COAs are supplemented with a patient-focused drug development approach, what are other ways meaningful change data can be captured? And there's been some discussion about this already today, and I'll just mention three different aspects that might be helpful.

Digital health technology measures, we have a guidance that we put out in December of 2023 called "Digital health technologies for remote data acquisition in clinical investigations" and this helps you in your development or modification of a DHT (Digital Health Technology) to ensure its fit for purpose.

And there are examples of DHTs being used in this context of use, which would be

schizophrenia, such as a virtual reality functional capacity assessment. So, those are considerations that can be made in terms of looking at meaningful change.

The next one is an observer-reported outcome. The caregiver-reported outcome measure would be quite valuable in this context of use, because there will be expected differences between a patient perception of change and a caregiver's perception of change.

And so it is important, if using this type of measure, to standardize rater training and to demonstrate adequate test and retest reliability. This would be part of providing the evidence of the reliability and validity of the proposed ObsRO.

And then, lastly, videos would be another means of assessing clinically meaningful change in terms of, for example, a task that has been directly linked through patient and caregiver clinician evidence to a negative symptom. And if that task is

considered to be meaningfully connected to that system, that can be conducted throughout the clinical trial via video, and centralized raters can be trained to rate that particular behavior, and in changes of that through the course of the clinical trial.

So, those are just three examples of alternative methods.

I would like to say in closing that identifying clinically meaningful change helps us to interpret treatment benefit, and supplying evidence that the chosen measure in your trial is fit for purpose, including patient-focused evidence, ensures that your clinical outcome assessment will reliably and validly measure the concepts of interest.

And then as Tiffany said earlier, whether you're assessing a current measure, modifying, supplementing, or creating a de novo measure, we recommend that you consult with the FDA early and often. Thank you.

DR. CAMPBELL: Great. Thank you,

both Eric and Laura, for your presentations. I think we are going to have a really dynamic discussion for the next 30 minutes. We also do have someone joining us virtually for our panel. So, they're going to bring her up on the screen to us.

But what I want to -- I want to give Laura and Eric a little bit of a break from presenting, and I want to turn to our panelists, and we have a really great representation on our panel.

So I'm going to ask our panelists to introduce themselves to the audience, and then provide one to two minutes of reflection of the presentations we've just heard, and I'm sure will also be reflective of our morning conversations. So Matt, may I start with you, please?

MR. RACHER: Yes. Thank you so much, and a pleasure to be here today. My name is Matt Racher. I'm an individual living in recovery from psychosis and schizophrenia. I'm

also a certified recovery peer specialist and a master's level social worker starting -- just starting a job as a clinician.

You know, I'm -- in reflecting upon the presentations today, you know, I think about my own personal experience, you know, kind of what my pathway from psychosis to sort of putting out the fires of psychosis with therapeutic help, medication management, recovery supports, and kind of the long pathway towards my recovery from certain domains of negative symptoms into feeling connected to, you know, meaningful, purposeful activity, et cetera.

So in thinking about this, you know, it was a long and challenging journey, and, you know, I think it's important to kind of asses what for me was -- felt like a loss of my former self, and a gradual return to connecting with family, with friends, with close peers. And I think in between that time, that's where kind of this long process of help was needed in

a sense of -- so I just wanted to introduce with that topic and that response.

DR. CAMPBELL: Thank you, Matt, for sharing.

Deana, I'm going to turn it over to you.

DR. KELLY: Hi, thanks so much for having me here, and thanks for the great talks, and the talks this morning. I'm Deana Kelly. I'm a professor at the University of Maryland at the Maryland Psychiatric Research Center. I'm also a PharmD, so I have the pharmacy perspective.

I've been taking a lot of notes today, so I'm thinking about a lot of things. And I'm -- so I'm going to -- I know you posed seven questions earlier for us to answer, or for us all to think about. I'm not sure I can answer any of those. But I'm going to add, probably, as Nina had said, add more questions to the mix as well.

I do think that starting off, like,

it's important in the real world for clinicians -- for us to understand that clinicians are short on time, and they will struggle sometimes between negative symptoms, depression, and we don't even talk about it a lot, but catatonia. And so that's an issue that's out there.

But also, this idea of primary versus prominent versus persistent is going to be even more challenging for them in the real clinic if we go down these pathways and try to define these symptoms.

So how we define research translates into how people are going to have to be thinking about this in the real world. So I think we do have to pay attention to these aspects as we design scales, we think about meaningfulness, we think about outcomes, and we have to, as Tiffany pointed out, put on the labels as well.

So -- and Dr. Jarvis sort of spoke to that earlier, about culturally, this does even get more challenging and different in

ethnic populations, potentially.

I think also, we haven't really talked about this today, but separating domains of symptoms and defining negative symptoms separately has allowed us, as a field, to accept that there's a set of symptoms that we can't treat. And that's how our clinicians feel. That's how, sometimes, we feel.

So regardless of how these symptoms actually occur or what causes them, sometimes we just throw up our hands and say we can't treat negative symptoms. So we have to get out of that mindset, too.

So I just want us to pay attention. As researchers, we talk about this all the time. But as clinicians, we forgot about treating negative symptoms, because we'll go after depression, we'll go after anxiety, we'll go after other things, but sometimes we're just not going to go after negative symptoms. So I think it's really important.

And Dr. Jarvis pointed out, too,

like, from the literature about people think there might be brain damage, and there's nothing we can change about that. So it's just important to think broadly outside of our research world on both the diagnosis and this idea that there's no way to treat negative symptoms.

I also want to reiterate that people with negative symptoms, regardless if they're primary or secondary, can be helped. If we change our thinking to align with the recovery focus or the recovery model, it helps us set aside just changes on scores and allows us to target behaviors, allows us to target attitudes.

Because we can indeed change negative symptoms. We can help people feel better. We can help people function better. And those are the things we want to do, as Brandon pointed out, thrive.

We can begin to poke holes in defeatist beliefs. We can increase competence

when there's actual performance deficits that are present. And we can assist people in initiating and engaging in goal-oriented behaviors.

Our hope is that pharmacologic treatments will be able to improve negative symptoms. But as Dr. Vinogradov pointed off -- pointed out, as other people pointed out, our best combinations might be treatments that help with motivation and help change, but also teaching people how to practice that, whether that's through CBT or CBSST or music therapy or whatever that is.

Like, I think that we're going to have to have study designs that are going to have to have the basis for teaching the skills or practicing the skills, in addition to improving care.

I think about a meaningful change. When we think about that, I think we have to pay attention to limitations. We talked a little bit about self-report negative symptoms

measures. And while they're incredibly important, I think about some of my patients and their inability to self-monitor and to not report what's there. So I think having input from clinicians, caregivers, as well as the patient, and take the best approach for all the information, however that may be, could be incredibly important.

Also, requiring informants in clinical trials. We really have to think about that. As you brought up, I have a clinical trial, a seven-site clinical trial we're running right now, and we have informants and we're looking at violence and aggression. And it is challenging to actually get reliable informants. So if we require that in clinical trials, we're actually going to diminish our ability to recruit people, too. So thinking about how do we go about getting all the information that we have out there together to inform negative symptoms, I think, is important.

And I think leaning towards some of these virtual technologies, as you brought up, are going to be possibly important for looking at some of these measures of functionality.

When we talk about what's important for meaningfulness, our team thinks about and I think about improving motivation, initiation, and engagement in goal-directed activities. And that's how we think about it. We think about how can we make someone feel better, how can we make someone function better?

I know that Dr. Correll, Christoph, didn't mention this today, but I've listened to him many times, talk about, like, these four -- some of these domains, occupational, functional, social, and family, physical health, living arrangements. Like, these are some of the things that really are important for people to function better. It may be what matters the most. But patient perspective, as you said earlier, is really going to matter. There wasn't a lot of discussion around

functionality today, but I think that we have to keep that on the table.

And cultural considerations, as Dr. Jarvis pointed out, they're extremely important. How do we incorporate this into our measurement of outcomes? How do we ensure that we pay attention to cultural norms, such as, as you mentioned, cultural mistrust, eye contact, dress codes, hygiene, how they differ by our different contexts?

And then -- we didn't talk about this, but I think as I heard about it today more and more, it's going to be important to ensure our research teams are diverse to ensure good interpretation and assessment of actual behaviors, roles, and measuring symptoms, and make sure that we're not narrowly focused, just coming from our own biased context, as we all have.

So we likely can do a lot better at that, and there's probably a lot more for discussion around how we can improve that. And

I loved the input from our lived experiences today. And I think that's critical for informing all of these outcomes that we're going to be talking about.

So we all know that negative symptoms impact people's lives. Negative symptoms impact the global functioning, and many different functional impairments in many different areas of people's lives.

We may be far too committed to just already the idea of negative symptoms change or this co-primary, as we've talked about off and on. But could we be interested in possibly another outcome of functional improvement, not necessarily co-primary, but another FDA indication, possibly, for functionality? No one's really talked about that. It might be silly. It might be too simplistic. But is it a possibility to think about, can we improve negative, but could we also have indications for medications that could improve functionality.

I mean, Dr. Keefe mentioned today about a good night's sleep. We can measure that, but how can you -- and then measuring how he functions the next day after a good night's sleep is something possibly a little bit different. So it's just a thought. So those are my thoughts from today. Thank you.

DR. CAMPBELL: No, thank you so much. Mark, I'm going to turn it over to you.

DR. OPLER: Thank you. Hi, everybody. Mark Opler, Chief Research Officer at WCG. I want to begin first just by addressing a comment from this morning. Somebody asked, very astutely, how do we know what the right level of volition is for an individual? The answer, actually, is five, moderate severe. That's the appropriate level of volition, if anybody wants to know [audience laughs].

You know, I want to -- on a slightly more serious note, I want to start by saying that much like me, our existing rating scales

haven't aged well over the last 20 years. They're creaky at the joints. They've lost that certain something.

You know, to give you an example, you know, in a lot of debate and discussion these days with folks who want to measure functioning and negative symptoms and the intersect between the two, we're talking about a scale called the UPSA, which many of you are very familiar with and have probably used a lot. Well, when it comes to the UPSA, my question is, you know, I've got folks on my team who probably aren't great at things like check writing, don't know what 411 is, honestly, have never dialed it and never will.

Our scales are getting older, and they no longer reflect even the dominant culture that we theoretically live in. That's a problem. So, you know, maybe some of the tools that were mentioned before, like the VRFCAT and other things, will be a better fit for the culture that we actually live in, and

also maybe more culturally adaptable. When we have to go to Thailand, we're looking at a completely different set of ideas, norms, daily life patterns. Our tools have to adapt. Otherwise, we're going to miss the boat.

The second thing I want to point out is that, you know, in addition to being somewhat creaky, a little culturally inflexible, it's very evidence also that our scales frequently don't measure what matters most. That's a phrase that's been said a lot today. I remember hearing it first in this context in the work of a friend of mine, Dr. Lawrence Yang. Look him up if you don't know him. He writes a lot about stigma.

But, you know, he started using the phrase "what matters most in a patient's life" as a way to think about what treatment means to them. A doctor might tell a patient, this is great, your voices are better, you're doing very well. Maybe. But if they're not living a life that feels important to them, have they

really improved?

The other thing I want to say, you know, there's been a couple of folks today, earlier this morning and recently, who've mentioned informant data, and how onerous it is to gather informant data on the PANSS, and for other scales. This is true. It's another checkbox that has to be checked.

Nevertheless, I would like to submit it's a vitally important piece of information and a vital perspective on the patient's actual status. In the PANSS, informant data isn't there by accident. It was put there very deliberately as a requirement, because when the folks who were writing it, sat down and looked at what they needed to rate, they realized they couldn't reliably judge the social performance of somebody they had only known for 20 minutes; they weren't the right person to determine whether or not what they were seeing made sense. They went to informants because they needed that perspective.

So, I'm sure folks will be happy to hear we're actually revising some of the informant tools for the PANSS. That is coming. I believe that the need for a better observer-reported tool that's relevant to schizophrenia will be found in the future. It might come from the past.

Last thing I'm going to say is, I think, you know, we also need to stop considering all of these measures -- observer-reported, clinician-rated, patient-reported, digital health technologies -- we've got to stop thinking of these things as separate and distinct.

Because I have a sense that what's coming in the future will be completely different. It will be a merging of these things in ways that we hadn't previously considered, whether it's the extraction of vocal biomarkers from clinician ratings or the incorporation of a virtual informant to help a clinician get to the right score on negative

symptoms. This stuff is coming, and it's not going to look the way we expected.

