FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH

JOINT MEETING OF THE PSYCHOPHARMACOLOGIC DRUGS AND
DRUG AND RISK MANAGEMENT ADVISORY COMMITTEES
(PDAC and DSaRM)

Thursday, November 1, 2018
8:00 a.m. to 5:01 p.m.

FDA White Oak Campus
Building 31, the Great Room
10903 New Hampshire Avenue
Silver Spring, Maryland
Meeting Roster

DESIGNATED FEDERAL OFFICER (Non-Voting)

Kalyani Bhatt, BS, MS
Division of Advisory Committee and Consultant Management
Office of Executive Programs, CDER, FDA

PSYCHOPHARMACOLOGIC DRUGS ADVISORY COMMITTEE
MEMBERS (Voting)

Walter S. Dunn, MD, PhD
Staff Psychiatrist and Assistant Clinical Professor
West Los Angeles Veterans Administration Medical Center
University of California Los Angeles
Department of Psychiatry
Los Angeles, California
Jess G. Fiedorowicz, MD, PhD
Associate Professor
Departments of Psychiatry, Epidemiology and Internal Medicine
University of Iowa Carver College of Medicine
Iowa City, Iowa

Satish Iyengar, PhD
Chair and Professor of Statistics
Department of Statistics
University of Pittsburgh
Pittsburgh, Pennsylvania

Felipe A. Jain, MD
Assistant Clinical Professor of Psychiatry
Department of Psychiatry
University of California, San Francisco
401 Parnassus Avenue, Langley Porter
San Francisco, California
Rajesh Narendran, MD

(Chairperson)

Attending Psychiatrist

Re:solve Crisis Network

Western Psychiatric Institute and Clinics

Associate Professor in Radiology and Psychiatry

Psychiatric Molecular Imaging Program

University of Pittsburgh

Pittsburgh, Pennsylvania

Kim O. Wictzak

(Consumer Representative)

Co-Founder, Executive Director

Woodymatters

Minneapolis, Minnesota
PSYCHOPHARMACOLOGIC DRUGS ADVISORY COMMITTEE MEMBER

(Non-Voting)

Robert R. Conley, MD

(Industry Representative)
Global Development Leader
Pain and Core Therapeutic
Team and Distinguished Scholar
Eli Lilly and Company
Lilly Corporate Center
Indianapolis, Indiana

DRUG SAFETY AND RISK MANAGEMENT ADVISORY COMMITTEE

MEMBERS (Voting)

Kelly Besco, PharmD, FISMP, CPPS
Health-System Medication Safety Officer
OhioHealth Pharmacy Services
Dublin, Ohio
Marie R. Griffin, MD, MPH
Professor, Health Policy and Medicine
Director, Vanderbilt MPH Program
Department of Health Policy
Vanderbilt University Medical Center
Nashville, Tennessee

Laurel A. Habel, MPH, PhD
Associate Director, Cancer Research
Division of Research
Kaiser Permanente Northern California
Oakland, California

Sonia Hernandez-Diaz, MD, MPH, DrPH
Professor of Epidemiology
Department of Epidemiology
Harvard T.H. Chan School of Public Health
Boston, Massachusetts
Martin Kulldorff, PhD
Professor of Medicine and Biostatistician
Division of Pharmacoepidemiology and
Pharmacoeconomics
Department of Medicine
Harvard Medical School and Brigham & Women's Hospital
Boston, Massachusetts

Steven B. Meisel, PharmD
System Director of Patient Safety
Fairview Health Services
Minneapolis, Minnesota

Anne-Michelle Ruha, MD, FACMT
Director, Medical Toxicology Fellowship Program
Department of Medical Toxicology
Banner University Medical Center
Clinical Associate Professor of Emergency Medicine
University of Arizona College of Medicine
Phoenix, Arizona
Terri L. Warholak, PhD, RPh, CPHQ, FAPhA
Professor and Assistant Dean
Academic Affairs and Assessment
College of Pharmacy
University of Arizona
Tucson, Arizona

TEMPORARY MEMBERS (Voting)

Jane B. Acri, PhD
Chief, Medication Discovery & Toxicology Branch
Division of Therapeutics and Medical Consequences
National Institute on Drug Abuse
National Institutes of Health (NIH)
Bethesda, Maryland
**David Cella, PhD**
Professor Chairperson  
Medical Social Sciences  
Northwestern University Feinberg School of Medicine  
Chicago, Illinois

**Stephanie Y. Crawford, PhD, MPH**
Department of Pharmacy Systems  
Outcomes and Policy  
University of Chicago  
Chicago, Illinois

**Harriet de Wit, PhD**
Department of Psychiatry and Behavioral Neuroscience  
The University of Chicago  
Chicago, Illinois
Kathryn E. Flynn, PhD
Associate Professor of Medicine
Division of Hematology and Oncology
Senior Scientific Director for Patient-Reported Outcomes
Center for International Blood & Marrow Transplant Research
Milwaukee, Wisconsin

Roxanne E. Jensen, PhD
Program Director
Outcomes Research Branch
Health Care Delivery Research Program
Division of Cancer Control and Population Sciences
National Cancer Institute, NIH
Bethesda, Maryland

Elizabeth Joniak-Grant
(Patient Representative)
Holly Springs, North Carolina
Brandon D.L. Marshall, PhD
Associate Professor
Department of Epidemiology
Brown University School of Public Health
Providence, Rhode Island

William T. Riley, PhD
Director
Office of Behavioral and Social Sciences Research
NIH
Bethesda, Maryland

FDA PARTICIPANTS (Non-Voting)

Robert Temple, MD
Deputy Director for Clinical Science
CDER FDA
Deputy Director (Acting)
Office of Drug Evaluation I (ODE I)
Office of New Drugs (OND), CDER, FDA
Ellis Unger, MD
Director
ODE I, OND, CDER, FDA

Mitchell Mathis, MD
Division Director
Division of Psychiatry Products (DPP)
ODE I, OND, CDER, FDA

Judy Staffa, PhD, RPh
Associate Director for Public Health Initiatives
Division of Epidemiology II (DEPI-II)
Office of Pharmacovigilance and Epidemiology
Office of Surveillance and Epidemiology
CDER, FDA

Tiffany Farchione, MD
Deputy Director
DPP, ODE I, OND, CDER, FDA
Dominic Chiapperino, PhD
Director
Controlled Substance Staff
Office of the Center Director, CDER, FDA

Daniel J. Lee, MD
Clinical Reviewer
DPP, ODE I, OND, CDER, FDA

Semhar Ogbagaber, PhD
Statistician
Division of Biometrics I
Office of Biostatistics
Office of Translational Sciences, CDER, FDA
## CONTENTS

<table>
<thead>
<tr>
<th>AGENDA ITEM</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Call to Order and Introduction of Committee</td>
<td>17</td>
</tr>
<tr>
<td>Raj Narendran, MD</td>
<td></td>
</tr>
<tr>
<td>Conflict of Interest Statement</td>
<td>23</td>
</tr>
<tr>
<td>Kalyani Bhatt, BS, MS</td>
<td></td>
</tr>
<tr>
<td>FDA Opening Remarks</td>
<td>27</td>
</tr>
<tr>
<td>Mitchell Mathis, MD</td>
<td></td>
</tr>
<tr>
<td><strong>Applicant Presentations – Alkermes, Inc.</strong></td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td>40</td>
</tr>
<tr>
<td>Lisa von Moltke, MD</td>
<td></td>
</tr>
<tr>
<td>The Unmet Need in MDD and Challenges in MDD Clinical Trials</td>
<td>51</td>
</tr>
<tr>
<td>George Papakostas, MD</td>
<td></td>
</tr>
<tr>
<td>Clinical Efficacy</td>
<td>65</td>
</tr>
<tr>
<td>Jerald Schindler, DrPH</td>
<td></td>
</tr>
<tr>
<td>Clinical Safety, and Risk Mitigation Strategies</td>
<td>97</td>
</tr>
<tr>
<td>Gary Bloomgren, MD</td>
<td></td>
</tr>
<tr>
<td>AGENDA ITEM</td>
<td>PAGE</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Clinical Perspective and Benefit-Risk Profile</td>
<td></td>
</tr>
<tr>
<td>Sanjay Mathew, MD</td>
<td>116</td>
</tr>
<tr>
<td>Conclusion</td>
<td></td>
</tr>
<tr>
<td>Lisa von Moltke, MD</td>
<td>122</td>
</tr>
<tr>
<td>Clarifying Questions to Applicant</td>
<td>123</td>
</tr>
<tr>
<td><strong>FDA Presentations</strong></td>
<td></td>
</tr>
<tr>
<td>Regulatory History</td>
<td></td>
</tr>
<tr>
<td>Tiffany Farchione, MD</td>
<td>155</td>
</tr>
<tr>
<td>Clinical Efficacy and Safety Overview</td>
<td></td>
</tr>
<tr>
<td>Semhar Ogbagaber, PhD</td>
<td>166</td>
</tr>
<tr>
<td>Daniel Lee, MD</td>
<td>181</td>
</tr>
<tr>
<td>Abuse Potential of Buprenorphine/Samidorphan (BUP/SAM)</td>
<td></td>
</tr>
<tr>
<td>Edward Hawkins, PhD</td>
<td>187</td>
</tr>
<tr>
<td>FDA Review of the Epidemiologic and Surveillance Data</td>
<td></td>
</tr>
<tr>
<td>Celeste Mallama, PhD, MPH</td>
<td>197</td>
</tr>
<tr>
<td>AGENDA ITEM</td>
<td>PAGE</td>
</tr>
<tr>
<td>----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>University of Washington Medical School</td>
<td></td>
</tr>
<tr>
<td>Department of Psychiatry and Behavioral Sciences Presentation</td>
<td></td>
</tr>
<tr>
<td>Depression Effects on Long-Term Prescription Opioid Use, Abuse and Addiction</td>
<td></td>
</tr>
<tr>
<td>Mark Sullivan, MD, PhD</td>
<td>207</td>
</tr>
<tr>
<td>FDA Presentations (continued)</td>
<td></td>
</tr>
<tr>
<td>Risk Management for Buprenorphine/Samidorphan</td>
<td></td>
</tr>
<tr>
<td>Somya Dunn, MD</td>
<td>221</td>
</tr>
<tr>
<td>Clarifying Questions to FDA and Dr. Sullivan</td>
<td>228</td>
</tr>
<tr>
<td>Open Public Hearing</td>
<td>268</td>
</tr>
<tr>
<td>Clarifying Questions (continued)</td>
<td>332</td>
</tr>
<tr>
<td>Charge to the Committee</td>
<td></td>
</tr>
<tr>
<td>Mitchell Mathis, MD</td>
<td>335</td>
</tr>
<tr>
<td>Questions to the Committee and Discussion</td>
<td>336</td>
</tr>
<tr>
<td>Adjournment</td>
<td>446</td>
</tr>
</tbody>
</table>
PROCEEDINGS

(8:01 a.m.)

Call to Order

Introduction of Committee

DR. NARENDRAN: I would first like to remind everyone to please silence your cell phones, smartphones, other devices if you have not already done so. I would also like to identify the FDA press contact, Sandy Walsh. If you're there, please stand. She's over there.

My name is Raj Narendran. I'm the chairperson for today's meeting. I will now call the Joint Meeting of the Psychopharmacology Drug Advisory Committee and the Drug Safety and Risk Management Advisory Committee to order.

We'll start by going around the table and introduce ourselves. We will start with the FDA to my left and go around the table.

DR. UNGER: Good morning. I'm Ellis Unger. I'm director of Office of Drug Evaluation I in the Office of New Drugs, CDER.

DR. MATHIS: Mitchell Mathis, director of
the Division of Psychiatry Products.

DR. FARCHIONE: Tiffany Farchione, deputy director of the Division of Psychiatry Products.

DR. OGBAGABER: Dr. Semhar Ogbagaber, Division of Psychiatry, Office of Biostatistics Division I

DR. STAFFA: Good morning. I'm Judy Staffa. I'm the associate director for public health initiatives in the Office of Surveillance and Epidemiology.

DR. CHIAPPERINO: Good morning. I'm Dominic Chiapperino. I'm the director in the controlled substance staff.

DR. RUHA: Hi. I'm Michelle Ruha. I'm a medical toxicologist and professor at the University of Arizona College of Medicine, Phoenix.

DR. WARHOLAK: Hi. I'm Terri Warholak, and I'm a professor and assistant dean at the University of Arizona College of Pharmacy, Tucson.

DR. KULLDORFF: Good morning. My name is Martin Kulldorff. I'm a biostatistician at the Brigham and Women's Hospital and Harvard Medical
School.

DR. HABEL: Hi. I'm Laurel Habel. I'm an epidemiologist and associate director for cancer research at Kaiser Permanente Northern California.


DR. IYENGAR: I'm Satish Iyengar. I'm a professor of statistics at the University of Pittsburgh. I also have a joint appointment in psychiatry at Western Psych.

DR. DUNN: Walter Dunn, psychiatrist, assistant professor at UCLA in the West Los Angeles VA.

MS. BHATT: Good morning. I'm Kalyani Bhatt. I'm the designated federal officer for this committee.

DR. NARENDRAN: Raj Narendran, psychiatrist, University of Pittsburgh Medical Center.

DR. FIEDOROWICZ: Jess Fiedorowicz. I'm a psychiatrist, associate professor of psychiatry, epidemiology, and internal medicine at the
University of Iowa.

MS. WITCZAK: Good morning. Kim Witczak.

I'm consumer representative and founder of Woodymatters.

MS. JONIAK-GRANT: Hello. I'm Elizabeth Joniak-Grant. I'm the patient representative. I'm also a sociologist, and I'm affiliated with the National Coalition of Independent Scholars.

DR. BESCO: Good morning. I'm Kelly Besco. I'm a pharmacist by background, and I currently serve as the medication safety officer for the Ohio Health Hospital System in Columbus, Ohio.

DR. GRIFFIN: Good morning. Marie Griffin, pharmacoepidemiologist and general internist at Vanderbilt University.

DR. MEISEL: Steve Meisel, director of medication safety, Fairview Health Services in Minneapolis.

DR. ACRI: Good morning. I'm Jane Acri. I'm chief of the Medication Discovery and Toxicology Branch at the National Institute on Drug Abuse.
DR. DE WIT: My name is Harriet de Wit. I'm at the University of Chicago, Department of Psychiatry and Behavioral Neuroscience.

DR. JENSEN: Hi. My name is Roxanne Jensen. I'm a program director in the outcomes research branch at the National Cancer Institute.

DR. FLYNN: Hi. I'm Kathryn Flynn. I'm an associate professor of medicine at the Medical College of Wisconsin and an outcomes researcher.

DR. MARSHALL: Good morning. I'm Brandon Marshall. I'm an associate professor of epidemiology at the Brown School of Public Health in Providence, Rhode Island.

DR. CRAWFORD: Good morning. My name is Stephanie Crawford. I'm professor at the University of Illinois at Chicago, Colleges of Pharmacy and Medicine, and I'm also executive associate dean.

DR. RILEY: Bill Riley. I'm the associate director for behavioral and social sciences at the National Institutes of Health.

DR. CELLA: Hi. I'm David Cella, also from
Chicago, but Northwestern University at Chicago. I'm a clinical psychologist and outcomes researcher.

DR. CONLEY: Good morning. I'm Rob Conley. I'm a psychiatrist. I'm the distinguished scholar for neuroscience at Eli Lilly and a professor of psychiatry and pharmacy science at the University of Maryland School of Medicine. I'm here as the industry representative.

DR. NARENDRAN: Dr. Jain, if you could introduce yourself.

DR. JAIN: Hi. I'm Felipe Jain, assistant professor of psychiatry at Harvard Medical School and a psychiatrist at the Massachusetts General Hospital.

DR. NARENDRAN: Thank you. For topics such as those being discussed at today's meeting, there are often a variety of opinions, some of which are quite strongly held. Our goal is that today's meeting will be a fair and open forum for discussion of these issues and those individuals can express their views without interruption.
Thus, as a gentle reminder, individuals will be allowed to speak into the record only if recognized by the chairperson. We look forward to a productive meeting.

In the spirit of the Federal Advisory Committee Act and the Government in the Sunshine Act, we ask that advisory committee members take care that their conversations about the topic at hand take place in the open forum of the meeting.

We are aware that members of the media are anxious to speak with the FDA about these proceedings. However, FDA will refrain from discussing the details of this meeting with the media until its conclusion. Also, the committee is reminded to please refrain from discussing the meeting topic during breaks or lunch. Thank you.

Now, I will pass it to Kalyani Bhatt, who will read the conflict of interest statement.

**Conflict of Interest Statement**

MS. BHATT: Good morning. The Food and Drug Administration is convening today's Joint Meeting of the Psychopharmacologic Drugs Advisory Committee.
and the Drug Safety and Risk Management Advisory Committee under the authority of the Federal Advisory Committee Act, FACA, of 1972.

With the exception of the industry representative, all members and temporary voting members of the committees are special government employees, SGEs, or regular federal employees from other agencies and are subject to federal conflict of interest laws and regulations.

The following information on the status of the committee's compliance with federal ethics and conflict of interest laws, covered by but not limited to those found at 18 U.S.C. Section 208, is being provided to participants in today's meeting and to the public.

FDA has determined that members and temporary voting members of these committees are in compliance with the federal ethics and conflict of interest laws. Under 18 U.S.C. Section 208, Congress has authorized FDA to grant waivers to special government employees and regular federal employees who have potential financial conflicts,
when it is determined that the agency's need for a special government employee's services outweighs his or her potential financial conflict of interest or when the interest of a regular federal employee is not so substantial as to be deemed likely to affect the integrity of the services which the government may expect from the employee.

Related to the discussion of today's meeting, members and temporary voting members of these committees have been screened for potential financial conflicts of interest of their own as well as those imputed to them, including those of their spouses or minor children, and for purposes of 18 U.S.C. Section 208, their employers.

These interests may include investments, consulting, expert witness testimony, contracts, grants, CRADAs, teaching, speaking, writing, patents and royalties, and primary employment.

Today's agenda involves discussion of the efficacy, safety, and risk-benefit profile of new drug application NDA210417 for buprenorphine and samidorphan sublingual tablets, submitted by
Alkermes, Incorporated for adjunctive treatment of major depressive disorder.

This is a particular matters meeting during which specific matters related to Alkermes' NDA will be discussed. Based on the agenda for today's meeting and all financial interests reported by the committee members and temporary voting members, no conflict of interest waivers have been issued in connection with this meeting.

To ensure transparency, we encourage all standing committee members and temporary voting members to disclose any public statements that they have made concerning the product at issue.

With respect to FDA's invited industry representative, we would like to disclose that Dr. Robert Conley is participating in this meeting as a non-voting industry representative, acting on behalf of regulated industry. Dr. Conley's role at this meeting is to represent industry in general and not any particular company. Dr. Conley is employed by Eli Lilly and Company.

We'd like to remind members and temporary
voting members that if the discussions involve any
other products or firms not already on the agenda
for which an FDA participant has a personal or
imputed financial interest, the participants need
to exclude themselves from such involvement, and
their exclusion will be noted for the record.

FDA encourages all other participants to
advise the committee of any financial relationships
that they may have with the firm at issue. Thank
you.

DR. NARENDRAN: We will now proceed with the
FDA's introductory remarks, presented by
Dr. Mitchell Mathis, division director.

DR. MATHIS: I think that's the wrong
presentation.

MS. BHATT: We're having some issues,
Dr. Mathis. Just give us a minute.

FDA Opening Remarks - Mitchell Mathis

DR. MATHIS: Thank you.

My name is Dr. Mitchell Mathis, and I'm the
director of psychiatry products in the Center for
Drugs here at FDA. I'd like to thank you all for
coming. It's good to see so many people so early.

Thank you for being here.

What I would like to do is give you an overview of what we're going to talk about today, to try and set the stage for the issues that my team and I think are important to talk about.

So by way of introduction, I will introduce the product, buprenorphine samidorphan; introduce the parts of the disease that the company is developing this drug to treat, the form of the disease; talk about current available treatments that we have; and we have to spend some time discussing the definition of substantial evidence. I think you will see that that will be important to today's discussion.

I'd also like to engage the committee for their thoughts on the management of placebo response. And you'll see as we proceed today that the company has employed some strategies to address placebo response. I'll go over the agenda for today and then the core questions for the committee.
The product is a combination product of buprenorphine and samidorphan. Buprenorphine is an approved drug. It's a partial agonist at the mu opioid receptor, and it's approved for the treatment of medication-assisted treatment in opioid dependence disorders, and it's approved to treat pain.

Samidorphan is the new molecular entity here. Samidorphan is an antagonist, a blocker at that same opioid receptor. And the pharmacologic idea between the combination of these drugs is to provide the buprenorphine and the samidorphan together in a bioavailable formulation.

Samidorphan binds at mu opioid receptor and prevents the opioid effects of the buprenorphine, whereas the other effects of buprenorphine, presumably which include treating depression, can be present without the dependency and abuse risk. That's how it was designed.

It's important to note that the indication that's being sought is the adjunctive treatment of major depressive disorder. Now, adjunctive
treatment of major depressive disorder, briefly, means this. The patient has been exposed to a monotherapy agent and has gotten better, but not good enough. Your choices then, as a clinician, are to stop the medication, start another, and see if you can get a single medication to work.

If you've done that one or two times and the patient is better but not good enough, then the technique employed is adjunctive treatment. So you can add a second drug -- this drug was developed for that reason -- to make the patient completely better, closer to remitted. And that's the idea behind adjunctive treatment of major depressive disorder.

The disease major depressive disorder is of course a debilitating and a chronic illness. There are millions of Americans right now suffering from major depressive disorder. It's a leading cause of disability worldwide. And unfortunately, partial response is more the rule than the exception.

The sequence treatment alternatives to relieve depression, study STAR*D, about a third of
patients get better the first time you try to make
them better with a monotherapy treatment. Now,
that means get better to a remitted state, and you
can define remission in different ways, but to
remission.

So there is a public health need for
effective add-on medications for the rest of the
patients who don't remit.

In terms of approved medications, we have
several drugs in different classes for monotherapy,
but there are only three drugs approved for this
indication: the adjunctive treatment of partially
responsive depression, quetiapine, extended
release; aripiprazole; and brexpiprazole.

I'm sure many of you have noticed that those
medications are all in the same class of drugs,
atypical or second-generation antipsychotics, which
means that they have safety risks in common.
Movement disorders, acute and chronic movement
disorders are a possibility with those drugs and
the metabolism syndrome: weight gain,
hypercholesterolemia, hyperglycemia. All three of
those are risk factors, cardiovascular risk factors.

So buprenorphine/samidorphan, should it be approved, would be in a different class to treat this partially responsive depression.

Buprenorphine/samidorphan, however, has its own safety risks, which we'll talk about today, and they're different than the other drugs, but remember that buprenorphine is an opiate, and samidorphan has reduced the dependency risk and the problems of the opiate, but not completely remove them.

Substantial evidence; this concept is important because we're going to focus on this a great deal today. In 1962, the Food, Drug, and Cosmetic Act was amended to require for the first time that drug manufacturers establish a drug's effectiveness by this concept of substantial evidence.

Substantial evidence is defined as evidence consisting of adequate and well-controlled investigations, and it's long been FDA's position
that Congress intended for investigations to be plural in the sense that at least two adequate and well-controlled trials, each convincing on its own, should be submitted to support a drug's effectiveness.

There are situations where a single study can be used to approve a drug. In 1997, the FDA Modernization Act, FDAMA, amended that same section of the Act to make it clear that FDA may rely, quote, "on data from one adequate and well-controlled clinical investigation and confirmatory evidence" end quote, to constitute substantial evidence if FDA determines that these data and evidence are sufficient to establish effectiveness.

So practically speaking, this single study with confirmatory evidence is reserved for situations where an important clinical benefit like a direct effect on survival is clearly -- when I say clearly, I mean statistically clearly and clinically clearly -- so that a confirmatory study, would be either hard to repeat because there, for instance, aren't enough patients, or there would be
an ethical reason not to repeat the study, to take
the time or the exposure to repeat the study.

That's not the case with adjunctive
treatment of major depressive disorder. As I've
already mentioned, it's unfortunate, but we have
many patients who have partially responsive
depression, and the other three drugs, including
this drug, have done multiple studies to assess
efficacy.

It's important to note that FDA and the
applicant do not agree on whether this drug has met
the standard for substantial evidence of
effectiveness. The applicant will argue that they
have two positive studies, study 202 and 207. To
provide substantial evidence of effectiveness for
buprenorphine/samidorphan, FDA will argue that
substantial evidence has not been provided, and my
team will provide our arguments to support this
case. We'll be discussing the efficacy in depth,
and I ask that you listen to both sides of the
argument and form your own opinion, which we will
ask you for later.
In terms of safety, the applicant will argue that the opiate effects of buprenorphine are largely but not completely blocked by samidorphan, and my team and I will agree with that. There are some opioid effects, but samidorphan seems to have largely done whatever it was designed to do.

Here are the trials. The applicant has conducted four trials designed to demonstrate efficacy, two they consider positive and two they consider negative. Since we agree on the negative trials and since it's not usual to see up to half of trials in MDD development programs fail, we need not discuss these negative trials further with regard to efficacy in my view.

I've asked my team to carefully discuss why they have concluded that studies 202 and 207 are not positive because there's disagreement with the applicant on this point and because this will be of obvious importance when we're asking you your opinion about substantial evidence.

Briefly, 202 was originally designed as a phase 2 proof-of-concept study, so it was designed
without a multiplicity plan to adjust for evaluating the efficacy of multiple doses in the same trial.

It's noteworthy to note that the study is negative at the higher dose, the 8/8 dose; 8/8 means 8 milligrams of buprenorphine, 8 milligrams of samidorphan. And it's negative if the 8/8 and the 2/2 doses are combined. It's also negative and it's important to point out if a single patient is removed from the analysis, and we'll talk more about that later.

The lack of an effect at a higher dose is very concerning to us. It undercuts whatever evidence there might be that a lower dose could be effective here. It's easy to propose reasons why a lower dose might work and a higher dose doesn't, but in our experience, we've never seen this.

If we were to approve this drug on the basis of two positive studies, with study 202 considered as one of the two, it's important to recognize that we'd be approving the first drug of an entirely new class, knowing that evidence of effectiveness
hinges on data from a single patient. Remove this patient and the nominal p-value is not statistically significant anymore. There will be more about this later.

Study 207 is positive only by an average of endpoints over multiple visits. FDA advised against this approach. And it is negative by the usual end-of-treatment endpoint on the depression rating scales.

It might not be irrational to average rating scales in depression over multiple visits, consecutive visits, to perhaps decrease variability, but we've never taken a close look at this, and we've explicitly told the applicant not to do it here.

In closing, both studies, 202 and 207, are of the SPCD design, Sequential Parallel Comparison Design. We'll talk more about this later. It's a design, however, that has not yet been determined to be statistically acceptable to FDA.

I'd like to talk a little bit about the management of placebo response. The sponsor has
done SPCD trials and a placebo run-in trial to manage the placebo response. The placebo response is, of course, a non-specific response to treatment that's not related to active drug. And while this is useful in the clinic in taking care of patients, it is destructive to psychiatric drug development trials.

It's a long-term problem in psych trials. The Sequential Parallel Comparison Design was set up to address this problem by putting more weight on placebo non-responder data, and we'll talk to you about that, but again, that's never been accepted in the division.

So today, the applicant will discuss the need for new treatments, the efficacy, safety, and the risk-benefit of their drug product. FDA and our guest speaker will then present our views of efficacy and safety. We'll talk about the abuse potential, the epidemiology of misuse and abuse, the relationship of opioid use disorders to major depressive disorder and how our risk evaluation and mitigation strategy, REMS, could help to manage the
risk of buprenorphine/samidorphan, should it be approved to treat MDD.

The committee will then discuss and vote on several questions. I put the core questions here so that you can see them before we get started. There will be other questions, but these are the ones that we'll need votes on.

Has substantial evidence of efficacy of buprenorphine/samidorphan been presented by the applicant? Has the applicant adequately characterized the safety of buprenorphine/samidorphan in treating major depressive disorder? And do the available data support a favorable benefit-risk profile of buprenorphine/samidorphan to support approval.

That's all I have. I hand the meeting back to the chair.

DR. NARENDRAN: Thank you, Dr. Mathis.

Both the FDA and the public believe in a transparent process for information gathering and decision making. To ensure such transparency at the advisory committee meeting, FDA believes that
it is important to understand the context of an individual's presentation.

For this reason, FDA encourages all participants, including the sponsor's non-employee presenters, to advise the committee of any financial relationships that they may have with the firm at issue, such as consulting fees, travel expenses, honoraria, and interest in the sponsor, including equity interests and those based upon the outcome of the meeting.

Likewise, FDA encourages you, at the beginning of your presentation, to advise the committee if you do not have any such financial relationships.

If you choose not to address this issue of the financial relationships at the beginning of your presentation, it will not preclude you from speaking. We will now proceed with Alkermes' presentations.

**Applicant Presentation - Lisa von Moltke**

DR. VON MOLTKE: Thank you, Mr. Chairman, and good morning. Good morning to the other
members of the committee, representatives from the agency, and other attendees. I'm Lisa von Moltke, and I head up drug development at Alkermes, and I'm going to be starting this morning with the sponsor's presentation. We appreciate this opportunity to present to you the buprenorphine/samidorphan program for adjunctive treatment of major depressive disorder.

Following my introduction, the rest of the sponsor's presentation will proceed as outlined here. Dr. George Papakostas from Harvard Medical School will take us through the unmet need in MDD. He will also outline the challenges in MDD clinical trials.

Dr. Jerry Schindler from Alkermes will take you through the clinical efficacy data. Dr. Gary Bloomgren will take you through the clinical safety data, and he'll also present the risk mitigation strategies that Alkermes will be proposing.

Dr. Sanjay Mathew from Baylor College of Medicine will give his clinical perspective and the benefit-risk profile, and then I will close the
sponsor's presentation.

We also have with us an external participant. Dr. Aparna Anderson is with us from Statistics Collaborative. Alkermes also has a number of participants, and they are listed here, and they will introduce themselves as they answer questions for the committee.

Now, there will be some key topics under discussion today that I want to outline before we get started. First, there is a major significant unmet need for patients with major depressive disorder. Specifically, there is a need for therapies that work by new mechanisms of action.

There are significant development challenges in studying MDD, including the high placebo response rates, which you've already heard about. And this warrants the use of advances in methodology, including the SPCD trial design.

You're going to see data indicating that buprenorphine/samidorphan has a positive benefit-risk profile. The data include substantial evidence of efficacy. It's comparable to other
adjunctive treatments, but with a completely
different mechanism of action. The safety profile
has been well characterized, and it's manageable,
and the abuse potential is low.

Finally, Alkermes realizes that it is
proposing a new opioid modulator in the midst of an
opioid crisis. Even with a profile that indicates
that the risk for abuse and dependence has been
mitigated by the addition of samidorphan, this will
call for a full commitment by Alkermes to the
obligations of education and monitoring, and we
will be presenting those to you.

I want to highlight some specific points
about the program to serve as a framework for
information that you're going to receive this
morning. And one of the first things is that I'm
going to be referring to buprenorphine/samidorphan
as BUP/SAM for the sake of efficiency.

The first point is that this program had its
genesis in the clinical observations that opioids,
and buprenorphine in particular, may have
antidepressant effects.
Samidorphan is being added as a way to mitigate the abuse and dependence potential that we know exists with buprenorphine. And as you've already heard, samidorphan is a mu opioid antagonist.

This therapy is being proposed for adjunctive treatment of major depressive disorder, so this is for patients who are not responding adequately to available therapies, and they're continuing to suffer with depressive symptoms.

The proposed dose is 2 milligrams of BUP and 2 milligrams of SAM, following a 1-week initial titration. And there will also be a 1:1 dose for special populations, and that's the patients that have renal insufficiency or hepatic insufficiency.

The BUP/SAM development regulatory history began with a pre-IND meeting in 2011. And during that meeting, there were initial conversations around study designs, including SPCD. There was an end-of-phase 2 meeting in October of 2013, and during that time, there were again discussions around SPCD study designs.
The clinical review that came out of that meeting indicated that, from a clinical perspective, the trials appeared acceptable. That same year, the program received fast-track designation since BUP/SAM has the potential to address an unmet medical need that serious. And this was based on the phase 2 study, and that's the 202 study that you saw presented by Dr. Mathis.

There have been two scientific exchange meetings, one in 2016 and one in 2017, and during that time, the results of phase 3 studies were shared. There was then a pre-NDA meeting in July of 2017, and the NDA submission content was discussed.

The applicant then filed in January of 2018, and Alkermes was very surprised to get a refuse-to-file in March. This was based on some factual inaccuracies in the agency's assessment, and the refuse-to-file was rescinded just 2 weeks later.

Before I leave the regulatory history, I want to note a few points here that are not on this slide. This process has been an ongoing multi-year
interaction with FDA with multiple points of contact. Throughout the entire program, we have been clear and FDA has acknowledged that our evidence of efficacy was going to rely upon SPCD studies.

At the pre-NDA meeting last summer, we presented our plan to submit based on evidence from studies 202, 205, and 207, and this was confirmed with FDA. And as you've already heard, today you're going to hear a couple areas of disagreement, and these are regarding the utility of the 202 study and the importance of the data from the 205 study.