Finally, to just close out and let us get back on with our work, there's a patient that I've gotten to know a little bit in the Bronx, and his name is Corey. He's a wonderful guy. And I once asked Corey, "Corey, what do you really want from the pills that all the doctors are giving you?"

And he said, "you know what I'd really like? I just want to look and sound like everybody else. When I'm on the subway, I don't want people to stare at me. I want to be taken seriously. I just want to look and sound like everybody around me."

And I've taken that to heart, and I think when we think about negative symptoms, let's not ignore the importance of that idea. So thank you.

DR. CAMPBELL: Thank you so much. I want to turn it over to our virtual participants.

Bonnie, can you -- first of all, can you hear us?

DR. KAISER: Yes. Can you hear me?

DR. CAMPBELL: -- unmuted yourself. We can hear you. So can you introduce yourselves and provide some thoughts?

DR. KAISER: Yes. And thank you so much for inviting me to participate and letting me join virtually.

So, I'm Bonnie Kaiser. I'm at the University of California, San Diego in the Anthropology Department in the Global Health Program. So my main focus of research is cultural adaptation of measurement tools.

And so I just wanted to kind of build on that thread that's been mentioned a couple times by folks just the importance of cultural considerations, particularly of when we do global studies, you know, of incorporating rigorous cultural adaptation of our assessment tools.

Researchers are sometimes reticent

to do a cultural adaptation process rather than kind of a simple translation/back-translation process. One, because it's time-consuming to do more, to do cultural adaptation, and particularly validation. And there's also concerns about moving away from using kind of a strictly translated version of a previously validated scale. Although previously validated, you know, usually means in the U.S. or in Europe, not the kind of local context where the research is going to be conducted.

And we've found that there are studies that show that culturally-adapted scales do perform better in subsequent validation studies. When we don't do cultural adaptation and we just do kind of translation/back-translation, and then trust, you know, that that's going to work, we end up with some confusing results. Like one said he found 97 percent of their study population had PTSD. Like, we just don't trust that on face value. That doesn't make sense.

And so when we're not confident that we're measuring what we're trying to measure, then it really matches or maps on the lived experience in the context of our study, we can't really be confident in any of our results. We don't really know what our data are telling us. We don't know what we can do with those findings.

So, I'm an anthropologist. I do a lot of ethnographic research that then feeds into mixed method studies, cultural adaptation studies, validation studies. So, I see kind of the ways that this research, the kind of preparatory research for these measurement tools can really improve our outcomes.

And there's also been, you know, studies that show that this actually ends up saving money ultimately in terms of how we're effectively identifying folks in need of care, effectively referring them for care.

And then finally, I just wanted to pick up on one point that Dr. Jarvis mentioned

in his study, that, you know, we talk a lot about culture and cross-cultural considerations, but really thinking about kind of the broader context of environment also includes thinking about structures, thinking about systems, and how that shapes possibilities, and how that influences behavior.

So, Dr. Jarvis gave the example of, you know, healthy cultural mistrust. But I just wanted to make sure that we think about those issues as well, and consideration alongside, kind of more specifically cultural considerations in global studies and, you know, not just in global studies. Thanks.

DR. CAMPBELL: Thank you, Bonnie, for your remarks. So, we're going to try to have a pretty fluid conversation amongst the panelists, and I'm going to start with Matt. And he knows this, so we did prepare for this with the first question.

And I want to thank him and Brandon

for being here and feeling comfortable to share their lived experience with us today, because it is extremely valuable to us.

So, we've been talking about clinical meaningfulness, and one of the most important things that we look at when we're reviewers, and we understand that everyone's lived experience is slightly different. And we have to have this understanding of what is that lived experience, and one of the best ways is through qualitative work.

So -- but what I would love to ask you, Matt, is what would clinical meaningfulness look like to you? What would success or improvement from a treatment look like for you in your everyday life?

MR. RACHER: Absolutely. Thank you. If I -- I'd love to preface the answer to this question with kind of a quick analogy for the experience of negative symptoms, if that's all right.

So, I have this sort of analogy that

I came up with over time, and I'd like to share it. It's, you know, imagine there's a beautiful, thriving community within a small town where people work together harmoniously, interconnected in their efforts to support one another, and at the heart of this community stands a central building, a hub that provides vital resources, serving as an essential cornerstone of the town's wellbeing.

So, one day, you know, disaster strikes. The building burns down in flames, kind of like psychosis. The fire department responds, extinguishes the fire, and this once vibrant center is reduced to a pile of rubble, almost like the experience of negative or cognitive symptoms.

So, this once-lifeblood of the community is now in this state. So this metaphor, I wanted to share it to kind of reflect after a severe episode of psychosis like the one I experienced in 2011, you know, it felt like this profound kind of -- at first

it felt like I was going to lose who my former self was.

And I really think the pathway towards seeing meaningful change or meaningful outcomes is really having the supports along the way to kind of show that, you know, I still have -- or people with schizophrenia and in recovery from schizophrenia still have the desire to work, to love, to find connection to purpose.

And that's a slow process, almost like a light dimmer. Not necessarily a light switch, on and off, but kind of a slow, gradual process to reach those goals and to become connected and to become -- to reinvigorate or re-instill a sense of emotional connection and purpose to passions. For me, that's music, you know, and helping others, and working in the field of mental health and social work, so.

DR. CAMPBELL: So -- well, thank you, Matt, for that. And I know when we talked, you talked to me, you gave me that

analogy about a light dimmer, and I think that was a really informative way of structuring this gradual aspect, right, of, you know, it may be a low light on that dimmer switch, but you want to increase over time.

And that may take time, but there is a range of what meaningfulness could also look like, depending where you are in that current moment. And so I really appreciate that analogy.

So I have been taking a lot of great notes throughout the day and throughout the session, and so I honestly don't know where to start, but I'm going to attempt.

So I think what we've heard today is that through various lived experiences is, how do we balance -- and something we heard earlier today is how do we balance improvement and the important concepts? So how do we really be able to capture what is meaningful, what matters most, whatever buzz term you want to use?

Particularly -- but how do we balance that with when people with lived experience with schizophrenia may not recognize their negative symptoms or the impact it really is having? And others can see it, but when we think about that meaningful change, we really also want to try to have some underpinning of what does that patient think?

And this is our struggle in a lot of our diseases and disorders in neuroscience, where lack of self-report can be problematic. And that's why we do have to rely on other informants and reporters to help us.

But the heart of it is, is what we're seeing meaningful to patients? And so I was wondering if our panelists had thoughts on that. So I'm going to start with Mark, because I see him head nodding. So -- because I know he's had some thoughts about this when we talked earlier.

But how do we really find that balance? Because that is part of this

conundrum of how do we then design the trial, and how do we incorporate all of those things into that trial design to be able to find an effective treatment?

DR. OPLER: Thank you for that question. I mean, I think two thoughts off the bat. You know, one, we have been trapped in the clinician's office for a very long time. And the more we can do to try to put context back into the work we're doing, especially in early phase and possibly in later phase development, the better of we're going to be in terms of developing treatments that actually mean something outside of the rarified environment of the clinical trial itself.

You know, one example that comes to my mind is the classroom study in ADHD. This has been a study paradigm in ADHD research for a very long time in pediatrics, and it's not revolutionary by any means. It's simply looking to see how kids are doing in an environment that matters, the classroom.

You know, what's the analogy for us trying to work on negative symptoms in schizophrenia? Is it a clinical interview in a small office? Or maybe it's a structured assessment in a group setting. Maybe the analogy for us in the world of schizophrenia research is group.

So, a new formulation of the PANSS, coupled with digital endpoints where patients are interacting with each other and with, you know, other folks in a group setting, might tell us something that we have been missing and introduce clinical meaningfulness and context back into the work we're doing.

DR. CAMPBELL: Thank you, Mark, for that. Does anyone else have any other thoughts about that? Deanna, I'm wondering if you may, just thinking about how you were trying to link the clinical practice with the research world and the trial world, and that some of it may have to go back to that practice balance as well, if you had any thoughts.

DR. KELLY: Yeah. I don't know the answer. I was hoping to come here and learn from others myself today. But, I mean, I started off talking about -- as I listened today and I thought more about this research context that we're all sitting in, that it was important to make sure we go back to that clinical perspective, that we go back to understanding, like, how people -- how the time that people have in offices with their physicians or their care providers, what they're assessing, what they're looking for, and what they're actually hoping patients -- or their patients might improve upon.

And it's going to be very different than, potentially, what we're looking at in the real world. But -- or what we're looking at in the research world. But I do think, like, I agree with what Mark had said. We have to figure out, like, what is meaningful? Like, I think about it as "are people feeling better? Are they functioning better? And how do we go

about measuring that?" I'm not entirely sure what the answer is. But I do think that we have to go for some of those outcomes as we think about what's really important. And I really, really think this time around, negative symptoms, including people's lived experiences and listening to people talk about what matters is going to be incredibly informative, and it's going to be critical for us as we think about that more. Thank you.

DR. CAMPBELL: Okay.

DR. OPLER: I can't help myself. I've got to throw myself in here.

DR. CAMPBELL: Go right ahead. And then I'm going to turn it over to Eric.

DR. OPLER: Very quick.

DR. CAMPBELL: Go right ahead, Mark.

DR. OPLER: You know, something that Deanna was saying prompted me. There's a fascinating old technique that's not used very much anymore called goal attainment scaling. I won't go into it now. If you don't know what

goal attainment scaling is, you should. It's coming back. We've used it a little bit in depression, and I think it's time to think about using it for negative symptoms in schizophrenia. I'm going to shut up now.

DR. CAMPBELL: Eric, do you have any thoughts to add?

DR. JARVIS: I do. I was talking in the break with Stephen about how we have our research protocols and our research structures and hierarchies, and we have to kind of follow things that are sort of in a certain way that will produce the -- or produce a result we hope will be replicable, and I think respected by our colleagues.

But I think from a cultural psychiatry perspective, it's all about adaptations, modifications, person-centeredness. It's about making exceptions. It's about trying to be flexible in how we apply the protocols and procedures and practices that we've learned and that we've

created.

So, it's a tension, and I think it's hard to implement often. But I think this is maybe going along with what you were saying, Mark, about group processes to try to find out how can we modify what we're doing to get a little different kind of input.

You know, so it kind of goes along with mixed methods ideas as well. Are there ways in our work that we can include a more person-level, or a life world kind of a reaction, or a life world input that can really nuance the findings that we're having? And I just worry that we haven't been able to do it so well, so.

DR. CAMPBELL: Well, thank you for that. I'm going to invite the audience, if you have questions, to start heading up to the microphone. But as I've been reflecting today, this conversation, and kind of bringing this back to drug development. So we know that most drug development's global, right? And so my

industry colleagues in the room will all head nod when I say they also have to work with other health authorities with their endpoints in their study design. But I think we would all agree that the hallmark of cultural adaptation and translation of these instruments that support those endpoints are sometimes thought about last, right? And unfortunately, I think it does -- the example that Bonnie gave, where we're not really investing in actually doing the qualitative work to make sure we're fully understanding that population we're going to go try to study in, in that country or that region, and enough time to have it be incorporated into trial design and endpoint selection.

So -- before we transition to the questions, how do we want to think about making sure we build in, and in the spirit of patient-focused drug development, early into that process of designing the trials, thinking about where our study sites are going to be, where do

we have to understand that meaningfulness to build that in early adaption?

And I know, Mark, you've got to have thoughts. Because we've talked about it a bit, and I think Eric's got some thoughts as well.

And then Bonnie, I'm not sure if you do?

DR. OPLER: Yeah. I'll try to be quick. I mean, I think in addition to other things, I also tell people a lot, don't overload your protocols. You know, the more measures you load in, the less likely you are to get data that means anything. I'm going to reverse that very slightly and say, you know, whatever we can do to strip away unnecessary endpoints that, you know, tell us stuff we already know, and replace them with opportunities to collect things like cultural formulation, information on, you know, what do you want to get out of this treatment? Things that are more culturally meaningful and more person-centric are a better bet than another

PRO that already measures something that you're capturing elsewhere.

DR. CAMPBELL: Eric, do you have any thoughts?

DR. JARVIS: I do have some. So in the current research we're doing, we spend a lot of time reaching out to communities, making connections to communities, the people we're going to be actually asking questions of, and try to figure out what's at stake to them. So this is, I think, it's a time consuming direction. I won't say it isn't. But I think it really changes the tone and the direction of the work you're doing. And once the communities you're working with, the people you're going to be studying trust you, and you can have real, honest discussions, they will really change what you'll be -- what they want.

They'll tell you that you have to change a lot of what you're doing. And it can be all the way from the title of your project, all the way down to the kind -- how you talk to

people, who needs to be in the room when you talk to people. I mean, it's major differences.

But the problem is, it takes time to get there, right? So people don't just come out with these problems quickly. It's a kind of a relationship.

We were just invited to an Afro-Caribbean parade in Montreal, and I was asked to give a talk on mental health at this parade. It was a really unique opportunity, but very different from what I was used to. So I realized that I was put into an unfamiliar position, like our patients are put into an unfamiliar one when they come to see us in a study setting or in a hospital setting. So I learned a huge amount from that one invitation.

Anyway, these things will happen slowly over time if you try to nurture those relationships.

DR. CAMPBELL: Bonnie, do you have anything you want to add? If you don't that is

okay. I don't want --

DR. KAISER: I agree.

DR. CAMPBELL: -- to put you on the spot.

DR. KAISER: Yeah. I agree with what everyone's been saying. I really appreciate those points. And I guess I'll just share, there's, like, a, I don't know, kind of trope in anthropology that we always get invited to join studies once they've gone wrong to try to explain why things are wrong. And I think there's increasingly a shift towards inviting anthropologists to the team earlier to try to avoid that happening. But yeah, obviously I'm biased.

But like, one way to approach it is that including anthropologists, including folks with, like linguistic expertise, you know, local clinicians that join the team from the planning stages.

DR. CAMPBELL: Thank you for that. So Heidi, really quick, do we have any online?

Okay. So here's how we're going to do question and answer. I'm going to start with an online question, and we have three people in the room that we -- I will come to you. And I just ask if you ask one question, if you have multiple, figure out what's the most -- the burning one you have. But Heidi, what is our question from online?

DR. WEHRING: Okay. Thank you to all our online participants. There are a couple that came in that I think might meld well with the next talk, but I have one here that's really, I think, impactful.

As mentioned by an audience member in the previous session, patient and external perceptions of functional outcomes don't always correlate. So from a regulatory perspective, how would you evaluate the discordance between outcomes resulting from patient observer and clinician raters. And of course, from a cultural and from a meaningfulness perspective, I think there might be some kind of rich

thoughts here.

DR. CAMPBELL: Yeah. So I also think that may be a good question for some of our next panel as well, or I'll take staff for you if you want me to.

So I think, you know, number one, I don't expect to have exact correlation among my different reporters. That is not the reality. I think what is important when we do see that discordance is the understanding of what was that perspective that they were providing.

So I think Mark gave a really great example of why that clinician perspective is important for certain things, because it helps with that perspective.

I think this is why it's important to do qualitative work and talk to Matt and to Brandon and those folks who can talk about what that experience was like for them and where they are at right now, to kind of give a way to help interpret and attribute the data we're seeing.

So I think that's really important for us to understand. I think when we take and review this data, we're looking at all of it coming in. But the more details behind how the attribution or what was really going on or things that are better defined in a protocol is extremely helpful when we're interpreting all that data that we get that comes in.

And this is why we have these -- ask and encourage for this early conversation with us, and frequent conversation, so that when a sponsor starts seeing that too in our trials, that -- what do we need to think about? Was this expected? Is something happening? Do we need to think about this further? Is there adjustments needed?

Or maybe this is just actually the reality of the treatment, and we need to make sure we have good documentation with that data for that interpretation piece. So Laura, and then I'll go to Eric.

DR. SWETT: Yeah. I just -- thanks,

Michelle. I just wanted to add to that, that this is such an interesting disease in terms of there's some constant symptoms and there's some ebbs and flows or waxing and waning, and there's some great benefit to getting information from patients, like Matthew had mentioned, that you had an episode in 2011, and then you have the different perspective now than perhaps if we had gotten your insight as a patient at that time.

And there's some real value to getting the post-evaluation of that and what would have been helpful now that somebody is back in their -- maybe just their normal sense of self. That would also be really helpful information to capture.

DR. CAMPBELL: Thank you, Laura.

Eric, did you want to -- go ahead. Go on.

DR. JARVIS: Very quickly. So in our studies, which are not clinical trials, but discordance is an opportunity for a discussion.

And it's -- it does take the time. But like you were talking about, it's a different kind of reporting from one stakeholder or a patient or a family member.

You're trying to triangulate data in qualitative studies often, and that's -- it's just the beginning of a very rich, sometimes very productive negotiation, I guess, of what the meaning is.

DR. CAMPBELL: Thank you for that. So, I'm going to start the first person up at the front microphone. I please ask if you can introduce yourself, so our online audience knows who's talking. Thank you.

DR. STRAUSS: Hi, everyone. Great panel. This is Greg Strauss from the University of Georgia. We've been doing a lot of research on culture just over the past year that we haven't published, and negative symptoms, and I wanted to make one comment and ask you all one question.

Comment: there are a few reasons why

there have not been cultural adaptations created for negative symptom assessment. So, one is that it's thought to be built into the assessment itself. So, raters are instructed to do is rate someone in relation to that person's demographic, age, sex, ethnicity. And that assumes that the rater has proper knowledge of those things, which is not always the case, of course. And their own cultural identity and awareness, which can interact with them.

There's no training that I know of to train raters to develop the type of cultural awareness and understanding of factors related to motivation, emotional expression, social behavior related to different ethnicities and other aspects of culture. And to increase validity, that needs to happen.

The second comment in relation to that is, the literatures lags behind because people assume that a deficit is a deficit, that the absence of a behavior comes from the same

process, regardless of what the absence is.

But there can be active cultural processes that differ across cultures. I'll give you one example, the strong Black woman schema. We have found that that is positively associated with the severity of all five negative symptom domains in people with schizophrenia.

And it's an active cultural process that occurs in that community that's very, very normative, non-pathological in general, but can contribute to some symptoms like depression and anxiety. But there can be active cultural processes that contribute to negative symptoms.

And here's the question I had for you guys. We're finding that context matters a lot. So for example, when you have incongruency between the ethnicity of a rater and the ethnicity of a patient, you see an increase in symptom severity.

We've even had the same patient interviewed by a White rater and a Black rater

-- for a Black patient, for example -- in the same week, and you find differences based on who is interviewing them. And the question is, are those genuine differences in the behavior? Do the people behave differently in the interview depending on the rater, or is it that the raters are rating the person differently because of their own culture?

So the question that I have for you is: how do you tackle that question, and how would you account for?

DR. CAMPBELL: Oh, wow. I'm going to -- I mean, that's a fascinating question in general, and we can apply it -- let's go global and all of that. I think that's a fascinating question.

I don't know if anyone had a quick thought about that? Okay. So, Mark, and then Eric. I mean, I'm sitting here rattling ideas off my head, and I'm like -- I've got a lot of thoughts. But Mark and Eric --

DR. OPLER: I'll be quick. You

know, back in the 50's, they did a big study in Manhattan, the Midtown Manhattan Study, like, one of the big first, you know, epidemiologic studies of mental illness on a population level.

And the doctors who were running it at Cornell realized a lot of folks there weren't necessarily from America. There are all kinds of languages and cultures. And for the first time in history, they said quick, call an anthropologist. Until today. We've got Bonnie.

Well, that was my grandfather that they called back in the 50's. And, you know, he realized a couple of things. You know, one was that there's social distance between the researcher and the subject. And we've stopped recognizing that.

Wouldn't it be interesting to start measuring that again, the level of social distance, you know, the cultural milieu of the sites where we do this work in? It's not that

hard to do.

To collect it as part of a meta-study would be an incredibly valuable thing for the question you're talking about. And I think it's -- this is data. We can't answer these questions until we start collecting data on it, and I would love to see that happen.

DR. JARVIS: Okay. That's really fascinating what you're saying, and the findings that you were just describing, I hope you can publish them. I think I may have cited one of your papers. The -- you're the Strauss. I said -- okay, excellent. I'll come and talk to you.

So anyway, the thing is that the finding you had about the strong African American woman, I would just run those ideas by members of the community and see what they say and what they think, and get a wide -- go out and get a wide -- that's just one idea.

The other one is that we work with culture brokers. There's linguistic

interpreters, and there's cultural interpreters as well. So it's kind of what Mark was saying. In this case, Bonnie is kind of like -- as an anthropologist, could be kind of like a cultural interpreter for certain kinds of things. But sometimes you need somebody much more specific to the community to help you understand what you're finding. You know, maybe an anthropologist wouldn't know or wouldn't have that kind of inside information. You know?

So those are a couple thoughts, just listening to the work you're doing, so.

DR. CAMPBELL: All right. So I'm going to take a question in the middle, and then I'm going to end with Nina up front. So person --

DR. KIRPATRICK: Comment based on --

DR. CAMPBELL: And who -- can you please introduce yourself?

DR. KIRPATRICK: Yes. I'm sorry. I'm Brian Kirkpatrick, University of Arkansas,

and with Quantic Innovation. I have reached a conclusion based on experiences with my wife, who is from a different country, has a strong accent, different ethnicity.

Watching her maneuver in her country, watching her maneuver in this country, I've come to the conclusions that Americans -- that a lot of things that we think are involved with ethnicity are really about social class, education, money. And I think that in our country, they're very confounded, and a lot of other countries, they are as well.

But I think that a lot of what we tend to attribute to one thing is really from another. And I think that we, in this -- in the research I'm hearing, I haven't heard that addressed. And I think it would be useful.

I would hasten to say, she came from -- she married down. So --

DR. CAMPBELL: Well, I think that --

DR. KIRPATRICK: I come from a bunch of rednecks, and she definitely does not.

DR. CAMPBELL: Well, I think that it actually kind of goes to the conversation that we just had about that, adding that on. And I think when we had our prep call, when we talked about cultural adaptation, and just diversity is, you know, we think about drug development globally, but within the U.S. ourselves, we have so much diversity, cultural adaptations, different thinkings that we need to -- we need to really be taking this account early.

DR. KIRPATRICK: Cultural adaptation is one thing. I'm talking about class.

DR. CAMPBELL: Yeah.

DR. KIRPATRICK: I'm talking about education and money.

DR. CAMPBELL: Yeah.

DR. KIRPATRICK: And I think for the people in this room, including me, to some extent, we tend to be blind to that in a way that people who are lower social class are less so, is my guess. So --

DR. CAMPBELL: Does anyone have a

quick thought on that?

DR. OPLER: Just that, you know, there's a culture of money and class as well. And I think, you know, we're talking about different sides of the same dice, if you will. It's part of social distance, and I think you're right, we overlook it. We like to pretend it's not there because it's uncomfortable.

DR. CAMPBELL: Okay. And last question, Nina.

DR. SCHOOLER: So I'm going to end on a more mundane note. This is a question to Dr. Jarvis. I was really fascinated by that little vignette you presented of the example of somebody asking for guidance, first of all, by the degree to which that person was really comfortable with the jargon of negative symptoms and so forth. And so my question is, what's your recommendation?

DR. JARVIS: To the referring clinician?

DR. SCHOOLER: Yeah.

DR. JARVIS: Well, I mean, we often don't give the recommendation right away. We kind of say "let's talk -- we need to discuss what your issues are." This is a new referral to our service. So I -- our impression is it's not psychosis, right? That's our impression. So we're going to have to find a way to talk to the team that's very convinced it is, you know, and we'll have to start a negotiated sort of resolution about how to treat the client.

So this is often the way -- we look at our work as mostly centered on the referring team, not on the clients or patients themselves, because it's more of a consultative model, you know?

So that's the answer I can give you. My recommendation is going to be that we're going to wait a while and then talk to them about some of our impressions and see if they'll accept that maybe -- part of it might be the patient as well. The patient may be

unhappy. I can't remember the full story. He may have been unhappy with the initial evaluation of depression. So that may be partially driving this kind of a -- what do you call it, sort of settling into a psychosis diagnosis.

But it's a good question. And the way we work usually, like I say, is we kind of -- we take our -- a bit of time, and we try to hear the needs of the referring team and see what may be driving the referral and making them have such a strong -- take a strong position.

Then we'll talk a little about stereotypes as well, and how, like, a lot of people from African communities are pushed toward the schizophrenia, you know, world. And we'll say we often -- we just want to try other possibilities, because he's very young and new to this psychiatric system.

So we might try an antidepressant trial and see if that's going to be helpful.

And that way, if it is depression, maybe we'll kind of clear the decks and he'll improve. You know?

DR. SCHOOLER: And just to clarify why I asked the question, I was thinking this person might be a candidate for a negative symptom study and wanted to negotiate that. Thanks a lot.

DR. JARVIS: That would be true if I was running one, you know?