The 202 study was robustly designed and met its prespecified primary endpoint. Now, FDA is going to assert that a particular patient cannot be included due to lack of eligibility. Alkermes has sent a team and conducted an extensive investigation and audit, and we were able to substantiate multiple sources documenting the eligibility, and these have been submitted to the agency.
The 205 study is an important supportive piece of evidence, and there is regulatory precedent in MDD for considering it as such. And we will be presenting the details of these studies, both 202 and 205, to you this morning.

Alkermes undertook the development of BUP/SAM because MDD is a major source of morbidity and disability, and this illness carries a risk of suicide. A significant percentage of patients do not achieve adequate symptom relief with standard therapies. All the approved antidepressants work via monoaminergic mechanisms, so they're all working on serotonin, norepinephrine, or dopamine, and only antipsychotics are currently approved as adjunctive therapies, and these can have significant side effects. They can be serious and they can be permanent.

So new approaches to treat depression, and adjunctive treatment in particular, are urgently needed. These options are needed by patients, and the physicians, and other healthcare professionals who care for them.
The BUP/SAM program is founded on evidence in the literature that exists over decades that opioids, and particularly buprenorphine, may have antidepressant effects. We know that opioid receptors are highly expressed in brain regions that are associated with emotional regulation. But a broader use in this indication has been limited by the concerns around abuse and dependence that we know exists with buprenorphine.

Samidorphan has been added to address this abuse and dependence liability, and it's an integral part of the resulting pharmacologic profile for this therapy. And that's in contrast to other antagonists that are added simply to deter parenteral use, for example.

So the resulting therapy would consist of buprenorphine, which is a partial mu opioid agonist and a kappa opioid antagonist, and it has bioavailability by the sublingual route. Samidorphan, as you've heard, is the new molecular entity here. And it's a potent mu opioid antagonist. It's been optimized for that high
potency at the receptor. It's also been optimized
to have high bioavailability by the oral and
sublingual routes.

Again, this contrasts to other antagonists
that are combined with buprenorphine that do not
have bioavailability by the intended routes. So
the combination is co-formulated in a single
sublingual tablet, and it would be packaged in
individual blisters, as you see here.

This has been a large program with an
extensive safety and efficacy evaluation of
BUP/SAM. There have been 34 clinical studies, 19
conducted with BUP/SAM together and 15 with
samidorphan alone. There have been 4 placebo-
controlled studies in the MDD patient population,
and then there's been a long-term safety study.

Over 2,000 patients, so our subjects have
received BUP/SAM together, and over 1500 MDD
patients have been treated with the specific 2/2
dose.

The study population for the development
program is comprised of patients diagnosed with
MDD. They had a mean lifetime number of major
depressive episodes that ranged from 4 to 7
depending on the trial. They had all cycled
through multiple therapies, and the median duration
of the current episode was anywhere from 9 to
10 months.

They had had 1 to 2 inadequate responses to
an antidepressant therapy in the current cycle and
episode. So they also continued on their
background ADT, so they were all also on an SSRI,
SNRI, or bupropion.

So this morning, you're going to see the
evidence of efficacy and safety. Dr. Schindler is
going to present to you the two studies that met
their prespecified primary endpoints. These are
202 and 207. He's also going to present to you the
supportive evidence of efficacy from a third study,
and that's study 205.

The efficacy seen in these studies is
comparable to that seen with the other approved
adjunctive treatments for MDD, and it achieves this
efficacy by a totally new mechanism of action.
Dr. Gary Bloomgren is then going to present to you the safety data, and you will see that this therapy is a generally well-tolerated therapy. Common AEs are GI and sedation related, and the abuse potential is low. The overall benefit-risk profile is positive.

Finally, because of the presence of buprenorphine, even in the setting where this abuse and dependence has been mitigated by the presence of samidorphan, Alkermes is going to commit to education and monitoring to ensure that the use of this therapy is appropriate, and you're going to hear about those propositions as well.

With that, I'm going to turn the presentation over to Dr. George Papakostas, who will take us through the unmet need in MDD.

Dr. Papakostas?

**Applicant Presentation - George Papakostas**

DR. PAPAKOSTAS: Thank you very much, Dr. von Moltke.

Good morning, everyone. My name is George Papakostas, and I am associate professor of
psychiatry at Harvard Medical School, director of
treatment-resistant depression studies at
Massachusetts General Hospital, and also a
scientific director at MGHCTNI. And essentially,
I'm up here this morning to convey two important
ideas that are clearly interlinked. So there's
going to be two components to this brief
presentation.

The first essentially echoes what the first
two discussions, the first two talks covered this
morning, is that major depressive disorder is a
serious and common medical disorder affecting
humanity. And in my point of view, and many would
agree, we fall terribly short in terms of
treatments for major depressive disorder. We have
a long way ahead of us to develop more treatments
for major depressive disorder.

It's a common medical illness. The lifetime
prevalence in adults in the United States is 16.6
percent as per 1 estimate. Because it involves so
many symptoms, it significantly impairs home, work,
life relationships, and social life, such that
almost 60 percent of patients with depression over 12 months report severe or very severe role impairment. So it affects functioning and quality of life at its core.

It's, as such, the number one contributor of disability worldwide. It's a very disabling illness. And being a medical illness, it increases the risk of other medical illnesses such as heart disease, diabetes mellitus, and cancer. And equally importantly, is a major risk factor for suicide. Suicide is climbing as a threat to humanity. It results in more deaths than terrorism and war put together. It is a major health threat.

So in my view, there are two gaps, there are two big gaps in our development of treatments for major depressive disorder. The first gap is that over the last 50 years, we focused extensively on developing monotherapies for depression. Monotherapy is a clinical term that essentially means giving a sole drug to treat major depressive disorder.

We didn't find out until late, until 2006,
how much that strategy falls short of patients' needs. Here is a slide from STAR*D. STAR*D is the largest clinical trial ever conducted in psychiatry, and it essentially looked at the following question.

If you took patients with depression who were treatment naïve and gave them different options in succession, how likely was it that patients would get better over these sequential treatments?

So as we heard from Dr. Mathis' presentation, treatment-naïve patients, adults with MDD, you give the first treatment -- in this case, it was an SSRI, citalopram -- one-third achieved remission; two-thirds remain symptomatic.

You then switch, which is a clinical term meaning you stop one therapy and you start another, second-line antidepressant monotherapy; three-quarters remain symptomatic. Then you go to a third-line antidepressant therapy; 80 to 90 percent remain symptomatic. You go to a fourth-line antidepressant monotherapy; again, almost
90 percent remain symptomatic. So the message from STAR*D here is clear. If you rely on antidepressant monotherapy, eventually it will become a futile effort.

The challenge, then, is that if there's partial or non-response in patients with depression, you get additional bad outcomes on top of the outcomes of MDD. Patients who are partial or non-responders, versus remitters, are at higher risk for their illness coming back. As you can see in the survival analysis on the left-hand side of this slide, in blue you have no persistent symptoms; in orange, you have persistent symptoms; and the further down you are means that less patients remain illness free.

There's a clear separation. The more you get through remission, the safer you are from this illness coming back. Patients who are incomplete, insufficient responders, are 2 times more likely to be hospitalized, are twice more likely to attempt suicide, and are 3 times more likely to use additional psychotropic medications compared to the
overall MDD population. So we need to get as many patients better as quickly as possible to avoid additional bad outcomes.

The second limitation of the development of antidepressants is that for the past 50 years, we've relied on essentially the permutation of a single mechanism of action; that is, blocking the reuptake of monoamines, predominantly serotonin, sometimes norepinephrine, on rare occasions dopamine, a very unimaginative approach and very limited approach.

Of course, clinicians have caught on to the idea that antidepressant monotherapies have become progressively futile and they use adjuncts. Up until 10 years ago, 15 years ago, there were no FDA-approved adjuncts. If you used something, there was limited evidence telling you it would work, even though that is the most common approach after 1 or 2 treatments fail, as we heard from the earlier presenters.

Fortunately, we have three FDA-approved adjuncts today. The limitation is that they're all
atypical antipsychotics and they're all monoaminergic. So essentially, if you rely on the evidence base for adjuncts, you have to have the difficult discussion with the patient that, if you try this medication, which has been proven to be effective in treating depression, depending on what you use, you may have to have a difficult talk about weight gain, the chances of weight gain, dyslipidemia, glucose dysregulation on rare occasion, but very important, new onset diabetes.

You have to talk about akathisia, the fact that they may experience restlessness, which will be apparent to others, Parkinsonism, stiffness. You have to talk about tardive dyskinesia. Patients with depression often need long-term therapy. You don't just get them better and take them off their meds. The illness will recur.

The chances of tardive dyskinesia, which is essentially motor ticks that are involuntary and permanent, increase over time. This is a difficult discussion to have with patients. This is a big burden for patients, and this drives patients away
from a potentially effective therapy, shunting them to adjuncts that are not really supported by substantial evidence, and of course the difficult discussion about potentially life-threatening neuroleptic malignant syndrome that's been observed with the antipsychotics, including the atypicals, that in some patients, body temperature can rise and this is potentially fatal; useful agents from an efficacy standpoint; a lot of difficult conversations, and risks of the safety side of things.

So this is where we stand today. We have a difficult and serious illness that's negatively impacting humanity more and more each year. We have a number of therapies. Patients are still symptomatic. Many patients are still symptomatic. Symptomatic patients have a higher illness burden than the general MDD population and especially those that are partial non-responders.

What we have today work on a single mechanism of action, so we need more treatments that replicate the efficacy of previous treatments,
are not associated with side effects, can be used adjunctively, and hopefully are more imaginative in terms of their mechanism of action, which brings me to my second talk and my second portion of my presentation, what is the challenge in getting there?

The major challenge in getting there is that the clinical research arena for major depressive disorder is an extremely difficult and technical arena. This is one of my favorite presentations, one of my favorite papers to cite when trying to illustrate this problem.

Essentially, this is an analysis from the FDA, from Khin and colleagues, looking at the following question. If you look at FDA-approved agents, how many studies in the registry are positive and how many are negative? Almost half of studies in FDA-approved drugs are negative. So that doesn't really count the number of drugs that we've completely lost that could have helped because of negative studies. So if you only look at the FDA-approved drugs, 50 percent of trials
fail. That is a very bad statistic. That is a 
sobering statistic.

So what is the response? We can't just 
watch this happen. What is the response? What's 
the reaction of the clinical research community to 
try to stay one step ahead, to go through 
thoughtful, carefully planned innovation?

Just to give you an example of how things 
have changed in the adjunctive treatment world, 
we've gone from simple randomization and that gave 
us quetiapine XR. In simple randomization, you 
especially take patients who you know have 
depression, who you know are on an antidepressant, 
and you know that that antidepressant has not 
worked by history.

So they report, or their doctors report, 
that this hasn't worked in these patients. That is 
what we call a retrospective assessment or 
confirmation of treatment, non-response or partial 
response.

So we've gone from that to a better design, 
which is a misnomer. It's called the placebo
run-in, even though patients at all times of the study are on an antidepressant and a dummy pill. This design gave us aripiprazole and brexpiprazole, two very useful agents, two atypical antipsychotics like quetiapine.

This was novel and considered revolutionary 10, 15 years ago. This is something we accept now as being part of the standard. So this gave us aripiprazole and brexpiprazole. Essentially, what happens is you have patients with depression. You treat them in the study with an antidepressant to confirm whether that works or not and, if it doesn't, you randomize them to aripiprazole, brexpiprazole, or placebo in these cases. This is what's called prospective confirmation of treatment, non-response or partial response.

A step ahead is what's called the SPCD design, which is kind of a blend of the two previous, and I'll explain how it refers and how it stands as a necessary improvement to its predecessors.

Let me put this in perspective. On the
slide here, you can see the predecessor design.

This is the antidepressant plus placebo lead-in design. So essentially, you have patients. You take them into the study. You put them on an antidepressant and adjunctive placebo. Those that do not respond are randomized to have the experimental treatment added -- back in the days of aripiprazole or, more recently, brexpiprazole -- or adjunctive placebo.

This is a great design, but has a serious limitation. The limitation here is that, depending on what study you look at, half of patients that come in or even up to three-quarters do not generate any data about the compound's efficacy, safety, or tolerability.

So you have patients coming in. They sign consent with the understanding that they're going to contribute to a potential new therapy, and their data is tossed out. This is an inefficient way of doing things. This really makes things more difficult from a practical standpoint. And one could argue if we can get away with -- if we can
evolve to use all patients' data, it would be a more ethical, in my opinion, study design because you would need fewer patients, and you would use the data from every patient who's consented with the intention of generating information that would generate a new treatment.

The solution to this is a simple tweak. Essentially, what you do is from the beginning, you add the possibility, then in addition to the antidepressant, patients get the experimental treatment as well.

Doing this, every single patient that comes into the study and randomized generates data about efficacy, safety, and tolerability. The patients that start in the study are retrospective partial non-responders, more similar to the quetiapine study. In clinical terms, this is similar to someone being referred to your practice who is on an antidepressant, who's depressed, and who hasn't gotten better.

On the right-hand side, you then move on to the more traditional design of 10 years ago, where
you have patients continuing through the study on
their antidepressant placebo and they are
prospectively confirmed as being partial non-
responders.

In clinical terms, this is more similar to
someone coming to your office, you treat them with
an antidepressant, you see what happens, and then
you decide on an adjunct. That is essentially the
change. The advantage here is that every patient
counts.

This is not a new approach. This approach,
first proposed in 2003, has been used not only in
major depressive disorder, in Alzheimer's, in
generalized anxiety disorder, in ADD and
schizophrenia. It has been used as a monotherapy
or an adjunct. It has been sponsored by companies,
by foundations like the Stanley Foundation, by the
NIMH several times. In fact, Dr. Mathews
conducted -- one of the speakers -- one of the
first SPCD NIMH studies and then many more. It's
recognized as a way of doing business.

So clinical trial challenges, we need to
move things along to get treatments for patients, demonstrating efficacy. The way we're doing things now is difficult, inefficient. We have to improve the way we do things to deliver treatments to patients who need them.

SPCD, which is essentially a blend of its two predecessors that have been used in approval process, in my mind is a more efficient way of doing things and allows for every patient and studies to count.

So with that, I'd like to conclude and invite the next speaker, Jerry Schindler from Alkermes, who is going to discuss the efficacy dataset.

Jerry?

**Applicant Presentation - Jerald Schindler**

DR. SCHINDLER: Thank you, Dr. Papakostas. I'm Jerry Schindler, and I lead the biostatistics group at Alkermes. And what I'd like to do is present the results of our four pivotal trials of BUP/SAM for the treatment of major depressive disorder.
These four pivotal trials were randomized, double-blind, and placebo-controlled, and I'll present data that show that 2 of the 4 trials met their primary endpoint. One study just missed the primary endpoint, but it still provided supportive evidence of efficacy. And 1 of the 4 pivotal trials did fail to meet its primary endpoint.

So first, an overview of the 4 pivotal trials. The program consisted of 1 phase 2 and 3 phase 3 trials. There were 3 SPCD designs and 1 placebo run-in design, and a 2/2 dose of BUP/SAM was evaluated in all studies. The primary assessments were based on either the HAM-D 17 scale or the MADRS scale.

Here's the SPCD design again, which you just saw. The highlighted areas showed the two randomizations, and the randomization in stage 1 was 2 to 2 to 9, with more patients being randomized to placebo to allow enough patients from stage 1 to roll over into stage 2. The randomization in stage 2 was even, 1 to 1 to 1, to evaluate the 3 groups in stage 2.
Remember, the formal statistical analyses are based on the randomizations. So the formal statistical analysis are the comparisons of the groups that are randomized in stage 1 and stage 2 combine together.

Now, one advantage of the SPCD design is that we can follow patients from stage 1 who received BUP/SAM into stage 2. For these patients, there's no formal comparison group, but we can still follow them anyway and see what happens, and I'll present some of the data for those patients as well a little bit later.

First, a brief mention of the statistical analysis. All of these studies were longitudinal trials with weekly efficacy assessments. To analyze the data, we used the MMRM model, which uses all available longitudinal data without imputation.

At each time point, the way this model works is that each time point calculates the mean change from baseline by treatment group and then compares this change from baseline between the two groups,
between BUP/SAM and placebo. And the statistical
test is to determine if this difference is equal to
zero or not.

Now, remember many of these, 3 of the 4
studies, are SPCD designs. So for the SPCD design,
we calculate the difference from stage 1 and the
difference from stage 2, and these prespecified
weights to combine the differences across both
stages for the trial. And the SPCD design, as
mentioned before, uses data from all the subjects
in the study.

One way to look at the data from the SPCD
design is to look at the end of treatment, and
that's a single time point, but you can look at the
end of treatment, and that's often done. And you
look at the end of treatment at the end of stage 1
and compare the difference between BUP/SAM and
placebo at the end of stage 1. And then look at
the difference between BUP/SAM and placebo at the
end of stage 2, and then combine these two
differences using prespecified weights.