DR. CAMPBELL: Well, I want to thank my panelists. I want to thank Bonnie for joining us virtually. I want to thank you for asking questions. So we're going to end our session. We are going to take a five minute break, so -- because we're -- because I want to make sure I have enough time for our next dynamic panel.

So we can return around 2:18 East Coast time, for our virtual folks. We greatly appreciate it. But thank you.

(Whereupon, the above-entitled

matter went off the record at 2:13 p.m. and resumed at 2:18 p.m.)

DR. WEHRING: All right. Hi, everyone. That was a really short five minutes. I apologize. But I know that everyone will be really interested in hearing what our next round of speakers have to say. So, as folks are filing back in, I'll just go ahead and get us started, introduce myself, and invite the panelists or respondents and our speakers to come on up and task a hot seat up here for Session 4.

So, I'm Heidi Wehring. I'm a Clinical Reviewer in the Division of Psychiatry. I'm a Clinical Reviewer here in the Division of Psychiatry at the FDA. But most of my pre-FDA career actually focused on the treatment of schizophrenia. And clinical research moving towards helping to improve the lives of persons with schizophrenia. So, this is a topic really near and dear to my heart.

And we have just a fantastic expert

panel of speakers and respondents that are filing up on stage. A lot of these folks actually -- most of these folks have already been cited in the earlier topics today. So, there are going to be some familiar themes that are going to come in here. So, basically, we're going to start with looking at the clinical outcome assessments for measuring negative symptoms of schizophrenia. And talk about some of the non-clinical outcome assessment measurements.

So, I'll start this session with Dr. Jack Blanchard, and he'll be beginning our talk with looking at the outcome assessments in negative symptoms of schizophrenia. And I'll let him give a little bit of his background about why we chose him to give us discussion on this topic. Thanks so much.

DR. BLANCHARD: Why did you?

(Laughter.)

DR. BLANCHARD: My career has been dedicated to understanding negative symptoms.

From graduate school, looking at anhedonia to assessing negative symptoms. And I'm going to talk about that research during the course of my presentation.

So, I'll be focusing on the clinical assessment interview for negative symptoms, the CAINS. Giving a little background which I think I can move through quickly because of the prior conversations.

But scale development in the 1980s allowed us for the first time to quantify negative symptoms and to begin to understand their clinical significance. Critically important. Allowed us to advance the field. But over the years, a number of concerns were raised about these instruments, despite the advances that they brought.

I'm highlighting a few of these here. Basically, what we're focusing on is the inclusion of items, the inclusion of constructs that don't seem to be central to the definition of negative symptoms. And therefore, risks the

introduction of error variance in how we measure and quantify negative symptoms.

So, some of these scales include items that really, as I said, don't fit with negative symptoms. One example here with the NSA, is this idea of emotional range, where the lack of anxiety, sadness, anger, is pathologized as reflecting the presence of a negative symptom.

Or other symptoms looking at cognitive impairment. Cognitive impairment can be associated with negative symptoms. It's not part of the core definition of these symptoms. So, to include things like attention or abstract thinking, may be problematic.

The other concern is that when we look at assessing negative symptoms, many of them are defined by their experiential component. How do you feel? Are you motivated to do something? Are you interested in doing something? Do you gain pleasure from doing something?

But some of these scales don't ask about experiential aspects. They don't ask the participant how they feel. Instead, they look at the heater. They look at performance and infer deficits in motivation, infer deficits in pleasure from those performance deficits.

The other concern is that in some cases, we see poor reliability, either at the scale level or item level. And then some of these scales lack really detailed anchors, interview of scales, and other concerns about making sure that we can use these collaboratively and consistently.

So, all of these concerns were noted over 20 years ago. We're having similar conversations. Twenty years ago, there was this conference about how do we advance the field in interventions for negative symptoms. Out of the conference came acknowledgement about these limitations. And the need to develop next generation scales.

And so, out of this really were two

scales that were developed. Each taking a very different approach to scale development. I'll be focusing on the CAINS, and then Greg is going to talk about the Brief Negative Symptom Scale, the BNSS.

So, for our approach, what we decided to do is we really felt like this undertaking had to be significant and required funding from NIMH. So, collaborators Ann Kring, Bill Horan, Raquel Gur, we came together and we had a multi-site, multi-PI study to address how we could develop the next symptoms scale.

And the modifications that we addressed in developing the CAINS, are listed here. And basically, we were trying to fix those things, that we'd come to learn about limitations, with other scales.

So, we removed item content that was unrelated to negative symptoms. We looked at discussing experiential deficits. And the approach that we took was to start with a large pool of items, just like drug development. You

may have ideas about what's going to work. You may have your best guess, clinically informed research, informed about what's going to perform.

But ultimately, we wanted data to adjudicate the decision as to what items survive. What items got trimmed, refined, altered? And so, we started with that large pool, understanding that we would end up going shorter on it.

We then used advanced statistical techniques like IRT to inform us. What items are working? What range of scale do we have? Is it a five-point scale? Is it a seven-point scale? We could guess, but we wanted the data to tell us what to do. Finally, we created a manual. We have standardized interview probes. We have training videos.

And so, the 23 items that we started with, again, large pool, tapping those five consensus domains that were talked about earlier today. I'm not going to go through all

the different content but again, we're spreading a wide net and trying to give everything a chance to perform. And if it doesn't perform, we're going to trim it out, as you'll see.

And we did this in an iterative fashion. We started with an early data version, that had those full 23 items. And that first study in 2010 was really just "what can we learn about it? What can we do in terms of developing our measures?" And then in the 2011/2013 studies, those were the grant funded studies. Starting off with Horan, with 23 items, five-point scale, 281 patients across our four sites. Looking at all those features that I have listed there. Taking that down to 16 items, revising the scale based on statistics data. And then taking that out again, multi-site with Kring et al., with 16 items, ultimately trimmed down to three.

And the final scale, you see here, you have nine items. Tapping that, motivation

and pleasure that you heard about earlier today. And four items, tapping into expressivity, blunted alpha et cetera. We found that the scales were internally consistent. We had good rater agreement. We had convergent validity, discriminant validity with depression, psychosis, cognitive impairment, and short-term test-retest reliability. And so, from this, we disseminated that 13-item scale. And the idea was, for purpose of that grant, was to develop a scale that would be used, that would have an impact on the field. One way to measure impact is to look at our citation impact. Over 500 publications have cited that 2013 paper. The vast majority, over 400, are in psychiatric journals. Is that surprising? But we also have neuroscience journals. We have circa from pharmacology journals et cetera.

So, since 2013, in addition to these citations, we can interrogate the validity to the CAINS. What have we learned about it that

might give us confidence to consider it in a drug trial? So, I'm going to walk you through each of these, pretty quickly because of time considerations.

But the first issue is about replication. Can other individuals, not in our hands, use this scale? And how does it perform? And probably the best study that we have is leveraging the MOSAIC that had over 500 participants, across 15 centers. This is not a drug trial. It was not an imaging trial. It was simply trying to understand the phenomenon of negative symptoms at a representative sample, and how it impacts these individual's lives. And the battery was the CAINS along with other negative symptoms scales. And replicated exactly what we reported in 2013. Showed those same two subscales, internal consistency, discriminate validity, convergent validity, and now an extended test-retest reliability with over 400 participants, extending what we had reported on previously. So, this is

reassuring.

The other thing that we can look at is patient reported experiences. That's something that we've come back to throughout today. Great, we're getting these clinician meetings, what does it mean from a patient's perspective?

So, here what I've done is captured results across a number of studies, just a sampling. And on the two far-right columns, you have clinician-rated CAINS for the MAP and expressivity. And then on that left column, these are patient reported questionnaires, self-report questionnaires. They're reporting on the constructs listed there.

And we can look at the relationship between CAINS and these different dimensions. And what you can see, is that clinician-rated MAP is associated with many features of patient reports. They're reporting worse quality of life. They are reporting greater social anhedonia, less social closeness.

Individuals who have more severe MAP also are reporting more loneliness, less social-emotional support, less feelings of affiliation in direct encounters with individuals in the lab. When we asked them to report on the size of their social network, they're smaller, not surprising. And their self-reported social functioning was worse when they have higher negative symptoms as rated by the clinician.

One pattern that you can see here, is that the CAINS MAP is more consistently and robustly related to these patient-reported experiences, compared to expressivity. That fits with a lot of other literature. Happy to talk about that later if there's time to answer any questions.

So, the other thing that we can look at is real-world experiences. I know Greg is going to talk about that some more in terms of EMA [Ecological Momentary Experience]. For those of you who are not familiar, we're leveraging the fact that we're all carrying

around a smartphone. We can ping someone multiple times today at random intervals, and ask you in-the-moment, who are you with? What are you doing? How are you feeling?

And so, we can look at these clinician ratings on the CAINS, and does that relate to those in-the-moment reports and experience? And here are two studies recently done. Both use the CAINS, focusing on motivation and pleasure. In both studies, what we're finding is that higher clinician-rated motivation and pleasure deficits are associated with in-the-moment decreases in anticipatory pleasure for Merchant et al (2022), overall. And then for Abel at al (2024), focusing specifically on anticipated social pleasure. So, clinician ratings have meaningful relationships to in-the-moment experiences, out in the real world, as these people are experiencing and navigating their social world.

The next thing that we can look at is -- we talked about clinician ratings, we've

talked about self-report, we've talked about in-the-moment self-report -- but what about behavior? This morning, we heard, I found a very compelling example of an individual with negative symptoms, struggling for employment because of the interpersonal consequences of negative symptoms and the behaviors that are part of that.

And so, in this study we looked at the association of negative symptoms and social skills and how that may cascade into social rejection. And so, on the far left you see our predictors that we had. We assessed paranoid ideation, because of its relevance to interpersonal functioning, marked with positive symptoms. We had CAINS. We also were assessing sleep in this study because our lab and many other labs, have now established that sleep problems can contribute to functional impairment as well as symptom severity. And then in the middle column there, we had objective-behavioral ratings from coders, video

tapes of the social interactions from our participants. And they are rating social skill. They're rating positive facial displays.

And then finally, we had naive raters watch those same video tapes, and they're not coding on the skill. They're not coding anything. They're simply reporting on subjectively, "how do you react to this video of this individual? Would you want to spend time with that person?" And what we found was that negative symptoms impacted ultimately social rejection through social skills deficits. So, CAINS clinician ratings are manifesting in social behavioral deficits and that ultimately is having an impact on social rejection. Showing the meaningfulness of those clinician ratings.

The other thing that we can talk about is neural responding. That was something that you head about in great detail this morning. I'm not going to go into all those

different models. But I'm just going to touch on findings indicating that clinician-rated CAINS are related to neural responding. And so, what kind of neural responding might we want to look at? This morning we heard about reward. I'm going to tap on that. But there's also another benefit of social affiliation that we all experience. And that benefit is social affiliation helps us cope with stress. It reduces the challenges that we have when we are encountering threats in our environment.

This is sometimes referred to as the social regulation of emotion. And we can study that in the scanner. And so, we can bring people into a scanner, and while they're in the scan, they are watching cues. And those cues can be safety cues, "Nothing is going to happen, relax." Or they can be cues of threat, "In our study there is a chance of shock."

And what you see in that upper-left brain scan is the green. It's showing that activation as an individual responds to cues of

threat. Not surprisingly, you have widespread neural activation.

We then used a paradigm from Jim Coan at University of Virginia, and he studies the social regulation of emotion in healthy individuals. And we borrowed that here. So, the hand-holding paradigm is simply, you're watching these images, the cues alone. And then we have another trial where a partner comes in. An affiliate partner comes into the scanning room, says nothing, and simply holds the person's hand. And Jim had previously demonstrated that if you look at couples, friends, and they do that, you see attenuation of neural activation in the face of threat. So, affiliative contact is attenuating neural response to threat. And we asked the question, "do motivation and pleasure deficits, are they related to not experiencing that benefit of social affiliation?" And that's exactly what we found. Those individuals who have higher motivation and pleasure deficits, have less

benefit from that affiliative contact. They continue to show that neural activation.

The other thing that we looked at was reward. And we looked at two forms of reward, monetary incentive delayed tasks that you see there. And those blue triangles are pointing to the ventral striatum reactivity, replicating prior studies. In a scanner, doing a task, interesting video of a monetary reward. Money is falling into a glass jar. You see that neural network becoming activated in anticipation of that reward.