Now, this is looking at the data at a single
time point, at the end of therapy. But we're really trying to improve the symptoms of depression in these patients, and what we'd really like to do is measure the effect over multiple time points.

We want to see what the benefit is for the patient over the course of treatment rather than just at that one single time point, so another way to look at the data, which is actually a little bit better way to look at the data, is to look at the data over multiple time points. And in this way, you do almost the same thing, but you do it over multiple time points.

So at stage 1, you look at the difference between BUP/SAM and placebo at multiple time points, at different weeks, and calculate that difference from the model to get the group different estimate from that model, and then combine those differences across stage 1 to get one estimate for stage 1 and do the same thing for stage 2.

Look at the differences between BUP/SAM and placebo in stage 2 and then combine those estimates
at the group level, not at the patient level, but at the group level from the model, combine those estimates and then get one estimate for stage 2. And now you have one estimate for stage 1 and one estimate for stage 2, and then you can combine the data the same way as before.

So statistically, the math is the same for either way to do it, but what this does is it looks at multiple time points and looks at the treatment effect over multiple time points. And it really provides a better estimate of the effect of the drug as perceived by the patient over the course of treatment. It also has the benefit of reducing some of the week-to-week variability in the data so that we get a more precise estimate of the treatment effect.

Now, let's look a little bit at the overall results. First, to orient you a little bit to this slide, this slide shows the difference between BUP/SAM 2/2 and placebo and their confidence intervals for the primary endpoints for each of the four pivotal trials. The zero vertical line shows
when the difference is equal to zero, when there's no difference between the two groups. Data on the left would favor BUP/SAM and data on the right would favor placebo.

You can see that all the mean values are on the left, which is the side that favors BUP/SAM, and for two of the confidence intervals, for the 202 study and for the 207 study, the entire confidence intervals are on the left-hand side and do not cross zero. These p-values are significant, so these two studies, 2 out of the 4 studies, met their primary endpoint, 202 and 207.

Study 205, the one in the middle, even though the mean value is on the left-hand side, the confidence interval just barely crosses zero, and the p-value is 0.1. So it was not significant, didn't meet its primary endpoint, although there's a suggestion that there might be something worth looking at for that study. Study 206, the placebo run-ins trial, clearly you can see the confidence interval crosses zero; the p-value was not significant 0.72.
What I'd like to do now is present the data for each study one at a time. Study 202, as you've heard already, was our first study. It was originally a phase 2 study because it was the first efficacy study, and it was an SPCD design. And it was a randomized, double-blind, placebo-controlled study, which was conducted under the same rigor as any phase 3 design. It evaluated 2 doses of BUP/SAM, 2/2 and 8/8. And the primary endpoint here was HAM-D 17, but we also collected MADRS-10 as a secondary endpoint.

So here are the results for study 202. On the top of this slide, just to orient you a little bit to the slide, are the changes from baseline. Gray is placebo and green is the group that received BUP/SAM 2/2. And we show stage 1. It's an SPCD design, so we show stage 1 and stage 2, and the evaluation of treatment effect was at week 4.

Then below this, just so you know what's on the slide, below this are the confidence intervals and the mean estimates for the comparison between each treatment group and placebo. You can see that...
there's clear separation between the patients who receive BUP/SAM, the green line, and the patients who receive placebo, the gray line, in both stage 1 and stage 2, pretty much starting by week 3 and week 4 in both groups.

When you look at the forest plot below, you see the statistical comparison of the difference between BUP/SAM 2/2 and placebo. And you see here the p-value is 0.014, so this is clearly a significant result.

When you look at the comparison between BUP/SAM 8/8 and placebo, the p-value is 0.7, 0.699, and the confidence interval clearly crosses zero, so this was not significant.

As you've heard, the FDA has raised two issues about this study, and I'd like to discuss them now. The first is that there was not a prespecified multiplicity adjustment, and the second was that one single patient might be driving these results. So what I'd like to do now is address some of these issues.

One note about study 202 is that there was
no prespecified multiplicity adjustment. Just to remind you what the p-values were from the previous slide, when you look at the comparison between BUP/SAM and placebo, the 2/2 dose, the p-value was significant, 0.014, and the 8/8 dose, the p-value is 0.699.

Now, there are only two treatment groups, so to apply a multiplicity adjustment, we're just doing multiplicity adjustment among two treatment groups. One way to do this is a post hoc Bonferroni multiplicity adjustment, which has the effect of splitting the alpha in half. So that changes the decision rule. It doesn't affect the p-value. It changes the decision rule from 0.05 to 0.025.

The advantage of a Bonferroni adjustment is that it provides two independent tests, so it splits the alpha error in half, so you have one half for one comparison, one half for the other comparison. It's very conservative, which we all know, because it splits the alpha in half, but it also requires no prior assumptions.
So it's assumption free and provides these two independent tests. And when we use a Bonferroni multiplicity adjustment here, the p-value of 0.014 is still significant, even with a much more strict decision rule, using 0.025 in the multiplicity adjustment.

Now, the FDA has suggested that potentially a hierarchical multiplicity adjustment would be of use here. Remember, this was the first efficacy study in BUP/SAM, so we didn't have evidence that either dose was better than the other dose at this point. Usually, when you apply a hierarchical adjustment, you usually have some information before you do the study that one dose might be better than the other.

We're also aware that there's a complex pharmacology of BUP/SAM because we're combining an agonist and an antagonist, and we had uncertain dose response. We might have an inverted U-shaped dose-response curve, where when you put the two together you might actually see that the dose response, rather than being increasing, as you
might expect in general, it may not be increasing. It may be an umbrella shape or an inverted U-shaped dose-response curve, in which case a hierarchical adjustment would definitely not be useful because we wouldn't know which one to put first.

So I don't see that there's any justification for wanting to use a hierarchical adjustment for this study, for the first efficacy study of BUP/SAM.

Now, the other comment that FDA made was that a single patient drives the results. And when we look at the data, what we see is that the patient that was selected as that single patient was the best responder from the study. So what FDA did is, looking at the data after the blind has been broken, the analysis has been done, and you've looked at the data, they've identified that patient that was the best responder from study 202 and taken that patient out.

Now you know when you take out the best responder, two things are going to happen. One is, the mean of the overall study is going to decrease
a little bit because you know the best responder is going to be contributing a bit to that mean. So the mean's going to decrease a little bit. You're also going to lose power if you take that patient out.

We also have looked into whether there's any legitimate reason to take that patient out, and the patient was appropriately randomized, appropriately selected for the trial, and a pertinent study was conducted in an appropriate, blinded, and unbiased manner. And we've gone back, and reviewed, and sent that documentation to the FDA. So there's no legitimate reason to take that patient out.

Removing a patient, as we all know, would be contrary to the intent-to-treat principle, so you wouldn't do that in general as a way to analyze the data. But it might be interesting to see what happens when you do take a patient out, and you look at the best responder, and do it as a type of sensitivity analysis.

So what I'd like to do is talk about it in terms of a sensitivity analysis, not so much as the
primary analysis, but talk about it as a sensitivity analysis. So down below, I have a little bit of the confidence intervals that I've shown before, the primary analysis, which I just showed, which is the mean value and the confidence interval of the 2/2 dose versus placebo where the p-value was 0.014. That was on the previous slide.

Now, from the FDA briefing book, their information is they took the patient out, which obviously has an effect on the mean, and you can see the mean gets a little bit smaller, gets closer to zero, and the confidence interval stretches a little bit because you're reducing the power of the study. All things you would expect; when you take the biggest responder out of a study, totally things that you'd expect, you'd reduce the power and you'd affect the mean.

But that's not the appropriate way to do a sensitivity analysis for this. If you want to see what is the effect of the best responder, the best way to do it is to take out the best responder and the worst responder, take out the top and the
bottom on both ends, and then replace that person that's missing, not decrease the size of the study, keep the study the same size so you have the same power, but replace those people with their nearest neighbor.

So you're bringing it in a little bit, but you're still doing an appropriate sensitivity analysis. And that's what the trim-and-replace one is, where you take the top and the bottom, the best responder and the worst responder, and then replace them with their nearest neighbor, and then do the analysis again.

You see, even when you replace the data with the nearest neighbors, you see basically the same thing. The result doesn't change very much. So the data really are not being driven by that one patient, but they'll be driven by the other patients in the trial. And the nearest neighbors are equal substitutes for that one patient.

Another way to look at the data is to take that patient out of the study, keep the sample size the same, but eliminate all of their data. If you
want to do a sensitivity analysis of that data, you eliminate all of their data and use multiple imputation to fill in for the missing data. But you still keep that person in the randomization scheme because they're properly randomized, and you want to make sure that you have the same amount of power.

So when you look at it as multiple imputation question -- you also do this as a sensitivity analysis -- you still see the same result. So the trim-and-replace method and the multiple imputation method, two sensitivity analyses, give essentially the same result as the primary analysis, which we presented earlier.

What that means is that that one single patient is not driving the result, that all of the patients together are contributing to the result, and that, overall, the result is robust.

So now, I'd like to look at the next study, study 205. Study 205 was another SPCD design. It looked at 2 doses, 2/2 and 0.5 and 0.5, and the primary endpoint here was MADRS-10.
Here are the results for study 205, same display as before with the changes from baseline in stage 1 and stage 2 at the top, and green is the BUP/SAM 2/2 dose, and below are the mean values and the confidence intervals for the comparison for the difference between BUP/SAM 2/2 and placebo and BUP/SAM 0.5/0.5 and placebo at the bottom.

What you see here when you look at that is that you see clear separation between the curves across the time points really starting after week 2 in stage 1 as before and pretty much throughout all of stage 2.

So you see clear separation between the curves, that there's greater reduction in depression severity for the group that receives BUP/SAM 2/2 relative to placebo.

Now, our comparison for this study, our prespecified primary endpoint, was the difference at week 5. And if you look at stage 2, you see a little bit of variability in the data, which you know we're going to see in this type of trial. And in week 5, the two curves comes together just a
little bit, especially in stage 2, and you can see that on the graph.

When you do the analysis, you see exactly the same thing. The primary endpoint is the change at week 5. And you can see that the difference between the two treatment groups, between BUP/SAM 2/2 and placebo, even though the mean value is on the left-hand side, which favors BUP/SAM, the confidence interval just crosses zero. So the p-value is 0.1 and not significant.

You see the same thing, only it's even more extreme. When you look at the 0.5/0.5 comparison, the mean value is around zero of the difference and the confidence interval clearly is on both sides, and the p-value is 0.9. It's clearly not significant for the 0.5/0.5 dose.

But the 2/2 dose just barely missed, and the reason it missed, you see it's some variability in the data. Especially you see the variability at week 5, where the points come together.

Now, the FDA briefing book ignored study 205 completely because it didn't meet the primary
endpoint. Even though the difference was not significant at that single time point at week 5, the data clearly show improvement in depression symptoms across all of both stage 1 and stage 2, and the 2/2 dose clearly looks like there's some difference from placebo in this graph.

So I believe that this study really does provide supportive evidence of efficacy. It didn't meet its primary endpoint, but it provides supportive evidence of efficacy.

Now, we learned some things after we did study 205. One of the things that we learned, when we look at 205 and 202 together, we learned that we saw a clear signal that there was some activity for BUP/SAM 2/2. So we saw early indications of efficacy for BUP/SAM 2/2.

We also realized that we had variability in our data, and we wanted to address that, and we also wanted to be able to address the patient's experience going through the trials.

So there are two ways to do it, two ways to look at the data. One way is to continue to look
at that single time point, and then we are at the mercy of the variability of the data. The other way to do it is to average the difference between the two groups at multiple time points.

The advantage, as I mentioned earlier, when you average the differences together, the advantage you get is that you reduce some of the influence of the week-to-week variability. You also get a more precise estimate of the treatment effect. And most importantly, you get the patient's experience over time, so you see what happens from the experience of the patient over multiple weeks rather than at one single time point. We wanted to pursue that in our future studies.

We also saw in the literature that there was some question that MADRS-6 might be a more sensitive measure of depression and might measure the core symptoms of depression severity better than MADRS-10. And MADRS-6 is just a subset of MADRS-10, so if we collect MADRS-10 data, we get MADRS-6 also. So we can look at both. We can look at MADRS-10 because we have the data for that, and
then we can look at the subset of MADRS-10, MADRS-6.

So we wanted to explore both MADRS-6 and MADRS-10 in our future trial, so we did that in study 207.

Here's study 207. It's another SPCD design. It looked at 2 doses of BUP/SAM, 2/2 and 1/1, and it collected MADRS-10, which also gave us MADRS-6. So we had the ability to look at both MADRS-10 and MADRS-6.

For this study, the basis for concluding efficacy was based on MADRS-6, where the difference was averaged over multiple time points, week 3 through end of treatment. That was the basis for concluding efficacy, which was prespecified in the protocol. But we also included MADRS-10 because we collected that data, and we averaged those differences over multiple weeks, the same weeks, week 3 to the end of treatment. So we average the differences from MADRS-10, and we included that as an endpoint.

We also had MADRS-10, so we could look at
the MADRS-10 at that single time point at the end of treatment. So we now had three endpoints that we could look at, although in our protocol, we prespecified the basis for concluding efficacy would be MADRS-6, where the difference was averaged over these multiple time points.

We did that in a hierarchical manner, looking at BUP/SAM 2/2 first, and then looking at the difference between 1/1 and placebo second. And we did a hierarchical multiplicity adjustment here because we already had evidence that the 2/2 dose seemed to be effective, and we wanted to get more evidence of that.

So here, a hierarchical multiplicity made sense where it did not make sense in the 202 study. Here, it does make sense, so we used a hierarchical multiplicity adjustment.

Here's the data for study 207, same format as before, where you see the data, the change from baseline at the top for stage 1 and stage 2. And here, you see clear separation between the two groups. The green is BUP/SAM 2/2. The gray is
placebo. And you see clear separation across all of the time points in stage 1, across all the time points in stage 2. The highlighted areas are the weeks that we incorporated into the analysis. We averaged the difference at those time points and included those in the analysis.

At the bottom, you see the hierarchy of the endpoints. The top one is MADRS-6, which was averaged over multiple time points, and the p-value for that one was 0.015. So that was the basis for concluding efficacy, so right away we know that this was a positive trial. It met its primary endpoint.

But we also wanted to continue the path down the hierarchy, so the next step in the hierarchy is MADRS-10, looking at the average difference for multiple time points, so MADRS-10, the average over multiple time points. And the p-value for that was also significant, 0.026. So both of these showed a significant difference from placebo for the 2/2 dose.

When we keep stepping through the hierarchy,
MADRS-10 at the end of treatment just barely missed, .07, and the p-value there extended a little bit over zero, and the p-value was not significant. So that stops our pathway through the hierarchy.

When we continue to look at the other data, because we can do it and we have it on the slide, the BUP/SAM 1/1 dose would not have been significant anyway, but we never got to that in the hierarchy. The only ones we got to were the top 3, and then we stopped when we got to MADRS-10 end of treatment. But the top 2 were significant, so this study clearly met its primary endpoint.

Now, when you look at the data, especially look at stage 1, we know we have variability in the data. And if you look at the data in stage 1 at week 5, you see a slight blip in the data, especially for the BUP/SAM group, where the curves start to come together a little bit. Also, the BUP/SAM group starts to look more like placebo.

You might have the question, is this indicative of potential waning of effect? The
treatment effect comes down for a few weeks and then starts to go back up. And that was something that we wanted to explore, and we can do it with the SPCD design.

In this slide, we actually followed these same patients into stage 2. What we're looking at -- and just to remind you what you're looking at -- the gray line is the placebo; the green line is the patients who receive BUP/SAM 2/2, and followed them from stage 1 to stage 2.

Now remember, the people who received placebo in stage 1, many of them are re-randomized into stage 2. So we're only showing the data for placebo up until the end of stage 1. But the patients who received BUP/SAM 2/2 in stage 1 continued to receive BUP/SAM in stage 2.

This is all blinded. The patients aren't aware of their treatment, and they're not aware of where they are in the study. So this is all completely blind data. It's unblinded now, but it was unblinded during the course of the study.

You can see that there is that little blip
at week 5, but then that data then come back and
return, and you can see a consistent reduction in
depression severity all the way through stage 2.
So this blip was just a little bit of statistical
wobble, and then it returned back to the path that
it was seeing before. And it is not any indication
of any waning of effect.

Now, we did another study as well.
Study 206 was a placebo run-in design, different
design, placebo run-in design, and it looked at
1 dose of BUP/SAM, the 2/2 dose versus placebo.
The primary endpoint for this study was MADRS-10,
evaluated at the end of treatment.