Well, we're really interested in social reward. Everything we've been talking about with negative symptoms, really focuses on social pleasure, social motivation. So, we took that same individual who held the hand in the prior protocol, videotaped them with positive social responses, smiling, thumbs up, great job. And instead of money, now they're seeing that brief video. And again, those blue triangles you can see the ventral striatum

being activated to social reward. We looked at the association between reward responses and motivation and pleasure deficits. And what we found was that MAP symptoms are related to blunted ventral striatum reactivity, to social reward, not monetary reward. And those associations held when we controlled for positive symptoms, depression. So, finding support for CAINS clinician-rated symptoms being related to neural activity to social reward.

The final thing I want to touch on has to do with sensitivity to treatments. That's obviously something of interest here. So, I just did a quick review. This is probably not comprehensive of the literature, and we have six studies here. I want to point out that five of these are psychosocial interventions. Some of these are RCTs, but if you look at the far right, I've indicated those negative symptoms where they're using CAINS. And where they're finding significant

differences between the active treatment and the comparator. And in these cases, the CAINS is detecting a significant difference.

There are a few studies, not many, and we have to be cautious about interpreting them, but there are a few studies that found CAINS detecting a signal, but other negative symptom instruments not detecting a signal. You want to be cautious about interpreting it, but it does lend credence to the sensitivity of the CAINS. And shows that it may be promising for future drug trials.

In terms of use in other settings, this is an unofficial list of CAINS translations. I've put an asterisk next to those that indicate that there's a published validation study. Other languages that you have there, those are investigators letting us know that they've translated it, but I don't think there's a publication on that yet. But the point is, is that it's been used in a number of countries.

So, just in brief, I've tried to overview some of the positive attributes of the CAINS and the research that we have so far, across multiple dimensions. Showing that it may have promise for use in intervention trials.

And I'm going to close by just acknowledging my collaborators, my students, current and former. And the grant funding from NIMH, that's supported this research. And I'll stop there, and I have time, actually.

DR. WEHRING: Thanks so much, Dr. Blanchard. So, we're going to go ahead and continue our session. I'll let Dr. Greg Strauss speak for himself. We are really lucky to have perspectives from different clinical outcomes assessments, as well as digital health and other techniques to be discussed next. Thank you.

DR. STRAUSS: Thank you, Heidi. So, I'll be talking to you about two things today. One will be the brief negative symptoms scale,

which we call the BNSS for short. And also digital phenotyping measures. So, I'm going to try to do double duty here and march you through both.

The BNSS is a 13-item clinical interview-based assessment scale that was developed in response to the NIMH consensus conference that Jack mentioned. And it was designed to measure the five core consensus domains. It also has one additional domain that is measured, the lack of normal distress.

It's rated after a brief 10 to 15-minute interview. And it has a very concise manual and workbook, which has been helpful for training raters to become reliable, especially, for clinical trials. And in conjunction with WCG, we also have professionally-developed training videos and gold standard ratings, that have now been used in over a dozen clinical trials, plus additional experimental psychology studies.

It's been translated into over 20

languages now. And the psychometric validation studies are strong. We have data supporting reliability in terms of inter-rater, internal consistency and test-retest reliability. And also, validity in terms of convergent and discriminant validity. And importantly, these good psychometric properties are replicated across the numerous translated versions of the scale.

Today I'm not going to spend much time reviewing those psychometric properties. What I'm going to do that I thought would be most helpful, is walk you through how the BNSS meets criteria for the FDA's eight COA criteria, A through H.

Criterion A, why should negative symptoms be assessed with clinical-interview-based rating scales? Jack already answered a lot of this, but I'll give you a couple of additional items. One is that the absence of an experience or behavior is harder to conceptualize than the presence of one. In

other words, you don't know what you don't have always. So sometimes it is helpful to have clinician judgment against what is normative. And this can be useful particularly in cases where people have less insight or awareness, perhaps due to cognitive impairment. [Second], traditionally, negative symptoms have not been measured through non-clinician collateral reports from relatives or caretakers, for example. Potentially, because these concepts are hard to understand, and also, people may not have complete access to all the information necessary, since several of the domains are not just based on behavior and observation, but also on experiential processes.

[Third], self-report questionnaires have been slow to be developed. I know of at least three that have been developed for negative symptoms. And they may not reliably assess or validly assess all five domains. For example, alogia and blunted affect which are based on observation, are very hard for people

to self-report on.

[Fourth], more recently, there's been a movement toward developing objective computerized behavioral tasks. For example, measures of reinforcement learning or effort-cost computation. And in my opinion, these are more intermediate phenotypes. They are things that are closer to mechanism than they are to clinical outcome.

So, clinical interviews hold a strong place in the literature still, and these are some of the reasons why.

So, does the BNSS assess all of the aspects of negative symptoms that are important? In relation to what clinicians and researchers have deemed core, from the 2005 NIMH consensus conference, clearly, yes. The scale was designed to assess the five domains according to modern conceptualizations.

But we've also, very recently, discovered through a qualitative study that we've done in my lab, that the BNSS also

captures on important aspects from the consumer and the relative/caretaker perspective.

I'll be showing you several slides from this study. And I'm going to focus on the patient data, but I'm glad to answer questions about the relative/caretaker data afterwards, if people have them.

What you can see here is we asked people a few things. We asked them, "do you agree with the definition we've provided?" And "should any of these not be a negative symptom?" And you can see with pretty high agreement, people considered the six domains assessed by the BNSS to meet their definition of negative symptoms, the way that we defined them. And they were defined in a more, slightly more colloquial way than what would be in the BNSS manual.

There was also little evidence that items or domains should be removed. You can see 27 percent for lack of normal distress, and 23 percent for alogia, which I think a lot of

clinicians, if they were to pick negative symptoms domains to alter, those would probably be the ones.

And we asked them, "well, how would you define these constructs, if they did define it differently than we do?" And the qualitative responses were interesting. For example, for alogia, people said things like, quiet, reserved, confused. For blunted affects: stiff, holding back. For avolition: dull, unfocused, not having it together. For asociality: shy. For lack of distress: tough, experienced, not caring, holding your own. For anhedonia: boredom, uninterested..

So many times, these are synonyms worded in a different way, in a more colloquial way than what Clinicians would use to describe the components of them. But there were some interesting additional facets that we gleaned from this.

We also asked them, "are there any additional negative symptoms that you think

exist in addition to the six we defined?" And the responses received were minimal, but the ones that came up multiple times were things like, apathy -- which by the way, I consider to be synonymous with the five domains. It's just developed in a different literature, mostly in neurology -- numbness, lack of energy, confusion or foggy thinking, and catatonia.

Well, we also asked people a number of questions and had them rate on an ordinal rating scale, going from zero, if not at all, 1 is slightly, two is moderately, and three is extremely, in relation to how important they thought the five domains were. So, their own subjective impression of the importance in several areas.

So, for example, "how important is it to you, to be doing well in each of these areas, or improve? How much does it bother it to have each of these negative symptoms?" And things related to functioning, "how much does having each of these make it difficult for you

to have a job, or go to school, or to socialize?" "How much does it keep you from having a good life?"

And the data -- I know these figures are small but forgive that -- the key take-home message is all of the domains except for, lack of normal distress, were rated as slightly to moderately important in these various ratings, related to functioning, quality of life, from the patient perspective. Lack of normal distress fell a little bit below that bar.

The higher domains were anhedonia, avolition and asociality. And blunted affect and alogia being a little bit lower, consistently. This raises a question in conjunction with the psychometric data we've seen on the lack of normal distress item on the BNSS, that have made us start to consider whether it's time to remove it, or at least, make it optional.

The intent was to create a proxy for the deficit syndrome because this has been the

primary means by which the deficit syndrome, or primary and enduring negative symptoms, have been studied in our field. And we've not been successful in doing so with that item, yet. So, maybe time to consider removing it.

Criterion C, do respondents understand the questions as intended by the measure developers? We basically asked people, "are these questions clear?" We gave them each of the questions and we asked them if they were clear. And if they said, no, we asked them to describe what they thought the question meant, in their own words. What were we trying to ask?

And you can see the percentage of agreement was very high. So, people with schizophrenia very clearly understood the probes on the BNSS.

They also gave us very useful feedback on what the sources of confusion were. So, we might be able to go back and refine some of the probes if we wanted to. And they even

gave us suggested changes to make.

Criterion D is, are scores influenced by processes that are not part of the negative symptom construct? Here there's a little bit more work to do. But in general, the answer is, no. Discriminant-validity correlations are low, in terms of secondary-negative-symptoms processes, such as positive symptoms, suppression, anxiety, et cetera. Item interpretation does not differ according to demographics. For example, we show measurement invariance across cultures and sex, in collaboration with Tony Ahmed.

And we have found support for the recall period use. There's an extensively reviewed and validated model in the basis affective science world, by Robinson and Clore that specifies why a one-week period for retrospective report, would be beneficial in this type of interview.

You essentially, get people to report on episodic memory or their actual

emotional experiences instead of semantic memory, or their beliefs about how they generally feel.

So, if you keep people within the bounds of their episodic memory, you're more likely to get accurate reports. And this is one of the reasons why the BNSS, and also, the CAINS uses that one-week timeframe.

Do fatigue or burden influence the assessment? Here we don't know. We've not done a formal tower ability study in relation to the trial. But given that the interview is brief, 10 to 15 minutes, we don't think so.

Does the motive assessment influence results? We've seen similar psychometrics between in-person and Zoom interviews. We did have higher ratings at the start of the pandemic. Those ended up normalizing, probably because the environment normalized again. But psychometric characteristics seemed pretty similar across modes of assessment.

And then expectation bias, this is a

future direction. We've not been able to formally deduce this yet.

Criterion E, is about scoring. Initially, we recommended a 2 Factor solution for the BNSS, based on the results of our initial exploratory factor analysis. But a few years later, we went ahead and we ran confirmatory-factor analysis and lo and behold one factor in our accepted two-factor solution offered a poor fit for the data.

And a five-factor hierarchical model was excellent. And interestingly, we found this not just for the BNSS, but also for other contemporary measures. Here you can see what the factors look like. Here on the left, you have the five domains. And on the right, you have the two-superordinate dimensions, the MAP (motivation and pleasure), and expressivity. Beneath which, you have the five lower-level domains.

And of course, we didn't believe it. So, we tried to replicate it. Here on Ahmed et

al, we found it across multiple studies and cultures. Six different cultures, multiple datasets, measurement invariance in thousands of people with schizophrenia in this study. We found it in both clinical high-risk and first episode patients. The same factor structure with five factors in the hierarchical model, with data from Hong Kong and America.

And then we also replicated it into our samples from America and Italy, using a different mathematical approach of network analysis in community detection. So, we've started to believe that this is probably the best structure for the BNSS, if not, for all negative symptom measures.

But there's a key question that Laura Swett brought up of, are all domains created equal? And here there's emerging evidence suggesting that they in fact, may not be. And here we find some evidence from our qualitative study that avolition may be deemed by consumers, to be more important than the

other domains. Here you can see an average across all of the different questions.

Here you can see avolition receives the highest rating, indicating that's the most important. And when you directly ask about importance itself, avolition is significantly higher than the others.

Now, interestingly, we've seen this more objectively too. We've run network analysis on the BNSS and we found that in people with schizophrenia, the domain that is most central is avolition and also, alogia.

Now, what does that mean? It basically, means that avolition, motivational deficits are driving the other symptom domains. They may have a causal connection and lead to a cascading effect of changes when motivational deficits are present.

We extended this in data from the phase IIb roluperidone trial, that you can see here, published in Schizophrenia Bulletin. And what we found is that compared to placebo, the

drug roluperidone was able to increase the centrality of avolition. And the extent to which it did so, dictated the magnitude of change in all the other domains. So, successfully treating avolition was key to the improvement of the entire negative symptom constellation.

And of course when their phase III trial was completed, we wanted to see if we could replicate this. So, here we saw there was a significant overall effect on negative symptoms this time, using the PANSS. And we used a different network analytic approach, called Network Intervention Analysis. And what this allows you to do, is basically isolate the treatment related effect, compared to placebo. And see which symptom is driving the overall improvement in the negative symptom constellation. And is that a direct or indirect effect?

What we found is that it was the PANSS item for emotional withdrawal, which is

the closest thing on there to avolition, rated on the newer negative symptom scales. So, we took this as being a nice replication of the earlier study. Interestingly, we ran follow-up analyses. Thank you, if any reviewers are in the room. That are in supplemental materials, where we basically, tried to see, are these roluperidone effects driven by secondary negative symptoms?

And they were not. They were not driven by depression, or positive symptoms. Interestingly, the drug did have a direct effect. You can see that redline there, from T to POS, on positive symptoms itself. Using this mathematical approach, it was able to reduce positive symptoms, even though it's, you know, not touted as an antipsychotic for that purpose.