Here are the results for study 206. Here,
you can see that the two lines are almost exactly
on top of each other, and the difference between
BUP/SAM and placebo is almost zero. And the
confidence interval clearly crosses zero. So this
study was not significant. This study did not meet
its primary endpoint.

Now, one way to compare the data across all
four studies is to use a common endpoint, and what
I'd like to do is compare these studies and look at all four studies. The advantage that we have is that all four studies included weekly MADRS-10 assessments. So we collected MADRS-10 in every study, and we can use MADRS-10 to look at the data across all four studies.

Now, as I mentioned before, there are two ways to look at MADRS-10 in a statistical way. The one way, which we prefer and really is probably the better way, is to look at MADRS-10 based on the difference between the two groups at multiple time points.

As I mentioned before, you calculate the difference at multiple time points from the model and average them together to get the overall estimate of the mean effect. What this does is it reduces the impact of the week-to-week variability of the data and also better reflects the patient's experience over time.

The other way to look at the data is the more traditional way, to just look at the end of treatment, just that single time point. And we can
do both. We've got the data. We've got MADRS-10
data, and we can look at both. And the end of
treatment, as I mentioned, is a more conventional
way to look at the data, but it has some of these
other issues.

So just to look at the data again, this is
MADRS-10 data for these three studies; study 202,
205, and 207; stage 1 first and then stage 2 for
each study, so you see 6 randomizations.

What you see is that all three studies show
clear separation between the treatment groups, and
they show a clear improvement in depression
symptoms for the BUP/SAM 2/2 dose relative to
placebo. And you can see that in every single
study, in every single randomization.

It's remarkably consistent across all six
randomizations. And this consistency is really
what's driving the results. It's not driven by
just one patient, as you may have heard in the
erlier presentation. It's not driven by one
patient. It's driven by all the patients. It's
driven by many of the patients. It's driven by the
patients in all of these trials. And we see similar effects in each stage and for each study across the entire program.

So when we look at the statistical analyses of the data -- what you saw before was the data, looking at MADRS-10. So here's the statistical analysis of the MADRS-10 data. And this analysis looks at MADRS-10, where we average the difference over multiple time points.

Here, we see that 3 of the 4 studies show confidence intervals that don't cross zero, so here now, 202, 205, and 207, three, when we look at MADRS-10, average over multiple time points, and confidence intervals don't cross zero.

The 206 study, which I showed you that didn't meet its primary endpoint, gives you the same result here. When you look at MADRS-10 average over multiple time points, 206 was also not significant, but 3 of the 4 were.

One way that is interesting to look at the data is to do a meta-analysis of the data. We recognize that three of the studies are SPCD
design. One's a placebo run-in. So it's not a perfect meta-analysis. And we're not using this as the primary endpoint of our trial, but it's still interesting because we have the data. It's interesting to look at what you get when you put all the data together and what kind of pattern do you see. So we have that for comparison.

What you see when you look at the meta-analysis, looking at MADRS-10 average, what you see is that the mean value is clearly on the side that favors BUP/SAM, and the confidence interval for the meta-analysis also doesn't cross zero, also clearly it doesn't cross zero, and also is on the side that favors BUP/SAM relative to placebo.

So all of these taken together, the MADRS-10 average, looked for in each of these studies, and the meta-analysis show that the 2/2 dose is better than placebo. But now, there's that other way to look at the data, looking just at the end of treatment, so we have that as well.

When you look at the data just at the end of
treatment, and here are the confidence intervals for that with the mean values and the confidence intervals, here you see that 2 of the 4 studies show confidence intervals that don't cross zero. Study 202 and 205, the confidence intervals don't cross zero. 207, the confidence interval just barely crosses zero.

Remember, we saw before that there's a little blip at week 5, and that's what's responsible here for that confidence interval crossing zero; and 206, again, the placebo run-in design that failed to meet its primary endpoint is not significant here, too.

The same story with a meta-analysis is used as a type of sensitivity or supportive analysis. It's really not the primary analysis, but we have four studies, and we want to show what happens when you combine them all into one statistic. So the meta-analysis is not our primary analysis for this program at all, but it really is a comparison of how do we bring all the data together, even though we have SPCD designs in parallel and a placebo...
run-in design.

So we don't have a perfect meta-analysis, but it's still something that's interesting to take a look at. We see the same result as we saw before, that the mean value is on the side that favors BUP/SAM, and the confidence interval here doesn't cross zero. So here's another indication that the BUP/SAM 2/2 dose is better than placebo.

So overall, what did we learn from our program? Overall, these trials showed that the BUP/SAM 2/2 dose is effective for the adjunctive treatment of depression. I've just shown you these slides before, the data before. What I'd like to show is that we've looked at the data many different ways.

First, at the top, we looked at the prespecified primary endpoints, and these were the endpoints that were prespecified in the protocol. And when we look at the prespecified primary endpoints, 2 of the 4 trials met their primary endpoint. And we see strong evidence of efficacy from a third study, study 205, which we mentioned...
earlier.

When we look at MADRS-10 difference, which is averaged across multiple time points, we see that 3 of the 4 studies demonstrate evidence of efficacy. And then, when we look at MADRS-10 at the end of treatment, at that single time point, we see that 2 of the 4 studies show evidence of efficacy favoring BUP/SAM over placebo.

Taken together from every direction, we presented data that showed that the 2/2 dose of BUP/SAM is effective for the adjunctive treatment of major depressive disorder.

Now, I'd like to introduce Dr. Gary Bloomgren, who will talk about drug safety.

**Applicant Presentation - Gary Bloomgren**

DR. BLOOMGREN: Thank you, Dr. Schindler.

Good morning, everyone. I'm Gary Bloomgren, head of drug safety at Alkermes, and I'm going to be providing the overview of the clinical safety program of BUP/SAM.

Our integrated assessment of the safety profile of BUP/SAM is based on data that we
collected from over 2,100 subjects treated with at least 1 dose of BUP/SAM. And of that, more than 1500 included patients with major depressive disorder treated on the 2/2 dose. And of those, more than 700 received treatment for in excess of 12 months.

This totals greater than 1100 patient-years of MDD patient exposure, which I'll be sharing details on as we go through the presentation.

Now, the safety data was pooled across 4 placebo-controlled studies, with each study having 2 randomizations, the first stage 1 at the beginning of treatment. And the second mid-study, when placebo non-responders were re-randomized for stage 2, the data that I'll be sharing with you on a number of tables that follow are summarized by those randomization time points.

Treatment-emergent adverse events across the development program were generally tolerability related. And what you see in this table is the treatment-emergent adverse events for BUP/SAM 2/2 that occurred at an incidence greater than or equal
to 5 percent and greater than placebo. And what
you see is the stage 1 to the left and then the
stage 2 randomizations and the data.

What you see is the rates of the TEAs were
less in stage 2 than in stage 1, but the events
were generally similar sorts of events. The TEAs
that we saw were generally gastrointestinal related
or associated with sedation.

Across the program, the majority of adverse
events were mild to moderate in severity and tended
to occur with treatment initiation. We saw no
meaningful differences by gender, age, race,
concomitant antidepressant used, or benzodiazepine
use subgroup. And there are no new findings that
we saw within the long-term study that went over a
year duration.

There were a few adverse events that led to
discontinuation across the program. And in this
table, you see those treatment-related adverse
events that led to discontinuation in an incidence
of greater than 2 percent. And again, with
stage 1, the incidence was a little higher with
BUP/SAM, 13.6 percent of subjects, compared with 2 percent on placebo, a significantly lower incidence of 3.8 and 1.4, respectively, for BUP/SAM and placebo in stage 2.

There were only 3 adverse events that occurred in an incidence of greater than 2 percent that led to discontinuation, and that included nausea, vomiting, and dizziness shown here, and that was only in the stage 1. We saw similar findings in the long-term study, and that appeared very similar to what we see here with stage 1.

There are few serious adverse events across the development program. In the placebo-controlled studies, 1.9 percent of the BUP/SAM versus 0.5 of the placebo patients had an SAE. There was no pattern to these SAEs and there were no deaths.

In the 1-year long-term study, 3.2 percent of patients had an SAE. The most common events were depression and suicidal ideation, each which occurred at an incidence of 0.2 percent.

There were 2 deaths in the long-term study, both assessed by the principal investigator as not
related. The first was a patient with chronic obstructive pulmonary disease, who died of respiratory arrest 47 days after the last dose of drug, and the second was a patient with hypertension and congestive heart failure who died of a cerebral hemorrhage on day 87 of study treatment.

There were no clinically relevant laboratory vital sign, weight or ECG changes in either the placebo-controlled or long-term studies. We saw no meaningful post-baseline changes or outliers of major clinical significance. And there was no evidence of risk of QT prolongation associated with BUP/SAM in a dedicated TQT study.

I'm going to talk now about some topics of special interest, and they really come from three areas; one either safety topics that were identified within the BUP/SAM program itself, those that are associated with buprenorphine as a safety concern by itself, or by other antidepressant treatments.

First, as it relates to the topics of
special interest associated with the BUP/SAM program, we saw CNS sedation as a major finding. This was mild to moderate in nature, typically was associated with treatment initiation, and resolved with continued use.

The other was one case of acute opioid withdrawal, which was precipitated with the first dose of BUP/SAM in a patient who had undisclosed pre-existing opioid dependence. This event was serious and assessed as related to treatment, and attributed to samidorphan, the mu opioid antagonist that's part of BUP/SAM, and is the reason why we are contraindicating its use in any population of patients who would either be chronically using opioids or are opioid dependent.

Additional topics of special interest included those topics that are associated with buprenorphine alone, and that includes respiratory depression, hypotension, orthostatic hypotension, and hepatic injury. And we saw no evidence by review of adverse events, changes in vital signs from baseline, or laboratory tests as appropriate.
that there was any change in this group relative to
the placebo plus ADT cohort.

Additionally, when we looked for items that
are potential safety concerns with other
antidepressant agents, including hypomania, mania,
sexual dysfunction, or suicidal ideation or
behavior, we again saw no difference in terms of
the AE profiles that we saw on treatment with
BUP/SAM plus ADT versus what we saw with placebo
plus ADT.

We also did a very thorough review looking
for suicidal ideation and behavior using the
Columbia-Suicide Severity Rating Scale or the
CSSRS, and what you see here is we saw less
suicidal ideation and behavior with BUP/SAM
compared with placebo. And this occurred with both
randomizations and had a similar finding in a
long-term study.

What you see in this table are the post-
baseline CSSRS events on treatment for each of the
stages for the suicidal ideation, suicidal
behavior, and the self-injurious behavior without
suicidal intent.

What you see is, numerically, these events are less on the BUP/SAM 2/2 than on placebo in both randomizations and that the prevalence of these events in a 1-year long-term study were very similar to what we saw with the BUP/SAM in the controlled studies.

Because we added samidorphan specifically to buprenorphine to mitigate the abuse potential, this was an assessment of abuse potential. It was a critical part of the development program, and I'll be sharing the data that we have collected across that as a part of our integrated assessment.

This included a dedicated human abuse potential study, which I'll be going into more detail shortly, and it also included, over 1500 patients with major depressive disorder that we studied, an evaluation of adverse events of special interest that were related to abuse potential, dependence, and withdrawal. We also did an objective assessment of withdrawal using the COWS or Clinical Opioid Withdrawal Scale following
abrupt discontinuation of BUP/SAM.

The dedicated human abuse potential study; our assessment was that the abuse potential is low. And what you see here in the human abuse potential study is a double-blinded, cross-over, 6-way study using non-dependent recreational opioid users and assessed treatments with placebo: the 2/2 dose, the indicated dose, supratherapeutic doses of BUP/SAM 4 and 8 times greater than the daily dose for BUP/SAM, as well as buprenorphine alone, 8 and 16 milligrams, which were used as positive controls.

The primary endpoint of this study is, at the moment, drug liking difference from placebo. And what you see with the dotted line at the 11 within the figure is the margin of clinical significance for BUP/SAM versus placebo.

This is established through multiple HAP studies looking at the placebo treatment groups and looking at what the upper margin of significance for placebo was across those studies. So anything to the left of that dotted line is consistent with
what can be seen with placebo.

What you see is the 2/2 dose is consistent with that of placebo. And even the supratherapeutic doses, the 8/8 and 16/16, were slightly greater than placebo but substantially less than that of the buprenorphine corresponding dose alone that was part of a positive control.

We also had a number of secondary endpoints that were part of this study, which included overall drug-liking and take-drug-again assessments. Both of these endpoints are highly predictive of real-world abuse liability, and it shows that BUP/SAM is similar to placebo. And that's not only with the 2/2 dose, but even the supratherapeutic voices. And this is substantially different from what we see with the active positive control of BUP alone, further supporting our assessment that there's a low abuse potential with BUP/SAM.

We also saw consistent evidence of low abuse potential across the MDD studies, and this, again, is with over 1500 patients with MDD. We curated
that dataset for abuse potential terms. And the
majority of the events were nonspecific to abuse
potential. This included dizziness, somnolence,
sedation.

Although these can be seen with abuse
potential drugs, they're commonly seen with drugs
that have no abuse potential at all. What's more
important is there was a low incidence of euphoria-
related events. With BUP/SAM 2/2, the incidence
was 1.6 percent versus placebo at 0.2 percent, and
that's across stage 1 and stage 2 combined.

The incidence in the long-term study -- now,
this is a year-duration study. The incidence was
very similar, 1.2 percent. The majority of these
events were associated with the first dose of
treatment, and none of them reoccurred.

The other important thing to note is that we
did not see a dose effect within the MDD
population. Half of the events of euphoria
occurred with the initiation with a 0.5/0.5 dose
with titration, the other half with a 2/2 dose.
And there were none on the 8/8 dose. We saw no
abuse behavior and no evidence of dependence across the development program.

We also did a thorough assessment for withdrawal and saw little evidence of withdrawal. And in these analyses, we took all patients that had a minimum of 4 weeks being on treatment, and then looked at the 2 weeks following abrupt withdrawal of treatment to see if we saw any evidence of withdrawal, and this slide summarizes those findings.

Our mean post-discontinuation scores in that 2-weeks following discontinuation in that population, was less than or equal to 1 on the COWS scores. COWS scores, no withdraw, are from 0 to 4, and that occurred in all treatment groups, whether it be placebo or active. And that was also seen in the long-term study.

If we look at the shift and mainly at the outliers that are really summarized, then, within the COWS scores table below, you see that in the placebo-controlled studies, BUP/SAM 2/2 and placebo, that 96 and 97 percent of patients had no
withdrawal; 2.7 of both BUP/SAM 2/2 and placebo had mild withdrawal. And there was 1 patient which made up the 0.9 percent in the controlled studies that had moderate withdrawal by COWS score.

In the long-term study, we saw similar findings. 94 percent of patients had no withdrawal by COWS score, and 4.9 and 0.7, respectively, had mild or moderate withdrawal by COWS scores.

Now, we looked at these 5.6 percent, the 4.9 plus the 0.7 patients, to see what we could assess clinically as to the meaningfulness of this. And what we found is that this accounts for 47 patients within the study. And within these 47 patients, 44 of those 47 had no or either required any prescription medication treatment whatsoever.

There were 3 patients out of the 47 that made up that 5.6 percent, and those 2 had anxiety; 1 had insomnia. Each of them required a benzodiazepine prescription, and I think puts in some context, at least, with what we did see in terms of findings as it relates to potential withdrawals, that this is easily medically managed.
In summary, we feel that we've done a very thorough assessment of the safety profile of BUP/SAM. The common AEs we saw were gastrointestinal and sedation related, mild to moderate in severity, and they typically occurred with treatment initiation.

There was no clinically meaningful changes in laboratory vital signs, weight, or ECG changes across the program, and we saw no evidence of increased treatment-emergent suicidal ideation or behavior.

Additionally, as it relates to our assessments across many facets for abuse potential, our assessment is the abuse potential of BUP/SAM is low. In the human abuse potential study, the abuse potential of BUP/SAM 2/2 is similar to placebo, and even supratherapeutic doses were slightly greater than placebo, but significantly less than the equivalent dose of the positive control for BUP alone.

Across the 1500 patients with MDD that we studied, the data was consistent with these
findings. There was a low incidence of euphoria typically with the first dose, none of which were recurrent, and there was no evidence of dependence during treatment and little evidence of withdrawal even with abrupt discontinuation, and that that discontinuation was well tolerated.

Now, as I've shared with you over the last several slides, our assessment of the abuse potential of BUP/SAM is low. But that said, BUP/SAM contains buprenorphine, and particularly during an opioid crisis, we want to ensure and we are committed to make every effort that BUP/SAM is used in an informed and appropriate manner.

This starts with how we've developed the product, and by selecting the lowest effective dose of BUP 2 milligrams and added samidorphan to mitigate that abuse potential, we've co-formulated the tablet as a monolayered tablet that's micronized to a homogenous mixture of BUP/SAM that cannot be mechanically separated. The tablets are packaged in cards in individual blisters, as Dr. von Moltke had shared
with you a photo of earlier in the presentation. And this is using F1 packaging, which is the highest quality packaging, to limit unintentional pediatric exposure.