Criterion F, do the scores correspond with specific health experiences that people with the illness have with regard to negative symptoms? Just like Jack

mentioned, same thing, correlations with EMA-based negative symptoms we found pretty consistently with measures in daily life.

Here again, do the scores correspond with individual experiences that the patients have? The answer is, yes. You can see that they indicated that all the domains are related to these important aspects of quality of life and functioning. And we've also demonstrated sensitivity to change. The BNSS has been shown to be sensitive to change in at least ten clinical trials. The majority of which were psychosocial, with only a few null findings.

And here you can see this busy table is from the roluperidone phase IIb trial. This is an anchor-based table, indicating the magnitude of change from baseline that's required to produce a CGI effect of a 1-point or a 2-point change in improvement. And the thing that I want you to pay attention to, is in the far-right column. What you see is effect size.

These are Cohen's d values and you can see they're in the medium to large range. So, a 1-point improvement on the CGI is producing a medium to large effect-size change across these various negative symptom domains. So, the BNSS is sensitive to change.

But as Dr. Swett mentioned, a critical question is, so what? Is that magnitude of change meaningful to people with the illness? Here the answer seems to be, yes, when I looked at this data I had available to me from the Roluperidone Trial. What you see in this table is data from our qualitative study. We essentially, asked people, we showed them the BNSS anchors, and we walked them through it and asked them, "where do you think you would rate yourself right now?" "Where would you want to be for your life to meaningfully improve?" And left that open to their interpretation of what meaningful meant.

What you see in this table is that the magnitude of improvement, difference

between the clinician rating on the BNSS, and the person's ideal rating for themselves was about the same magnitude of difference as what roluperidone produced on a 1-CGI-point change. So, 1-CGI-point change on this drug was equivalent to what people with the illness were saying, would be ideal for them to change as well.

In other words, roluperidone does seem to be improving negative symptoms to a level that's meaningful not just to clinicians, but to people with the illness.

So, as a quick summary, here you can see the eight COA fit-for purpose criteria. There is still some work to do on the BNSS. But there is some evidence for at least each of the criteria supporting that it meets the recommendations.

Now, in Part 2, I'm going to talk to you about digital phenotyping. And this involves the use of technology to measure symptoms in the real world, or from clinical

interviews. It's, typically, divided into active and passive approaches. Active, simply refers to something that people with the illness must initiate. Such as an ecological momentary survey on the phone, or an ambulatory video, or a cognitive test performed on the phone.

Passive methods in comparison are basically unobtrusive. They're collective in the background, usually through sensors of a smart band, or a smartphone, while people are going about their daily lives.

To start off, I'll show you here, the conclusion. We're probably about halfway there, in my opinion, in terms of what's needed for the assessments to meet the FDA COA criteria. But we've made a lot of rapid progress in a short period of time.

So, Criterion A, why should negative symptoms be assessed with digital phenotyping? As wonderful as the clinical rating instruments are, and I do of course think that they have a

very important purpose in the field, there are some limitations. One is subjectivity. There are social desirability effects. Cognitive impairments can influence the retrospective and prospective reports. Halo effects, cultural biases, there is also imprecision, right. These are made on a, usually it's zero to five, zero to six ordinal rating scale. So, the level of precision can be limited by that.

They also have lower resolution in terms of time and context, right. With digital phenotyping, you can get hundreds, if not thousands of data points per day. And you can drill down into the exact context in daily life that may matter the most to an individual symptom profile.

They may also be less sensitive to treatment effects. And often require very large n's for studies to be completed. Digital phenotyping offers you much more power and it's yet to be determined how much more cost effective it is, but it does have that

potential.

Here you can see an overview graphic of what some of the measures are. These are just a few examples from what we've been using. We've used smartphones and bands to measure things like accelerometry, which is the measure of movement. It's how much variability and magnitude of movement is there.

Geolocation, which is a measure developed from GPS coordinates. So, you nearly continuously monitor someone or get data each time they move a certain amount in space. And you can use that to map different variables of interest. Like percentage of time at home, or distance traveled from one time point to the next or number of location clusters.

We've also had measures of speech that are collected from the internal sensors of the phone. So, rather than directly recording people's speech and samples from the ambient noise, what we've done is we have a program that can basically, on the fly, calculate a

value.

So, it's not recording a sample that's stored. On the fly, it's calculating whether human speech is present in the background and the level of intimation of the voice of that speech, and spitting out the variable that gets saved automatically into a dataset.

We've also paired that with EMA surveys, where we ask about location, activity context, social context, interest, pleasure et cetera. And have people perform an ambulatory video at the end of the survey. So, for example, holding the phone up in front of them, and responding to a probe, such as “tell me what you did over the last hour” or ”How you felt, and about your symptoms.”

We can then decode that later for various measures of interest, like facial affect intensity and frequency and variability. Acoustic measures related to intonation in the voice, and pitch. And all sorts of other

acoustic properties.

You can also interestingly, take data from your old video recorded interviews, and plug it into software that has automated algorithms for things like facial and vocal affect and acoustic properties.

Here's an example of one measure that we think is really promising from a collaborator, Alex Cohen of LSU, called speech latency. This is essentially, a measure of verbal response time, as an objective marker of a number of elements of psychopathology. This might be like a g-factor for cognition, if you will, that relates to negative symptoms more broadly.

So, here for example, this is a fake video of me with one of my grad students. But you ask them a question on the BNSS, something like who did you spend time with this week? And you look at the pause after the interviewer stops their question, and the time it takes the participant to start their response. "I saw my

uncle and two cousins this week." "Oh, how often did you see them?" Pause, "Only on Monday." "What did you do when you got together?" "We watched Georgia beat LSU in football," right. So, the term latency we know is a critical predictor of negative symptom response. And as I'll show you later, is a critical predictor of sensitivity to change.

Criterion B, do these assessments capture all important aspects of the negative symptom construct? They have an inherent face validity. So, these are measures of movement collected in daily life, behavior collected in daily life, social activity, emotional expressivity, all collected in a more ecologically valid way. So, they do have an inherent face validity and a ground truthiness, if you will to them.

But the digital phenotyping measures are, generally, modestly correlated with clinical rating scales. And you can see an example from data from our lab here, that our

values tend to be about .3/.4 between BNSS, anhedonia and avolition and asociality with their corresponding EMA survey measures.

The magnitude of correlations with passive measures is pretty similar, usually about .3 to .6. But this is really quite a bit lower than what you would expect. For example, if you had two clinical rating scales that you were looking at correlations with, you'd expect something like .8 or higher.

And there are a number of reasons for this. One is methods variance. You would not expect the correlations to be as high as they would with the clinical rating scale, due to methods experience. But also, temporal resolution and context, right. How far do you zoom in for these passive measures? Do you go into every hour? Do you average every minute, every second? Do you average across one week to make them comparable to clinical ratings scales? These are issues that the field is still grappling with.

In our self-perceptions qualitative study, we also asked people about digital phenotyping measures. We asked them for example, "do you think that these measures are relevant for measuring negative symptoms?" "Can you measure negative symptoms through smartphones and smart bands?" And the responses on average were, either around moderately, or in between slightly and moderately.

And we asked them about things like geolocation, accelerometry, ambient speech, the ambulatory videos, turn latency, and the EMA surveys. So, we had a decent, but not overwhelming, list for them to evaluate. And we gave them descriptions of what each one was, so that they knew what we were actually talking about.

We also asked them, "How do you think your life would change if you were to improve on these types of measures?" And use the same rating scales before. And what you

see is that they did consider these things to be important. They thought that if they did show changes on these digital phenotyping measures, that their functioning would improve; that their quality of life would improve, at least slightly or moderately so.

Now, one thing that people often wonder about these digital phenotyping measures, is "what concerns do people have when using them? What concerns particularly do people with schizophrenia have about them?" And what we did, is we asked them, "Do you have concerns about using any of the technology described, to measure your negative symptoms?" And you can see for each of the measures, the responses were below, slightly. So, there was minimal concern. But when they did have a concern, we had them "Tell us about them." And most consistently what they asked about, what they said was, being audio recorded. They were afraid that they would be recorded all the time. They thought it should only be when

asked. That it could be invasive if frequent. Not being recorded, not recording others unknowingly.

And, essentially, we found through this and other studies, when you explain to people that it doesn't do that, they're much more comfortable with the technology. But there are apps out there that people are using, that do record actual samples. So, that's something to consider.

They also, worried about whether the data obtained would be natural or forced. Had questions about the privacy and security of the data. People didn't like the idea of being dependent on using a phone or device throughout their lives, if their clinicians wanted to measure their symptoms.

They worried that the smart bands might be glitchy or uncomfortable. Whether they would be reminded of having the illness by completing the EMA surveys and whether that would make them more ill. And some people

worried about being tracked continuously by GPS. They wanted to be able to turn on and off, which some apps do allow that.

We also, asked them whether the questions we were providing in the EMA surveys were clear or not. And here the response, 94 percent of the time, was that they were clear. And we had them describe to us what they thought we were asking about. How to make it better in instances where they thought it was not clear and did receive some helpful feedback.

Now, are the scores of digital phenotyping influenced by processes that are not part of the negative symptom construct? Here discriminant validity tends to be good. The correlations are generally low with measures not part of the negative symptom construct, either in terms of clinical ratings scales, or concurrently collected EMA. For example, measures of depression, anxiety, or positive symptoms. But I will say that they're

not quite as good as the clinical ratings scales. So, the correlations are a little bit higher than what you typically see there.

Does item interpretation differ according to demographics? We don't know. There's not been a lot of work done on this. And we need to do some work on measurement and variance. But one of the advantages of digital phenotyping is that you don't need to necessarily make a retrospective report.

So, you can completely eliminate the need for recall by asking people how they feel in the moment. You know, their activities in the moment. Are the scores influenced by processes, not per the construct? Well, what about fatigue or burden?

We studied this systematically, and throughout it all we did study on adherence and tolerability for EMA and paths of digital phenotyping. What we found is that people with schizophrenia and healthy controls, found it highly tolerable. They rated the experiences

highly positive and not very negative.

But we did find that adherence dropped during one week, not just for EMA surveys, but even for some of the passive measures, and wearing the band. And so, this is something where we do need longer term studies. We've only done this for a week, but we need more methodological studies to track it long-term.

But importantly, negative symptoms were not the predictors of lack of adherence. It was things like age, being busy, like whether people had children or not, right. These were some of the things. Life being disruptive rather than not having motivation.

If anything, the patients with higher negative symptoms we found were more likely to be compliant, because they were sitting at home and life was not as disruptive to performing the activities on the phone.

Does motive assessment influence results? Here there's a lot of work to be

done. We know that incentives probably matter. How much money you give people in the EMA survey for example. Do you pay them for doing the passive data collection? It might influence whether they wear the band or keep the phone on them. Are you providing a phone or having them use their own phone, right. We don't know how much that matters yet.

How much do different operating systems matter, right? We know that Android versus Apple, collected different parameters. They allow you to collect some things on Android, that cannot be collected with Apple, right. Methods variance may also matter. Here's an example of this. We did an accelerometry study where we had people wear a band and also, collected accelerometry through their phones.

They were supposed to have both of them on, concurrently. And we could tell when they weren't wearing the band. What we found was that the phone was able to differentiate

people with schizophrenia and controls. They were group differences, but the band could not.

In contrast, the band had correlations with negative symptoms, measured through the BNSS, whereas, the phone did not. So, there are discrepancies, both in terms of the magnitude of group impairment and the connection with negative symptoms. So, modality or mode of measurement may matter here.

We don't know yet about expectation bias, and I think this is critical to study. So, we need long-term studies to be done. And we need to do a lot of work on scoring, right. So, what is the right level of temporal and spatial resolution? How far do you zoom out or zoom in? How do you combine EMA and passive measures to drill down into the situations of greatest interest?

So, for example, do you only get VOX or speech measure during instances where they report having been in a social interaction?

Also, do we need normative data on healthy controls in people with schizophrenia to facilitate the interpretation of this type of data by consumers and clinicians alike?

Would you know what a change in three meters, from one time point to the next, meant? Of course not. Or a certain change in accelerometry? We need norms to put these into more interpretable value sets, similar to neuropsychological tests.

We have interestingly, Tony Ahmed and I have been playing around with some of the digital phenotyping data we've collected. And we have over 100 people at this point, and we wanted to see, could we include measures we thought should be relevant to each of the five domains, and find either a two-factor or five-factor solution?

Here, what we found is again, one and two-factor solutions were not a great fit. And the five-factor was not quite as good as the clinical rating scales, but it was the most

optimal. I suspect that there will be the more measures you put in, there will be more and more factors that will emerge.