We're committed to continuous monitoring of the distribution system from manufacturer, to wholesaler to pharmacy, to monitor for any evidence off any suspicious activity that might suggest diversion.

We're also very much committed to educating healthcare providers and patients to the appropriate use of this product in the marketplace. And we're also committed to implementing additional post-marketing safety initiatives, which includes a REMS, which I'll be speaking to briefly.

A centerpiece to our risk mitigation is education. The objective of this is to alert healthcare professionals and patients to important safety information and appropriate use. This includes the risks and precautions associated with concomitant use of opioids, either prior to considering initiating treatment as well as during
treatment.

It includes instructions as it relates to safe use, storage, and disposal, and because BUP/SAM contains buprenorphine, also alerting to the risks of abuse, misuse, diversion, addiction, overdose, and death that can occur with opioids.

As it relates to information to go to healthcare professionals on the left panel, that information will be predicated on the details that are part of the product label as well as the product REMS. And it will include selection of appropriate patients for this sort of therapy, but also educational materials that include patient counseling and a placebo training tool, since this would be the first sublingual delivered medication in this sort of a population.

It includes a commitment for trained medical staff available to respond to queries via a call center or in person and a website and interactive web-based training for physicians, whether they be in rural or urban areas. And as it relates to patients, a medication guide and a medication I.D.
wallet card that ensures that they have the appropriate information so that they can share this not only with family members, but also with other healthcare professionals, who may be treating them with other medications.

It also includes additional product website materials to help educate patients around the important risks and precautions that are associated with this therapy.

As it relates to a risk evaluation and mitigation strategy or REMS program, we've done a lot of thinking about this, and we've modeled this after other buprenorphine-containing therapies. The objectives of this REMS would be to mitigate the risk of misuse, which includes abuse and accidental exposure.

It includes a medication guide for patients, a REMS website, and call center for healthcare professionals, communication materials, which include the healthcare HCP letter which would go out at the time of a product launch, to ensure that we get out in front of prescribers with this
important information, HCP brochure and appropriate use checklist.

We're also committing to product-specific active monitoring for misuse, abuse, dependence, or diversion using RADARS. This is a nationally accepted, multifaceted approach for monitoring that's used by FDA as well as many other manufacturers in this space to assess for any evidence of end user inappropriate use of BUP/SAM specifically.

We're committed to identifying new healthcare professionals on an annual basis and ensuring that they have the information that I've shared with you above. And of course, with any REMS, we will be providing periodic assessments of the effectiveness of the REMS. And based on new data, if there is new information that suggests that we need to take a different stance, that will be a data-driven discussion and decision that we will be making with the FDA to make sure that the product is used appropriately.

Now, I'd like to turn the next stage of the
presentation over to Dr. Sanjay Mathew.

**Applicant Presentation - Sanjay Mathew**

DR. MATHEW: Thank you, Dr. Bloomgren.

My name is Sanjay Mathew. I'm a professor of psychiatry at Baylor College of Medicine, and I direct the mood and anxiety disorders program. And I'm here today as a consultant to Alkermes. I've been compensated for time and travel.

I'll be speaking as a practicing psychiatrist from a clinical perspective and give my assessment of the benefit-risk profile for BUP/SAM, and in more general terms, who would be an appropriate candidate for an adjunctive therapy.

These would be patients usually with longstanding depression. They may have had multiple episodes. They've had little or no success with successive monotherapy antidepressants, perhaps from a variety of classes, SSRIs, SNRIs, bupropion, despite being dosed at adequate doses and for adequate durations.

The current antidepressant that they're on has helped them, but certainly has not enabled them
to achieve remission. And they have persistent symptoms that go beyond the 9 symptoms of a major depressive episode, to include significant impairment in their social, occupational, and just in their family functioning.

Finally, a patient has to be willing to consider adding on another therapy. They've tried perhaps several monotherapies, but now it's time perhaps to add something on to the existing antidepressant.

So a typical case for a person where BUP/SAM may be considered would be a patient like J.D. She's in her mid-40s. She's experiencing her second major depressive episode. Her first episode was in her 20s, and she responded at that time to an SSRI, fluoxetine.

She's been in this episode for approximately 9 months. The first trial she had was an SSRI, escitalopram, which was pushed to the FDA max dose of 20 milligrams for about 3 months. And at baseline, she had a high very severe score on the Quick Inventory of Depressive Symptoms scale, which
is a self-report instrument for depression, a score of 21.

So with the escitalopram, she did have some improvement. Scores came down to 16. But that's still a severe score. So at that point, she was switched to an SNRI, venlafaxine, and that dose was pushed to a fairly high dose, 300 milligrams, for about 3 months, and her scores hovered at the 15-16 range, so there was no significant improvement, and she had some side effects with the venlafaxine.

We discussed what would be the options at this point, given the fact she had persistent depressive symptoms that continued to impact her ability to function. She's a mother of two school-aged children. She really became socially isolated, stopped going to church. She stopped preparing their lunches, and really stopped engaging with family and friends in her community. So beyond the depressive symptoms, there was this marked social withdrawal.

When we think about what is available for patients like J.D., we think of the atypical
antipsychotics. These are the on-label options. The most recently approved one was brexpiprazole, and here is a graph just looking at the comparison of the effect sizes between BUP/SAM across the 4 studies that were presented this morning and brexpiprazole, essentially a comparable effect size. And BUP/SAM does so through an entirely different mechanism, so it's not a "me too" approach. It's certainly not a monoaminergic approach.

So we would consider initially augmentation with what's available; aripiprazole, quetiapine, perhaps brexpiprazole. However, in this case, one of the main concerns for J.D. is she absolutely does not want to have an agent that could contribute to weight gain. She's struggled with weight most of her life. Her current BMI is 30, and she's a borderline diabetic. She was also worried about daytime sleepiness and somnolence, as one of the symptoms is residual fatigue.

So looking overall at the safety summary from these studies that were presented, it appeared
the common adverse events were GI or sedation. The sedation appeared to be limited and usually upon treatment initiation, and was not a reason for drop-out. And these were generally mild or moderate in severity.

The low potential for abuse, of course, is an important consideration, and this was evidence from both the human abuse potential study and from the patients with MDD, both short-term as well as the long-term 12-month study.

So overall, my assessment of the benefit-risk profile in terms of what are the risks; I mean, clearly, with buprenorphine, has that issue been mitigated with the presence of samidorphan, the mu antagonist, and it appears it has. The abuse liability significantly mitigated.

The adverse effects that were reported in the studies, mainly G.I., appeared to be manageable, transient, and did not lead to significant drop-outs. There was a lack of weight gain, metabolism impact, and movement side effects, which certainly would present a contrast to the
existing therapies.

Then with any antidepressant therapy, you're worried about induction of hypomania and suicidality, and that did not appear to be the case in the studies.

In the benefit considerations, of course you want something that's effective robustly, and the studies here showed consistency of those effects across 3 of the 4 studies, clinically meaningful efficacy in these difficult-to-treat patients, recognizing these patients have had multiple episodes, have been in this episode for many months, and have high baseline scores on the MADRS, and then the efficacy consistent as mentioned with what's available.

So in conclusion, then, in my view, this drug has a positive benefit-risk profile. Its efficacy is comparable to what is available in terms of adjunctive therapies, but offers a new distinct mechanism of action. It has a favorable safety profile, and in my opinion, patients like J.D. need these options. And as clinicians, we
need to have additional options for our patients.

Thank you.

**Applicant Presentation - Lisa von Moltke**

DR. VON MOLTKE: Thank you, and we appreciate the attention to the presentations. And I want to close by briefly summarizing some key points from this morning.

First, there is a significant unmet need for patients with MDD, and there's a specific need for therapies that are working via a new mechanism. BUP/SAM works via opioid system modulation and would bring a new mechanism of action to this indication. This therapy maintains the antidepressant activity of BUP while mitigating for risk of abuse and dependence.

You've heard about the significant challenges in studying MDD and the utilization of SPCD to meet those challenges, and you've seen data today from three important studies indicating that there is a positive benefit-risk profile for BUP/SAM. There is substantial evidence of efficacy comparable to other adjunctive therapies, and
there's a well-characterized and manageable safety profile.

Finally, Alkermes is committed to extensive education and monitoring to ensure appropriate use, and with that, we'll close the sponsor's presentation. Thank you.

DR. NARENDRAN: Thank you.

If we could have the people who came to the table introduce yourselves, Dr. Temple and Dr. Lee.

DR. LEE: Daniel Lee, clinical reviewer.

DR. TEMPLE: Bob Temple, deputy director of ODE I.

Clarifying Questions to the Applicant

DR. NARENDRAN: If we could now open the session for clarifying questions. We're already running a little bit late, so I'd like to remind the panel to be very focused in your questions. It's really just clarifying questions, not to engage in a discussion. Be cognizant of the other panel members who may also want to ask questions.

So clarifying questions, very brief, try to get the answer; please do not engage in a
discussion. There will be plenty of time for a
discussion later.

Dr. Warholak, your question?

DR. WARHOLAK: So my first question, I'm not
really sure who to address to because I have a
question about the clinical outcomes assessment
used as an endpoint. So I would address that
usually to a psychometrician, but I don't know if
there's one over there.

My question is, why did you decide to use
the MADRS-6, and what evidence of content validity
does that have compared to the 10? Also, can you
also give us some evidence about using the average
over the duration of the study as opposed to the
endpoint?

DR. VON MOLTKE: Sure. I'll ask two
different colleagues to come up and address that.
The MADRS-6 was evident in the literature, and
that's where we got our interest in it. I'll ask
Dr. Pathak to come up and walk you through some of
the differences. And then, with regard to
averaging, I'll ask one of our statisticians as
well as one of our clinicians to come up.

DR. PATHAK: Sanjeev Pathak, psychiatrist, Alkermes. Thanks for that question and the opportunity to clarify. We picked up MADRS-6 based on emerging literature, and the literature demonstrates that it is sensitive in evaluation of depression symptoms and improvement. And also recent literature suggests that it may be a good measure for assessing adjunct treatment when every patient is receiving a background antidepressant, where some of the symptoms may have been addressed. And the validation was published way back in 2002 by Beck [ph] and colleagues.

DR. SCHINDLER: Could I ask you just to clarify? I'm not sure I totally understood your question of evidence for averaging.

DR. WARHOLAK: Yes. It appears that the usual way to look at the -- it is the MADRS-10, and it's at the end of the therapeutic evaluation as opposed to averaging every week over the study period. So what comparative evidence do you have that that's a valid approach?
DR. MATHEW: Just to clarify, we're not averaging at the patient level. We're averaging at the group mean level, so we're using the MMRM model to estimate the treatment effect at each week, and then averaging that treatment effect among multiple weeks.

That's a method that's been used before. Coming as a statistician, not a psychometrician, when you calculate data using an MMRM model that is longitudinal data over multiple weeks, one method is to look at the difference over multiple weeks, and then average those estimates together to stabilize the estimate of treatment effect. And that's a method that's been used before.

DR. SCHINDLER: I'll just add, from a clinical perspective, looking at the average over time makes perfect sense because that's exactly what we're doing when we make treatment decisions. We're not basing it on an end of treatment because there really is no end of treatment. We're continuing to assess patients longitudinally over time.
DR. NARENDRAN: Next question, Ms. Grant?

MS. JONIAK-GRANT: Elizabeth Joniak-Grant. There's a lot of discussion here in changes from baseline or the difference with the BUP/SAM versus, for example, placebo. So there's lots of talks of minus 1, minus 2, minus 3. And I can't help but think, what does this mean practically. What does this mean for the actual person? When you say there's this change of minus 2, what does that mean for an individual's life, so to speak?

Does that mean they're sleeping 30 minutes less a day? What's the clinical significance of it?

DR. VON MOLTKE: So I would ask Dr. Papakostas to give us his perspective on how that translates.

DR. PAPAKOSTAS: Thank you for your question. George Papakostas, Massachusetts General Hospital, Harvard Medical School. I'm here as a consultant to Alkermes, but not an employee of the company.

It's a great question because studies vary
in terms of the patients that they have, and some
studies have more severe patients or less severe
patients. So how do you extract the clinical
significance of a statistically significant
finding?

The best way to do that is with a
standardized effect size, preferably by pooling the
data for both positive and negative studies, not
just cherry-picking.

So if that's applied to the current dataset
in one of the slides that was shown -- slide up,
please. This is actually essentially that. It's a
standardized effect size of all of the studies of
BUP/SAM, both positive and negative, and all of the
studies of brexpiprazole, both positive and
negative, that essentially show that the efficacy
of this program is similar to the efficacy of its
predecessor, that is, a typical antipsychotic
agent.

So it's pretty much comparable to the effect
that the atypicals have in clinical practice. So
depending on the patient that you use, you can
expect response rates of 25, 30 percent in clinical practice. Thank you.

DR. NARENDRAN: Next question, Dr. Fiedorowicz?

DR. FIEDOROWICZ: Yes. My question is, was the integrity of the blind assessed in any of the 4 studies.

DR. VON MOLTKE: Yes, it was. And to speak more specifically to that, I'll have Dr. Pathak come up. I can tell you that, most recently in the 202 study, because we had teams going back out to assess looking for documentation, et cetera, we reconfirmed the integrity of that blind even a second time.

DR. PATHAK: Could you please clarify what would you be looking for in regards to interrogating the integrity of the blind?

DR. FIEDOROWICZ: Were participants specifically asked what treatment they believed they were receiving?

DR. PATHAK: We did not ask the question. Overall, though, we want to emphasize that all the
evidence suggests that the blind was maintained.

DR. NARENDRAN: Next question, Dr. Meisel?

DR. MEISEL: Thank you. Steve Meisel with Fairview in Minneapolis; a couple of questions. I think they'll all be brief. First of all, just to clarify, is the postulate that the kappa and antagonist quality of buprenorphine is responsible for the efficacy? Because otherwise, you're combining matter and antimatter with a mu agonist and antagonist. Is that the postulate?

DR. VON MOLTKE: To some degree, that is part of it. We certainly have the kappa antagonism still going on. The amount of mu that's been eradicated, so to speak, has been empirically driven down such that we don't see on vast scales, et cetera. So whether there is some residual amount or nothing, we really can't say. But clearly, the kappa antagonism remains.

DR. MEISEL: If I were a person who was intent on diverting and abusing this product, have you tested whether or not it is possible to chemically or physically separate the two
chemicals? If I were to get a pile of these pills, and there's no doubt that there will be some internet-based recipes to separate them out, have you tested whether that's possible or not?

DR. VON MOLTKE: Yes. Our chemists did conduct a limited set of conditions, but those conditions were specifically picked because they felt that they were optimally able to likely separate those two based on physiochemical properties.

So they picked the conditions where they thought it would be easiest, and they did find a couple of conditions where you could separate buprenorphine, but the yield was extremely low and it was variable.

So starting from a 2-milligram pill or even multiple pills, the highest yields were 23 percent, and that was on a good day. It was very variable.

DR. MEISEL: Twenty-three percent sounds high to me, but okay. And then the last clarifying question; can you explain or postulate why the 8/8 dose is less effective than the 2/2 dose?
DR. VON MOLTKE: Sure. I think there are two reasons in this case. First, with buprenorphine, there's pretty good literature, both in the preclinical and then in vitro, that it has a U-shaped curve, especially on the nociceptor effects, but even in some of the other paradigms that has been looked at. So it's a complicated pharmacology just for BUP.

In addition, in psychiatry, using these symptom-based scales for efficacy, if your side effect profile starts to bleed into affecting those efficacy measurements, you're going to see a hit on your efficacy.

So for example, we have nausea, vomiting, things like that. That's going to impact the HAM-D assessment of things like appetite. So that's been well described, going back to the antipsychotics. It's a known phenomenon.

DR. NARENDRAK: Next question, Dr. Iyengar?

DR. IYENGAR: I have two questions about the sensitivity analysis that you've done. One is, did you check the assumptions about the lack of
correlation between the two stages in the SPCD design? And the second question relates to the missingness, the missing data issue.

No matter what you do, there's some assumption about the missing data. It sounds to me like MAR is the assumption that has been used here. Is that correct? And also, did you do any sensitivity analyses to check to see if some non-ignorable mechanism might be operating?

DR. VON MOLTFKE: I'm going to ask Dr. Schindler to come up and walk you through those questions.

DR. SCHINDLER: Yes, we did assume that when we do the MMRM model, that we have missing and random MAR. And as you know, when you have missing at random, the estimates from the MMRM model are unbiased. We do have some sensitivity analyses.