We also, again, asked, you know, do these scores correspond to specific health experiences? And remember, patients thought they related to quality of life and functioning, so they do think that digital phenotyping is relevant to their health experiences. And there has been some evidence for sensitivity to change already.

One way to measure this is through context effects. So, here you can see EMA data on anhedonia. That is what's on the Y-axis. You can see that in some activity types, in some locations, people with schizophrenia do and don't have anhedonia.

For example, when at a family member's home, they do not have anhedonia. When in public, they do not. But when they're at their own home, they do. When they're out running errands, they don't have anhedonia, but

when they're engaged in a recreational activity or eating, they do.

So, context, this is one way of measuring sensitivity to change across various activities and locations. You can do the same thing, pairing the active and passive data.

Here we show that both people with schizophrenia and healthy controls have more social activity identified through the speakers of the cellphone with our VOX measure, when they self-report being in a social interaction. So, that's helping to validate that particular ambulatory measure.

This is data from Bill Horan, Phil Harvey and colleagues at Karuna. You saw a little bit of this earlier in their open-label 12-month study of KarXT. Here they show sensitivity to change in terms of improvements, in terms of decreases in unproductive activities, increases in productive activities, spending less time at home. So, their drug, KarXT was able to cause these changes. And EMA

was sensitive to be able to pick them up, using these survey-based measures.

This is data from Alex Cohen, from the phase III trial of brilaroxazine. And what they did, is they took over 2,000 audio clips, audio recordings from clinical interviews and Alex processed them for certain acoustic and speech variables. One was turn latency, that I mentioned to you earlier.

He was, essentially, able to segment out participants who had a certain magnitude of turn latency deficit, which they termed "vocal biomarker positive," and compared that to people who did not have this vocal biomarker. And what they found is that the originally negative results of the trial, null results, became positive when you identified, when you stratified patients based on this digital-phenotyping-based biomarker.

So one, an alternate way to use these methods is as an enrichment tool to identify your biomarker of interest that you

think should separate people on negative symptoms.

And I will stop here on this slide and just conclude by saying, again, a bit more work needs to be done, but it is a very promising measure. And I think the field has a lot of excitement on using these digital phenotyping measures, especially in conjunction with clinical rating scales and other measures.

Thank you, and I'd like to acknowledge NIH and my team and collaborators who helped conduct this work. Thank you.

DR. WEHRING: All right. Thank you so much, Dr. Strauss and Dr. Blanchard. With our time remaining, I'm really excited to turn it over to our respondents. Let Dr. Blanchard and Dr. Strauss rest for a moment. And I'd like to just kind of go down the line, maybe starting with you, Dr. Horan, and get your reflections, response based on your area of expertise, about the discussion they started.

And also, if anything stuck out to

you as something that was really a poignant point, or something that was missing in our discussion, to kind of help move us forward. Oh, and please introduce yourself. I think I forgot to say that. Thanks.

DR. HORAN: Thanks, Bill Horan with the EMS in UCLA. I'll mention two things. Number one, I had never seen all that qualitative data before, that Greg Strauss collected. And that's, it's really impressive and really encouraging. Particularly for things like perceived participant importance of avolition, and how that relates to functioning and daily life. Also, acceptance of things, like EMA measures.

My second point is, I continue to be impressed by the rapidly-growing data on EMA, particularly as a measure of avolition symptoms. It's really those avolition symptoms that seem to be most strongly related to functioning in daily life. Participants perceive them as very important.

And with EMA, in the moment, you can just ask people, "What are you doing?" "Is it a productive activity, or is it passive?" "Where are you? Are you at home or are you out of the home?" "Who are you with? Are you with other people, or are you alone?" And you can also ask about their emotion. "Are you feeling happy? Are you feeling sad? Are you feeling content?" You can collect all that information in the moment. And those are all direct behavioral outputs, correlates of avolition. We're getting to the point where these are not just things that are being done in sort of small academic studies, but some of those studies are involving hundreds of patients, in large clinical trials.

And finding reasonable associations, correspondents with things like the NSA, or the personal and social performance scale. And then even sometimes seeing them converge with measures like, steps taken, or activity levels. So, I think it's getting closer and closer to

the time where maybe we can start working with regulators to understand what it will take to validate these measures for use in clinical trials, as endpoints. What should be the gold standards? Is it going to be the PANSS' negative symptom factor that we need to find correlations with, or do we need to think of other things? If it is something like the BNSS or the PANSS' negative symptom factor, what does the correlation need to be? It's, you know, as Greg was showing, it may not be exactly what we're used to, using correlations between clinical rating scales. So, we seem to be getting to that point, where maybe we can start working toward using these as endpoints. Thanks.

DR. KIRKPATRICK: Greg, forgive me, the brilaroxazine was significant for positive and negative symptoms. The data you showed was correct about what happens with vocal. You may be thinking about another study, where turn latency made something significant that was not

previously significant.

We have, you know, those of us interested in digital phenotyping, have kind of a story, we have a pitch, which is, "Oh, my gosh, these clinical rating scales are subject to all kinds of weaknesses, recall, rater bias, patient's willingness to tell you what's really going on." We say about the digital phenotyping, "This is great. I mean look at all these wonderful ways that they're so much bigger than the scales." They say, yes. You'll say yes, but how they valid? "Well, they predict clinical rating scales." That's it? That's embarrassing. And what's interesting is the correlations, as you pointed it out, tend to be 0.33, maybe 0.5. So, they're clearly not redundant. So, then the question becomes, is one of them better than the other?

Well, maybe one kind of scale -- well, it's not necessarily the case the digital phenotyping measures are better. There may be something very complicated that human beings

are doing that we can't articulate and can't teach a machine to do, possibly. Maybe one kind of approach is better for certain things. And the other one is better for other things. We simply don't know.

So, we need to get out of this circular reasoning of, "This is bad, this is better because it is as good as what's bad." And a way to do that, I think, is to go to other sources of information. And, preferably, a basket, because every other source of information about assessing someone's function is going to be flawed as well.

So, I think it would be things like, patient report, informant report, families or people who know someone very well. I took it that you were not big on functional capacity, but I think as part of the basket that's a reasonable contributor. And then another one is sensitivity to change, including sensitivity to treatment effect, of things that we know already work.

So, another piece of that basket might be independent verification of key things. So, for instance, a clinical trial or some other research, a staff member could call and say, and ask if this person, in fact, lives independently. To what extent does he or she take care of all activities of daily living, including going to the grocery store? Or do they, in fact, have an intimate relationship with someone? Things like that. So, and then maybe we can get out of this circle -- kind of embarrassing circular reasoning.

Another thing about the latency measure, it's interesting because of the background of it. There is fairly extensive literature in the psychological literature. a lot of research on that, outside of the context of clinical trials, suggesting that, you know, we already know a fair amount about what turn latency may mean.

So, you had kind of an a priori reason to go in and look at it. And that

includes looking at the effect of cognitive load in normals, compared to people with a broad range of serious mental illnesses. The other thing that's nice about it is that it's clinically interpretable. It is a kind of psychomotor retardation. And that makes good clinical sense. And so, to quote Mark Opler, who I think just left, so I can claim he said whatever I want to say, he claimed. But he has frequently said these numbers are arbitrary. But his point is really excellent. There are a thousand digital phenotyping measures. There are 12 that matter.

If you go in and you do machine learning, and it's unguided in any way, with any kind of a priori thought, you're going to get wonderful predictions. And they're not going to replicate, because they're not part of the 12. And I think to get to the 12, as soon as possible, we need to have some a priori basis for doing so.

Last, but not least, can we have a

moment of silence for Item 4, lack of normal distress?

(Laughter.)

DR. KIRKPATRICK: Not a moment of laughter, I was asking for silence and respect. All right. You know, I think there is a phenomenon there

DR. WEHRING: Thanks so much, Dr. Kirkpatrick. Dr. Marder.

DR. MARDER: Yes. Steve Marder from UCLA. You know as I look at this session, which is really outstanding, there are two themes. One is, we've made real progress in developing two very good clinical assessment instruments that are based on clinician ratings, the BNSS and the CAINS. Both of them are substantial advances.

The problem is, as somebody who does a lot of these ratings, particularly with the CAINS, one can see that there are limitations to clinical ratings that really need to be addressed. And Brian and others, and Greg,

referred to them. It's, people can recall the severity of their hallucinations, at least I believe they can do it relatively well.

I don't have great confidence having done these ratings, that they could really recall with adequate precision, how motivated they were to work during the past week. I'm not - that motivation varies. The time of day that you do a clinical assessment, may not be the time of day that they can get the best information about that person's motivation.

And things like social interest, really depends upon context that the person is in. And then when you put them in this situation, of a clinical interview, everything gets changed. So, what I would like to do is to sort of reframe the question of what are the limitations of our clinic-based assessment instruments that could be addressed with new technology. The problem is there's chaos out there, with multiple different measures. And I don't think we should talk so much about what

we can measure, because we could measure a lot of things. But -- how can we take the limitations of our current instruments, and improve precision using digital instruments?

And how can we reach as a field, and sort of help FDA decide what the best kind of multimodal instrument is? And I think that's a problem that I don't think we have a plan to resolve. But I can say how to resolve -- I mean, I think we would know how to do it. It would require sort of groups of people evaluating the current instruments, developing clinical trials, perhaps using industry settings in order to compare them in real clinical trials.

And I think there are methods for moving the field forward from what's now, the methods chosen are too often dependent upon the salesmanship of the person who developed the measure, then the actual quality of the measure. And its ability to measure what it says, with precision. So, I think we have a

problem about how to make our assessments of negative symptoms more precise. But I think there are probably strategies for addressing it, if we move in that direction. I'll stop there.

DR. WEHRING: Thanks so much, Dr. Marder. Dr. Ahmed. Thanks.

DR. AHMED: Yes. Hi. Anthony Ahmed, Weill Cornell Medicine. It's a real honor to be part of this session with a lot of people like Bill and Steve and Greg and everybody else, and Jack, who was my honors graduate school professor. So, it's a real honor to be having this discussion with you.

I'm also an inpatient psychologist. I work with patients with schizophrenia in an inpatient psychiatric rehabilitation facility. In that context, I want to make one clinical point, and then I'm going to circle back to the psychometric issues that we're struggling with.

In this program, this is a psychiatric program where we provide a lot of

evidence-based psychosocial rehabilitative interventions, skills training interventions, like cognitive remediation and social skills training and everything.

But programs like this are sort of going out of the market, primarily because there's a lot of pressure in terms of the length of stay of such programs. This is probably one of maybe two or three tertiary psych rehab facilities.

Who are the patients that we see in this program? Well, these are patients who struggle with the most functional deficits, with the most skills deficits. Others are patients that struggle with engaging in the community and staying in the community and maintaining community tenure, because, well, they don't have the skills. Most of these patients are relatively symptomatically stable when you think about processing those negative symptoms and emotional distress. But they are struggling with engaging in their community and

being part of the community.

A lot of the data that was shown earlier today about the prevalence of patients with permanent negative symptoms, like, yeah, you'll getting rates of about maybe 20 percent, maybe just under 20 percent. But in settings like ours, those rates are even higher. And those are the patients that we struggle with the most. Other patients, like, do relatively well, for the most part, when we discharge them or we put them through our rehabilitative interventions.

But the problem is we don't have enough time for those patients who struggle with these persistent negative symptoms and permanent negative symptoms, to really gain within that length of stay period.

Now, programs like ours are continually under pressure to continue to reduce our length of stays. And what that means is that, you know, like, well, our rehab -- and I really appreciate Jack's slide about,

you know, all of the behavioral health psychosocial interventions that show improvement, or contribute to improvements in negative symptoms. But if you look at the length of those interventions, the length is quite long. We're talking about 8 weeks, 12 weeks, 10 weeks, 20 weeks. We don't have that kind of time, you know, in rehab programs, okay?

If we don't come up with something that can support the work that we're doing in behavioral rehabilitation, we will see the last of the programs like the program called the Second Chance Program. So we need -- there is a little bit of urgency here.

Now, to the psychometric points. You know, I think the information is clearly a theme. But I do want to make the point that -- and certainly in the context of some of the data that we published on the two-factor and the five-factor models -- you know, the hierarchical model that we really settled on,

that we think has a lot of, you know, like, validity for predicting cognition, for predicting function, for predicting everything, you know, like, you know, we really think -- I think the conclusion that we're drawing from it is that the two-factors are important, okay? And we think that that's a good place to start . And we also think that the five-factors are important, but we don't think that you need to pick one against the other, necessarily.