The other question was about the correlation of the estimates between the two groups. And even though you have patients who were on placebo in stage 1 and they rolled into stage 2, but the advantage of the SPCD design is the differences...
that you calculate are actually not correlated between stage 1 and stage 2.

So at the difference level, the correlation goes down to zero, and that's been demonstrated. There was actually a paper by Yeh Fong Chen from FDA, et al., who have shown that there isn't a correlation between the difference.

I can show a little bit of some sensitivity analyses for the missing data. Slide up.

Here's the study 202, and what you see here is the HAM-D data and the MADRS-10 data at end of treatment. The primary analysis for HAM-D is what we presented earlier. MADRS-10 was also presented. And then when you use the MMRM model, that's the top line where we see the primary endpoint -- there's other ways to deal with missing data, and those other ways include multiple imputation.

So when we do multiple imputation using the other members of that same treatment group to fill in the data for the people who are missing, that's the second line, and you see for the HAM-D 17 at
end of treatment, that's virtually the same;
MADRS-10, end of treatment, also very similar for
multiple imputation.

But there's an even more conservative
approach to multiple imputation, as you know, and
that's where you replace the missingness with the
placebo. So you flip the missingness and use the
placebo means to replace the data in the multiple
imputation, which is much more conservative, which
would mean that the members who are missing are
having the same effect as placebo.

We know that's much more conservative. The
HAM-D 17 at end of treatment, the mean value is
still on the side that favors BUP/SAM. The
confidence interval spreads out a little bit across
the zero, but we know that's a very conservative
multiplicity, very conservative multiple imputation
method.

For MADRS-10 end of treatment, we see the
same result as before, so it doesn't cross zero.
So this is reassuring that even with our
sensitivity analyses, we still see a robust effect.
DR. IYENGAR: Just one quick follow-up; roughly how much missingness was there?

DR. SCHINDLER: Well, these are longitudinal studies, so the missingness -- for people who are lost to follow-up, some may be lost over time. So the most extreme would be towards the very end of the study, the very last visit. So the MMRM model uses all the data that's there at the earlier visits. So the missingness is not the same at each visit. The most extreme is about 20 percent at the last visit.

DR. NARENDRAN: Next question, Dr. Jain?

DR. JAIN: Thank you. I have two questions. One pertains to efficacy and the other to side effects. In terms of efficacy, at the individual subject level, have you calculated response rates and remission rates from the placebo-controlled trials?

On the safety side, you presented some data related to the COWS in terms of withdrawal scores in which participants were categorized into mild, moderate, and severe categories.
Do you have that data available in a summary fashion, a continuous fashion, such that we could understand what symptoms participants were experiencing and at what level subjects within the mild category were experiencing withdrawal?

DR. VON MOLTKE: Yes. I will have Dr. Pathak take your first question about response and remission, and then we'll have one of our safety people walk you through the COWS data that we have.

DR. PATHAK: Sanjeev Pathak, psychiatrist, Alkermes. Yes, we did explore or evaluate a response in remission in the placebo-controlled trials. I do want to emphasize that the duration of the trials was 4, 5, or 6 weeks.

What we found in the overall patient population; that the median time to remission was approximately 3 months or so, and this we can glean from following the patients as they rolled over to the longer-term study. And overall, there was clinically meaningful benefit.

DR. CONLEY: Next question?
DR. STANFORD: Hi. Dr. Arielle Stanford. I'm a psychiatrist. I work at Alkermes. We looked at the item levels for the COWS scores, and there was no particular pattern of which items were particularly common across the different subjects. The items were generally G.I. in nature, which is not uncommon for what we see following discontinuation. But that said, there was no pattern, and the individual item scores were low for any particular item.

DR. NARENDRAN: Next question, Dr. Dunn?

DR. DUNN: Hi. Walter Dunn. One follow-up and two quick questions. Follow-up; in terms of the averaging of the MADRS, what was the rationale for averaging weeks 3 through 5 as opposed to 4 and 5 or 2 through 5?

Second follow-up or second question; do you have the individual data for that patient who the FDA is proposing to exclude from study 202, specifically their week-by-week response? And the third question is, do we have any evidence that, for patients who are on naltrexone, would BUP/SAM
still be effective for antidepressant?

DR. VON MOLTKE: We don't have data on the naltrexone, but we would assume that being on a complete blocker to start with might be something that would not be something we would recommend.

I'm going to then march up to your second question on the individual data for the patient that was proposed to be excluded, and then I'll ask Dr. Schindler to come up and address your first question.

So we do have the single assessments for the single patient. Slide up, please. So you can see here, we've got both the HAM-D, which was the primary endpoint as well as the MADRS-10. The patient came in and had a pretty robust response following both of those indices.

The patient then started a taper, which was part of this particular study at week 4. You can start to see the loss of response, and then continued on placebo following that. So it looks pretty consistent to us. And in addition to that, the patient did meet all eligible criteria.
There had been some question as to how long the patient had been on fluoxetine to give you an idea of what the question was. And we found actually 5 separate independent corroborators that the patient had been on fluoxetine for more than 8 weeks, which was the defining line.

Slide down, please. Dr. Schindler?

DR. SCHINDLER: The earlier question about averaging, why did we pick the weeks that we picked; from the protocol, we used titration at the beginning of the study, so we did want to average during the time period of titration. But we did look at different ways of averaging over different weeks.

Slide up. This shows that what we wanted to do was to really see what was the treatment effect during the duration that there really was a treatment effect to measure, but we looked at different time points.

So this is the study 207, which I presented, and this is MADRS-10 data. We looked at week 2 through end of treatment, week 3 through end of
treatment, which was the primary endpoint. We looked at 4 through end of treatment. They give essentially the same picture. I think you could pick any of these if you wanted to.

We chose to look at the time point that included the largest duration after titration.

DR. NARENDRAN: Ms. Witczak?

MS. WITCZAK: Kim Witczak, consumer representative. I'm curious a little bit about the ages of the people that were in the clinical trial and were any of them seniors, just because that's going to have real-world consequences.

DR. VON MOLTKE: Sure. The upper end of the age range, I believe, was 70.

MS. WITCZAK: How much weighted, I mean --

DR. VON MOLTKE: Yes. There were not many that were over age 65, for sure, and we would be advising caution if a patient had underlying concomitant serious illnesses, which is what was done in the trial as well.

MS. WITCZAK: With multiple medications?

DR. VON MOLTKE: Yes, and things like heart
disease, lung disease. There'd have to be some caution and some extra consultation with their physicians.

MS. WITCZAK: Great. Thank you.

DR. NARENDRAN: Dr. Hernandez-Diaz?

DR. HERNANDEZ-DIAZ: Hi. Sorry for going back to the question of what is clinically meaningful, probably for Dr. Papakostas. The question is not about the statistically significant or standard deviations, but clinically, what would you consider getting better?

Or in the plot, you put a remission. In the scales that you use clinically and the ones that have been used for research, is decreasing 15 points in the scale, 10 points in the scale, is that the kind of question? I would like to understand.

DR. PAPAKOSTAS: So historically, antidepressants, the traditional antidepressants, if you use them as monotherapy in a regular population, not selected for treatment failures, if you use the MADRS, you tend to see differences from
placebo, 3 to 4 points.

In augmentation studies, depending on how many failures there have been, how difficult to treat the population is, typically you see between 1 and a half and 2 and a half MADRS points of a difference. This seems to fall between those two.

DR. HERNANDEZ-DIAZ: So Dr. Mathis presented a case for patient J.D., presenting it as a patient that had been on treatment and went from a scale of 21 to 16, so 5 points different, and that patient was considered to be not responding, so a treatment failure. So 5 points difference in clinical practice is considered failure, but a difference of 1 in a trial is considered effectiveness?

DR. PAPAKOSTAS: I got it. I wasn't clear enough in your question before. Let me try again. So when I talk about 3 or 4 for monotherapy and 1 and a half to 2 and a half augmentation, I don't mean improvement. I mean improvement in drug, minus improvement for placebo.

So if you treat a patient, and they're very severe, and their improvement is 5 points, that's
not enough. But if you have a decrease in
depressive symptoms that's greater than placebo in
augmentation studies of 1 and a half to 2 and a
half points, that is a strong signal.

DR. HERNANDEZ-DIAZ: Yes. But since in
clinical practice, we are not going to leave the
placebo, then we are going to have only the one
point unless we are going to leave placebo, and
placebo meaning in this context we're only taking
care of a patient being closer or whatever we are
doing in the trial. We are not going to do that
when we prescribe. Right? So we are going to have
only the one point.

DR. PAPAKOSTAS: Let me try a different way.
Is there a slide that shows how much of a change in
MADRS there was on the drug during a study?

DR. HERNANDEZ-DIAZ: Slide 46, I think.

(Laughter.)

DR. PAPAKOSTAS: Thank you. Slide up.
Excellent. Slide up, please.

So a change from baseline here, the average
here in stage 1 is 7 points and the average in
stage 2 is 3 points with a very low placebo response. But that's the average. That's looking at patients where this didn't work in patients and where this did work. That's meaningful, and the difference is also meaningful. Thank you.

DR. NARENDRA: I think we're clear on that.

DR. MATHEW: I'm sorry. Just a quick clarification. In the case, it was a QIDS scale, so that's a self-report instrument, so we weren't addressing the MADRS.

DR. NARENDRA: Thank you, Sanjay.

Dr. Kulldorff?

DR. KULLDORFF: Thank you. My name is Martin Kulldorff. I'm a biostatistician at Harvard Medical School. I have questions for Dr. Papakostas and for Dr. Schindler. Maybe we can start with Dr. Papakostas.

On slide 16, if you can get it on the screen also, I understand the first step there after the first treatment, 63 percent are still symptomatic and 37 percent we had success. The subsequent one, is that 75 percent of those 63 that are still
symptomatic?

    DR. PAPAKOSTAS: Correct. So if the first doesn't work, the chances of the second working in the patient where the first did not work goes down to 25 percent. If you give another one in those patients where it did not work, the success goes down even further.

    DR. KULLDORFF: So using the Excel spreadsheet, I did a quick calculation to see what is then the overall success rate after the four steps. And my number was that the success rate is at 66.1 percent and there are 33.9 that are still symptomatic.

    Would you agree that that would be a reasonable number based on your experience?

    DR. PAPAKOSTAS: I didn't do the calculations, but I think a better way of orienting this -- because I think I see what you mean. A better way of orienting this is what do we do in the clinic if you have a patient where the first or the second treatment didn't work?

    So let's say that you start from the end of
step 2. If you give them a monotherapy, the chances that the person will remain symptomatic after you optimize the dose duration is going to be close to 90 percent if you use a monotherapy. So you need to use an adjunct, and that's where we need to grow. Thank you.

DR. KULLDORFF: Thank you.

DR. NARENDRAN: Next question, Dr. Crawford?

DR. KULLDORFF: I am confused if we can compare slides 43 and 54. The 205 study, if we look at slide 54, if we can get that on the screen -- thank you. The 205 study there has an effect size of about minus 2 and a half and is statistically significant. But if you do it on 43, there's an effect size of about minus 1.8, and it's not statistically significant.

Why is there a difference there? I'm missing something.

DR. SCHINDLER: It's the time point that's different. There's two slides. On this slide, we're looking at MADRS-10. Can you do the last slide? This slide, just to remember, is for 205,
looking at MADRS-10 at the end of treatment. So it's hard to keep track of the excellent time points that we're making the comparison. So this is a comparison at the end of treatment.

Now, the other slide, 43, if you bring that slide up, that's at week 5. And you notice the dilemma for this -- and this is exactly the reason why we want to use the average difference across multiple time points. At week 5, see at the top of the slide, it says MADRS-10 evaluated at week 5.

At week 5, there's variability in the data, and the curves -- especially you notice it in stage 2. The curves start to come together a little bit, and that difference is a little bit smaller than it is at different time points. At week 6, you can see the week right next to it, the curves separate even more.

So then the question would be, when you look at end of treatment, the end of treatment is looking at week 6, but this evaluation was looking at week 5.

So if your goal is to understand what is the
treatment effect, which week you picked, if you only pick one time point, which week you pick is important. If you average some of the differences together, then it becomes less important, and you're really measuring the treatment effect. You're not measuring the variability of the data.

I think that's what you've uncovered, that if you look at just week 5 versus the end of treatment, you see a slightly different picture. And that's just due to the variability of the data.

DR. KULLDORFF: So on slide 54, you looked at week 6 only then?

DR. SCHINDLER: Week 6, and you notice -- if you look at this picture, you see stage 1 is 5 weeks and stage 2 is 6 weeks. So the end of treatment for stage 1 is that 5-week period, but the end of treatment for stage 2 is 6 weeks because they're one week different.

DR. KULLDORFF: So slide 54 is a combination of weeks 5 and 6.

DR. SCHINDLER: It's a combination. Right. It's the end of each stage. It's the last week in
each stage.

        DR. KULLDORFF: That leads to my second
question, which --

        DR. NARENDRAN: Can we get one second? I
just want to -- we have two more questions, so
we're running out of time. I think we're clear on
that point.

        Dr. Crawford?

        DR. CRAWFORD: Thank you. The applicant has
provided extensive information and data on
buprenorphine and samidorphan as fixed-dose
combinations. The true new molecular entity is
samidorphan, and prior to the agency's
presentation, because I do see something there, I
wanted to ask the applicant, do you have any
comment of your findings you might have of the
efficacy and adverse effects of the NMEs,
samidorphan alone?

        DR. VON MOLTKE: Yes, we've had an extensive
characterization of samidorphan. It actually has a
place at another development program for us as
well. It has some similarities in its AE profile,
things like nausea, G.I. upset, some constipation.

Its efficacy wasn't tested partly because its purpose here was primarily to mitigate the mu agonist activity, but it also is part of a group of antagonists that we know from other studies have not shown to have antidepressant activities.

So we really had no reason to think that samidorphan on its own would have efficacy.

DR. NARENDRAN: Next question, Dr. Riley?

DR. RILEY: So as I understand it, what you are doing is a mean or means on the MADRS average. Did you also consider or do analyses on the patient level instead of the group level? And if so, did you get comparable results?

Related to that, I'm more familiar with mixed models for repeated measures over time and treatment effects. Did you consider those or do those as well?

DR. VON MOLTKE: I'll ask Dr. Schindler to come up.

DR. SCHINDLER: Just to answer the endpoint, the last point, I think the MMRM model is a mixed
model, so I think that, yes, we use the MMRM model. And recognize that what the MMRM model does is it looks at all of the visits, all of the data from all of the visits, and calculates a visit mean for each treatment group as part of the model. And that's part of what you estimate in the model, and then you look at the contrast between the visit mean for one group and the visit mean for another group. And that's part of the way we analyze the data.

What we don't do is analyze individual patient data over multiple time points, and recognize that there's a certain amount of missing data, not a lot, but a certain amount of missing data.

So if we average 2, 3, 4, 5, and 6 weeks over individual patient data, some of those might be missing, and so we didn't do that. That analysis, I don't think, would be the fair analysis for this type of data. The MMRRM model uses the available data to estimate that, what the treatment effect is for the study at each time point, and
then uses sort of a variance/covariance structure
to really estimate the trajectory of the change
from baseline, and that trajectory is what you see
as part of the model.

So that's the way we did it. We did not
average individual patient data over time. We just
averaged the treatment effects estimated from the
model at multiple over time at multiple time
points.

DR. NARENDRAN: I have two focused
questions. Do you have in vivo receptor occupancy
data for the kappa and the mu opioid receptor,
either in human or primates, for the 2-by-2 dose?

DR. VON MOLTKE: We do have in vitro data.

DR. NARENDRAN: In vivo, in vivo, in PET
imaging.

DR. VON MOLTKE: No, no. We do not. I'm
sorry.

DR. NARENDRAN: The second question is, the
human abuse potential study, you looked at in
controls and also you showed in major depressive
disorder. The drug liking scores, were they done
at week 1 or were they done at -- were they done
after 7 days when the metabolite accumulates?
Because it said samidorphan is a metabolite that is
a full mu agonist. What time points were they?

DR. VON MOLTFKE: I'll have Dr. Pathak walk
you through that.

DR. PATHAK: Dr. Narendran, to clarify, your
question is about the human abuse potential study.
Right? These are single administration, and the
evaluation of drug liking is done at multiple time
points for a prolonged period of time. So in this
study, it was 48 hours and it was done repeatedly.

DR. NARENDRAN: So just to clarify, you did
not do multiple dosing and you did not get it at
7 days, where the mu agonist would have accumulated
the metabolite?

DR. VON MOLTFKE: Right. The half studies
were single dose.

DR. NARENDRAN: Thank you.

I think we'll take a 15-minute break. We're
running a little bit over time. We'll try to make
up that 15 minutes during the lunch break if that's
okay.

Just to remind panel members, please remember there can be no discussion in the break amongst yourselves. So we'll meet back at 10:45. Thank you.

(Whereupon, at 10:32 a.m., a recess was taken.)