Now, there are domain-specific effects that were seen, especially given the impact of avolition. You know, but I think you still need to collect, you know, administer the comprehensive scales. You know, I think, you know, the BNSS and the CAINS certainly, you know, ideal if you have those. But we can still capture the motivation and pleasure and emotional experience factors even with the PANSS. You know, we have data that we're analyzing. And, in fact, a few have actually published the two-factor approach to the

matter, the seven items in the PANSS, and even in clinical trials data was seen that, you know, those still capture, you know, to a certain degree, the scope of those two factors.

And, as you can see in Greg's data, looking at the impact of avolition is very central, that there is enough there to show some process in terms of how this network of symptoms change in the context of our treatment studies.

So, the individual domains, like, important those granular domains, but we don't also want to forget about the forest itself, the forest of the scope of symptoms. And so we should continue to collect that data whenever possible.

One final point has to do with, like, how the scope of work that Jack Blanchard and Greg and Brian Kirkpatrick have done with getting translations of the CAINS and the BNSS. And one thing that's clearly missing, and I think it's pretty obvious to everyone, is that

there are no translations in languages from the African continent. I think that goes to collaborations that we need to establish with researchers in South Africa and Nigeria and some of those countries where their academic institutions are well-established.

And, yeah, and I'm going to stop there. Pass it on to David.

DR. WEHRING: Thanks so much, Dr. Ahmed. And, Dr. Reasner, we let you anchor this. We thought it would be nice to kind of come full circle to address some of this from the regulatory perspective. So, you know, no pressure. Thanks.

DR. REASNER: Okay. Then I'll start with my disclaimer. No, actually, I wanted to say that these comments are really from the measurement lens, because that's my role in the Division of Clinical Outcome Assessment. And regulatory decisions are a multi-disciplinary process.

Also, I think some of these comments

are really about the really exciting ideas that I heard today. That these are about research hypotheses, research designs, and I have to think a little bit more about the implications for us sponsoring late phase. And how this might change, because we sort of have the short-term.

We have the tools we have today. We have important ongoing programs. And then we have sort of our aspirations about where we might go, in terms of the tools. And what sorts of differentiated profiles we might be able to recognize in the future. So, those are not in the same timeframe. Although Mark warned us this all could change very quickly.

So, one thing I wanted to just mention in terms of the validity-evidence table, which you very kindly used in your presentation. In Item B, which is about capturing all important aspects of the construct, in this case, the NSS construct. I just want to say at least from my perspective

as an individual reviewer, it doesn't all have to be in one assessment, in one tool.

I think sometimes we do need a primary endpoint, in order to declare a positive trial and move forward past the inferential gate, to talk about the secondary endpoints, and even exploratory endpoints. But I think that if the measurement strategy is comprehensive, that's really, I think, the primary interest. Because there may be an older tool that omits certain concepts that are now recognized. It doesn't mean it's not useful. But maybe you back fill those, supplemental concepts in a secondary or exploratory assessment, or endpoint.

So, I think, think about assessing a patient completely, but build your endpoint hierarchy in a way that's practical with statistical power, right. Maybe putting reliable endpoints higher in your hierarchy.

And on a few points, and I want to blend a few different questions, we had some

questions from folks online, and of course some questions came up earlier, and some questions have been posed in this session, regarding the digital health technology, I will say that in this context, it seems to have sort of great potential. And so, what I would point out, and some of this work is already ongoing, I'm not conversed in that health literature. But I would say things such as the concepts underlying blunted affect, seem really amenable to a digital health technology.

And some of that, I think, would enable a comprehensive assessment of the patient by combining different recorders. You have your digital recorder, maybe for blunted affect. You have your observer who can observe behaviors that the observers can observe. And you have an, you know, your learned clinicians and that context. But that can move your assessment into different places.

And, you know, some of the digital health technology can be applied in clinic, and

some can be out, you know, free ranging. And this idea of that, you can also probe, either through EMA, so it could be random phone calls, or activity triggered phone calls, or something like geo-mapping, you know. I think that's very powerful and don't usually use the example of dyspnea.

So, your phone can tell you when you've just walked up the stairs. And if you got a phone call, and someone called you to rate your breathlessness, that would be a very powerful way of looking at patient symptoms. So, similarly in our context in schizophrenia, I think that could be very powerful.

So, yes, it's complicated to apply the regulations and guidance to digital health technology, but much of the work that you all are familiar with about, like reliability, you know, will apply to these digital technology endpoints as well. So, you can apply your experience. And I think that that has great potential.

Oh, and I wanted to say that in terms of the multicultural aspect, they may also remove some of the cultural and demographic variables from the room. And that could be very interesting. I'd love to see those data. That would be, I think, helpful.

There was a little discussion earlier about, you know, is it worth looking at individual items and whatnot. And I think the answer from a research perspective, again, I don't how to bring it up to an endpoint that has alpha control that you present to a regulator. Now, that might take some thought. But I think, the individual items will be of some interest.

And an example that I thought of, was within anhedonia. So, we know that the profile on anhedonia differs between different indications. And so, you might have patients with comorbid conditions, like depression or undiagnosed comorbid conditions, or prodromal syndromes, right.

Then you could use individual items like a few items like anticipatory anhedonia and a few items of consummatory anhedonia. You can profile those patients and you could either exclude them from your trial, or you can recruit them. But maybe they're not in your primary analysis or maybe they're in a stratum.

But anyhow, understanding that that profile is different at the individual item level, not necessarily the main domains that the developers anticipated, but maybe informed by today's research. I think that would be, you know, something that's worth doing. And, you know, you should continue to look at those things.

Just one thing about, I guess there's just a natural tension that came up a couple of times, I would mention. You know, we want to reduce variability. We want to maintain generalized ability, and it matters what phase of development you're in, right.

So, asking a question in phase II,

and which I think was discussed in an earlier session, I think is, "is the ideal place?" And there's not a lot of room to provide advice on what assessments to use, or how to score them, or how to construct endpoints. If the first conversation is, you know, shortly before your investigative meeting, and you're planning those things.

So, I think building that conversation -- I'll say it, right -- early and often, you know, is really helpful. And the Division of Clinical Outcome Assessment actually do a lot of consultation at pre-IND stage. And I think that's very helpful. Because there's a lot of investment in the program, and the patients are waiting.

And then along that line, in terms of sources of variability, so much changes when the patient enters the trial, right. You have sort of physician/patient alliance. You have the beliefs about treatment. You have apparently different rates of rescue, used by

site. A lot is going on. So, I think that it's, you know, important to recognize that you're going to see those shifts.

And if you can account for those variables, maybe to put on my statistician's hat for a minute, you know, stratification, a lot gets put into the site effect. So, maybe rescue methods end up in the site effects, not explored.

Also, in terms of this idea of providing sort of a psycho-behavioral package for every patient. Consider how patients are treated across sites and whether those are equivalent.

I know if you can standardize, maybe that's a little too aspirational. But I've often thought that we're neglecting a little bit, making sure the patient experience is consistent across sites. So, that's a few preliminary thoughts. And thank you for your time and the invitation.

DR. WEHRING: Thanks so much, Dr.

Reasner. Do we have time for a question, before closing remarks? Great.

Dr. Campbell, do you have some online participants?

DR. CAMPBELL: Yes, we have some, a couple online questions. So, I want to ask two questions. I think they're kind of important. The first one is "Given that the PANSS recognize the importance of informant data, why was informant data not included in the CAINS?"

"Are there problems with the CAINS ability to identify negative symptoms in clinical trials as a result of the lack of informant data?"

DR. BLANCHARD: So that decision I think was just driven about something that was raised earlier, is the availability of collateral informants is challenging. And the data that we've collected from direct report from a participant, the accumulated validity data would indicate that, it's not a problem.

Could you enhance certain aspects of

the assessment, if you included those other assessments? Possibly. But the evidence that we have now, with a clinician rating based on just the participant, is pretty compelling.

So, I don't see it as a problem for including it in clinical trials.

DR. CAMPBELL: And one other question that I think is important when we think about leveraging prior data and wanting to pool data to help us better inform trial development going forward.

And the question is, "How shall we interpret previous clinical trial results, where generalized, a more general schizophrenia patient population was recruited but not predominant negative symptoms. And they used the PANSS negative symptom scales, was used in a short trial, and they claimed the drug benefits while they may not have been a more predominately negative symptom group. So, how can we leverage that prior data to help inform, going forward?"

DR. STRAUSS: I can make one response here. I think we should try re-analyzing data, using cut point criteria. Similar to what some of the more recent trials have done. So, separating out patients, identify the subgroup that meet the unique predominate or persistent subtypes. And then just analyze it that way and see if the results hold.

DR. WEHRING: All right. Any questions from the audience? We're almost out of time, so last chance.

Well, you guys were so clear that we don't have any more questions. Thank you all so much for your participation and for the outstanding discussion. Really appreciate it.

DR. FISCHER: All right. So, I did budget some time in the schedule for a wrap-up, and the wrap-up is also a cushion, because I knew we were going to go over. So, I had initially planned on taking some notes and PowerPoint, and then having some slides. Kind

of summarizing the day and all of talks. But there's way too many slides. There's way too much and I don't know if my notes would make sense. So, I'm just going to wing here.

But I really want to first of all thank all of the participants, all of the panelists, all of the speakers for today. Everybody who joined us virtually, people who submitted questions, I think this day was very successful. I'm not judging the day by whether we have all the answers we need, because of course, we still don't have all the answers we need.

But I think we've identified a lot of the questions that we need to answer, which is a very important part of this day. In the past, there have been a number of academic communities and collaborations taking a look at negative symptoms, and how best to define the group? How best to do clinical trials?

But I think even though the FDA had been participants in some of those past

activities, this is the first time the FDA is actually initiating a statement about this, about negative symptoms, and treatment development. And looking forward to answering some of these questions from a regulatory perspective.

So, I think that this was really important. Some of the take-home points that I want to emphasize are that it's great to hear about the lived experience from people who have schizophrenia. And of course, with what Michelle had said earlier, with our patients' listening sessions, and with some of the things that we're trying to do to figure out if a drug has a clinical, meaningfulness to people. We want to hear from, directly from patients, from caregivers. We want to know what makes a difference in people's lives.

I think it was great to hear about the neurobiology and how we're starting to see the interaction. That we can't look at negative symptoms really in isolation, that

they're really part of a constellation of symptoms. And we want to look at them in conjunction with things, like cognitive impairment in schizophrenia.

We definitely have some problems that we need to address with study design. But I think there are also some things with study design that we've figured out. We really need to think about who to enroll in studies, how to design the studies, things like, active comparators, placebo controls, some things to think about there. But how to best design a program to show an effect?

When it comes to scales, some of the important cultural considerations, you know, not just back translating a scale, but actually in addition to having a scale that is culturally sensitive, having a rater that's culturally sensitive as well.

So, they can kind of interpret things and rate things with that in the back of their mind, to figure out whether the things

that they're seeing should be considered negative symptoms or should be considered part of the culture of the person that they're interviewing.

I think it's also important to figure out where we have the interface, between clinical ratings, and digital phenotypes. So, we had some great presentations today to hear about some of these new exciting scales that were developed because of the, I guess, the things we noticed with some of the older scales, that could be improved upon.

But now, we're kind of in this brave new world of digital phenotyping and machine learning and EMA. And I think that it's not going to be an either-or question. It's going to be how do we integrate these two concepts together, to come up with the best way to measure negative symptoms. So, I think these are exciting times to figure out how these digital phenotypes are going to inform our clinical ratings.

So, finally, I just wanted to mention that for all the people who asked questions virtually, that we didn't get to today, those questions are very important to us. Because it also shows what people are interested in, who have attended this session. And so, those are going to factor in to our thoughts, moving the meeting forward, thinking about things.

As far as a product from the meeting, we are going to post the slides on the website. And we are going to have some kind of summary. We haven't decided yet whether this is going to be a journal article or a white paper.

Maybe even a guidance for industry from FDA. So, we're still in internal talks about what the work product will be from this. But there will be a product that will summarize the meeting for people.

So, I just want to thank everybody again, for their attendance. And just remind

everybody who has luggage in the back to just make sure that it's your suitcase, because many of them look similar. So, again, thanks for coming and safe travels home.

(Whereupon, the above-entitled matter went off the record at 3:54 p.m.)

E

F

O

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: Evaluating the Negative Symptoms of Schizophrenia in Clinical Trials

Before: US FDA

Date: 08-16-24

Place: Silver Spring, MD

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate complete record of the proceedings.

Court Reporter