DR. NARENDRAN: Thank you. We will now proceed with the FDA presentations, starting with Dr. Tiffany Farchione, deputy director's presentations.

FDA Presentation - Tiffany Farchione

DR. FARCHIONE: Hi. Good morning, everybody. I am going to begin the FDA presentations this morning by providing the regulatory history of this development program and the interactions that the company had with the agency.

I plan to highlight a few key interactions between the agency and the applicant. So beginning with the pre-IND meeting in February of 2011; moving on to the end-of-phase-2 meeting request
that was scheduled for October 2013; the written
guidance that we provided on the statistical
analysis plan in July of 2015; a guidance meeting
related to the completed studies 205 and 206 in
September of 2016; and guidance on the revised
statistical analysis plan and endpoints in February
of 2017; the pre-NDA meeting that included feedback
on the MADRS-6 in July of 2017; and culminating
with the NDA submission in January of this year.

I'm also going to highlight a few other
interactions in the year or so preceding NDA
submission, specifically a receipt of the amended
protocol and statistical analysis plan for
study 207 about a week before that September
meeting in 2016; the breakthrough therapy
designation advice that we provided in March of
2017; and the applicant submission of their dossier
for the MADRS-6 that they submitted in April of
2017.

The first interaction that we had with the
applicant related to this program was the pre-IND
meeting in February of 2011. And at that time,
they described their plan to develop buprenorphine and samidorphan for the adjunctive treatment of MDD.

We discussed the general design elements necessary to support that indication, like selecting patients with inadequate response to antidepressants and ensuring that the patients continued their antidepressant on which they had suboptimal response while enrolled in the double-blind treatment phase. So in other words, the studies were to involve an add-on design, which they did.

We also discussed the support needed for a combination drug and the difficulties involved with conducting a full factorial study for this particular product. So normally, if you have a combination drug, we expect that the company would provide evidence of efficacy for both components separately as well as together, but in this case, the samidorphan was added for safety and not for efficacy.

We also noted the potential to produce
opioid toxicity in opioid-naïve patients with MDD or potentially opioid dependence if patients were treated with buprenorphine alone. The applicant did follow our advice on study design in this regard and did not conduct the full factorial study for the safety reasons noted.

The phase 2 study that's been referred to today as study 202, was discussed during this meeting as well. At that time, the applicant specifically described it as a proof-of-concept study. There was no a priori hypothesis as to which dose would be more likely to be effective, and there was no plan for multiplicity adjustment to control for type 1 error or the chance of a false-positive finding.

So because of that, we said that we had no objection to using sequential parallel comparison design in the proof-of-concept trial, and we encouraged the applicant to provide a detailed statistical analysis plan and seek feedback prior to starting the study if they intended to use this study to support an efficacy claim.
We received the statistical analysis plan for this study in August of 2013 along with the study report.

Following the completion of study 202, the applicant requested an end-of-phase-2 meeting that was scheduled for October of 2013. In their background package for that meeting, they described their planned phase 3 studies, and, as is typical, submitted questions in advance related to that plan.

The agency provided preliminary comments a few days ahead of the scheduled meetings, and in one of those comments, we did express our concern related to the SPCD analyses. We noted that, from a statistical perspective, even though the design appeared to be reasonable, there hadn't been any analytical proof for the validity of the associated analyses, particularly when they were missing data.

Following the receipt of our comments, the applicant cancelled the face-to-face meeting, noting that our preliminary responses were sufficient and they did not require further
discussion.

Our next notable interaction was in July of 2015. The applicant requested a guidance meeting to discuss the statistical analysis plan for the phase 3 studies, and we provided written responses to that request. And at that time, we again noted our concerns related to using SPCD in the pivotal trials.

Again, we reiterated that we hadn't endorsed any analytical method for SPCD in a confirmatory setting. We were actively engaged with the academic community to try to understand the pros and cons of that design from a regulatory perspective.

So we actually encouraged the applicants to collect efficacy data from both stages because, then, they would have sort of a back-up plan. Right? So if the primary endpoint at the end of the study using SPCD was not positive, we said that we could use the data from stage 1, which was a more traditional design if we needed to, if the analyses were still unsettled at the time of the
The applicant continued their phase 3 program, and when 2 of the 3 planned short-term trials were completed, they requested another meeting to discuss the results and to plan their next steps.

During the face-to-face meeting, this was in September of 2016, we discussed studies 205 and 206 at that time. So again, 205 was another SPCD study and 206 was the one with the placebo lead-in that you heard about earlier.

The applicant acknowledged that neither study met its prespecified primary endpoint and they presented some additional analyses that they said would inform modifications to the other ongoing short-term phase 3 study, which was study 207.

Actually, a week prior to the face-to-face meeting, the applicant did submit their modified protocol and statistical analysis plan, but that wasn't enough time for us to review it prior to the meeting. So they ended up requesting a follow-up
meeting and that meeting was held in February of 2017.

So at that meeting, we specifically focused on the changes in the primary endpoint and the statistical analysis plan. So at this time, they changed the primary endpoint from the end of treatment MADRS-10 to the three endpoints that you heard about earlier, to be evaluated in a hierarchical fashion.

So the MADRS-6 average, using the change from baseline to week 3 through the end of the efficacy period; the MADRS-10 average change from baseline, week 3 to end of efficacy; and then the traditional change in MADRS-10 from baseline to end of treatment, and again noting that the end of treatment was week 5 for stage 1 and week 6 for stage 2.

So in response, we noted that we have not previously accepted the MADRS-6. This was a novel endpoint to us. So we recommended that the applicant submit a dossier for us to evaluate so that we could actually look at the MADRS-6 and
determine whether it was fit for purpose for an antidepressant trial.

We did not agree to the averaging strategy. We actually didn't say that it was a review issue. We said that, "We do not agree with your averaging strategy." We also noted that the protocol amendment introduced additional complexity into the SPCD analyses because the stage 1 and stage 2 now have different durations, so you had 5 weeks versus 6 weeks.

Then at that time, the applicant also proposed pooling studies 205 and 207, but the suggestion was made with the data in hand and not prospectively. And we stated that the pooled analyses could only be considered exploratory.

So despite this feedback, there was actually little that the applicant could do to address our concerns because the database lock for study 207 occurred in October of 2016.

The applicant also requested informal breakthrough advice. This is a relatively new option, which allows companies to ask the Division
leadership whether a product might qualify for a breakthrough therapy designation or what additional information they might need to support a breakthrough request.

Now, they did mention earlier that they had a fast-track designation, so that's one of the other programs that we have for promising applications, but the bar is higher for the breakthrough therapy designation. So we actually advised the applicant that it would be difficult to grant a breakthrough designation for this program because, at that time, we still had not yet determined the acceptability of the MADRS-6, and we noted that any statistical significance in phase 3 relied on post hoc analyses.

The applicant did then submit the dossier for the MADRS-6 in April of 2017, and we consulted our clinical outcome assessment staff at that time. At the pre-NDA meeting that was held in July of 2017, we did discuss the COA review of the MADRS-6. We noted that the MADRS-6 was not fit for purpose. We said that it could not replace the MADRS-10 for
use as a primary endpoint because it excludes concepts that are relative and important in major depressive disorder.

So the missing items in the MADRS-6 are related to reduced sleep, reduced appetite, concentration difficulties, and suicidal thoughts. These are core features of depression and can't be excluded from an endpoint in a trial that is designed to assess antidepressant efficacy. We informed the applicant that any analyses of MADRS-6 would be considered exploratory; again, not a review issue. We said they would be considered exploratory.

That brings us to the January 31, 2018 NDA submission, and as has been widely reported, we did initially refuse to file the application. The sponsor noted and we relayed in our background package as well that there were some factual inaccuracies in the letter in which we refused to file, so at that point, we did have a conversation with the applicant. They clarified the analyses that they used that were intended to support the
application, and at that point, we agreed to file
the application, which brings us here today.

At this point, I'm going to hand things over
to Semhar so that he can talk to you about our
evaluation of the efficacy analyses. Thank you.

**FDA Presentation - Semhar Ogbagaber**

DR. OGBAGABER: Good morning. I'm Semhar
Ogbagaber from the Division of Biometrics I. I
will be presenting the FDA review of efficacy data
for buprenorphine samidorphan, which we call
BUP/SAM throughout the presentation.

Four studies were submitted in this NDA,
which investigated efficacy of BUP/SAM for
adjunctive treatment of major depressive disorder.
Both applicant and FDA agreed that studies 205 and
206 were negative. So in this talk, I will focus
on the other two studies, 202 and 207, that had
some concerns. Study 207 would be considered
positive if MADRS-10 average were accepted as
primary endpoint.

Since both studies 202 and 207 utilized SPCD
or the Sequential Parallel Comparison Design, I
will first briefly introduce SPCD before going through the efficacy results, then present general questions or concerns related to SPCD, and in the end, touch on meta-analysis.

The SPCD was initially proposed to address high placebo response observed in major depressive disorder trials. Unlike the conventional parallel comparison design, which has only 1 treatment period, SPCD evidence is based on efficacy data from 2 stages. More patients are allocated to placebo group in stage 1, so more placebo non-responders will be randomized in stage 2.

The SPCD estimated treatment effect is essentially a weighted average of estimated treatment effects from both stage 1 and stage 2. Within each stage, you compare drug versus placebo as in any conventional design. Stage 1 consists of all incoming patients and stage 2, placebo non-responders from stage 1.

Here is a hypothetical illustration of how SPCD treatment effect is computed. As shown in the previous slide, it's a weighted average of
treatment effects from both stage 1 and stage 2. As illustrated in the table, supposed weight allocation of 60 percent in stage 1 and 40 percent in stage 2; suppose in the first stage we have 150 patients, which many of them are randomized into placebo arm, the estimated treatment effect for stage 1 would be 2 units.

At the end of stage 1, 70 patients were identified as placebo non-responders. In stage 2, the treatment effect is 3 units based on the 70 placebo non-responders, and the SPCD estimated treatment effect is computed to be 2.4 as indicated in the numbers in red.

This is the estimated treatment effect when the two stages are combined. First, we calculate the SPCD estimated treatment effect. To derive the p-value, we need to calculate the variance associated through this estimate.

From this illustration, it is apparent that different weight allocation affects SPCD-estimated treatment effect. If we expect a larger treatment effect in stage 2, it would probably make sense to
assign more weight to stage 2. However, from a
statistical perspective, you might want to do it
the other way around because the sample size in
stage 2 is relatively smaller in stage 2 than
stage 1.

But clinical relevance should play a key
role in weight allocation and also which population
is relevant may deserve some attention.
Regardless, it's uncertain whether it is reasonable
to combine estimates from both stages because
stage 1 enrolls the usual patient population and
stage 2 an enriched population.

In addition, from this illustration on this
slide, there are 70 placebo non-responders entering
stage 2. These 70 patients also contributed data
in estimating treatment effect in stage 1. So
these patients appear to have more influence in the
overall estimation of the SPCD treatment effect.

Now, let's look at study 202. Study 202 was
a phase 2 proof-of-concept, multicenter, randomized
trial using SPCD that investigated BUP/SAM 2/2 and
8/8. The primary endpoint was changed from
baseline to week 4 in HAM-D total score.

In the applicant's primary analysis, there was no multiplicity adjustment prespecified to control the overall type 1 error rate. This was an exploratory study to generate hypotheses for study 207. It's uncommon to draw a final inference based on exploratory objectives.

Results for study 202 are summarized in this table. For the low dose, 2/2 compared to placebo estimated treatment effect was minus 2.8, and for the high dose, it was essentially neutral.

Although the nominal p-value for the 2/2 dose was 0.014, we do not consider the efficacy was demonstrated because there were 2 doses compared against placebo and multiplicity adjustment was not prespecified.

Also, in a post hoc exploration, conclusion becomes inconsistent when using Bonferroni or fixed sequence testing procedure. If we use Bonferroni procedure by splitting alpha, the 2/2 dose would have made it. But if we use fixed sequence testing procedures, starting with the high dose, which is
the prespecified procedure in the other studies, no
dose would have made it.

Typically, we would expect a larger
treatment effect on the higher dose, but in this
trial, estimated treatment effect was essentially
neutral and p-value was really high. If the 2/2
dose and 8/8 treatment groups were combined and
compared against placebo, result was not
statistically significant. In FDA's view, without
the prespecifications, study 202 is not an adequate
and well-controlled trial and statistical
interpretability is undermined.

In this exploratory study, site 124 was
identified as influential by the Office of
Scientific Investigations at FDA. This site
enrolled a single patient, randomized in the 2/2
dose group in stage 1 and had an extreme result.

This quote here has been amended, and I want
to emphasize that there is a slightly different
quote here. The quote on this slide, according to
the Office of Scientific Investigations, should be
the subject's eligibility for the study cannot be
determined due to incomplete and contradictory source information. Therefore, we recommend DPP to conduct sensitivity analysis for patient 124001 in study 202.

To explore the impact of this patient, we compared the response profile of this patient against the treatment groups. This plot displays the change from baseline in HAM-D 17 score at each visit. The red curve is the single patient who was randomized to the 2/2 dose group in stage 1. The three curves are the group means for the three treatment arms.

This patient was randomized to the 2/2 dose group. The mean response profile for this dose group was represented by the solid gray curve. This clearly shows that the patient is quite far from the mean plot of the assigned treatment arm.

We also explored the potential impact of this site on efficacy results by removing the site. By excluding this site, the magnitude of estimated treatment effect dropped from 2.8 to 2.2 and nominal p-value goes up from 0.014 to 0.057. The
strength of evidence for the 2/2 dose diminished.

In FDA's conclusion, study 202 was not an adequate and well-controlled study. There was no prospective plan for type 1 error control rate. The nominally positive result for the 2/2 dose depended on a single patient. This further undermined the strength of evidence.

If study 202 had been adequate and well controlled and if statistical significance of results for the 2/2 dose had not been sensitive to the contribution of a single patient, the neutral results with the 8/8 dose would undercut the persuasiveness of the results for the 2/2 dose.

This is study 207 that was a phase 3, multicenter, randomized, and investigated BUP/SAM dose 1/1 and 2/2. It also used SPCD design that had two different periods in the first stage and second stage. The first stage had 5 weeks and the second stage had 6 weeks.

The proposed primary endpoints were a change in MADRS-6 total score using average of change from baseline to week 3 through end of efficacy period,
week 5 for stage 1 and week 6 for stage 2, change in MADRS-10 total score using average of change from baseline to week 3 through end of efficacy period, change in MADRS-10 total score from baseline to end of treatment period.

So initially, there was only one primary endpoint specified, which was the third one. Toward the end of the study, the applicant added the two top endpoints and proposed to test them sequentially in the order it appears, and both added primary endpoints measured average improvement over several weeks, and the top one is based on MADRS-6.

The clinical outcome assessment staff at the FDA had a concern with the MADRS-6 total score as a primary endpoint because of the omission of clinically important symptoms included in MADRS-10 and those symptoms were reduced sleep, reduced appetite, concentration difficulties, and suicidal thoughts.

This is MADRS-10 total score at each week or visit, and visit zero corresponds to the baseline
visit for stage 1 and stage 2. There's a general improvement in depression in both stages, but the improvement seems to level off or worsen after week 4 in stage 1 or week 5 in stage 2.

In addition, comparing the two stages, the baseline scores were generally lower in stage 2, with an average of 32 in stage 1 and 27 in stage 2. The observed improvement in stage 2 seems very tiny compared to stage 1. Recall that stage 2 are all placebo non-responders from stage 1.

This is the mean change from baseline in MADRS-10 total score. At each week, we see similar patterns to the previous slide; that is, there was a general improvement in depression in both stages, but improvement seems to level off or worsen after week 4, stage 1 and week 5 in stage 2. Also, larger improvement is observed in stage 1 than in stage 2.

This is the applicant's efficacy result within each stage and combined SPCD results. The top portion is for the MADRS-10 average endpoint and the bottom portion is for MADRS-10 end-of-
treatment endpoint. Among those comparisons, there was only one comparison leading to statistical significance, as shown in the cell with red numbers. This was the 2/2 dose group on MADRS-10 average and only when the two stages were combined. The estimated treatment effect was minus 1.9 with a nominal p-value of 0.026.

In conclusion, efficacy on the 2/2 dose was supported based on MADRS-10 average and not MADRS-10 end of treatment, which was a conventional primary endpoint used in most MDD trials. For the MADRS-10 average, the magnitude of estimated treatment effect was 1.9 on a 60-point MADRS scale.

FDA had a concern with MADRS-10 average as a primary endpoint. In particular, averaging the change in MADRS-6 or MADRS-10 tends to obscure possible drop-off in drug efficacy after the first few weeks of treatment. This is a clinical call regarding acceptability and/or interpretability of MADRS-10 average, SPCD in unequal durations in both stages.

The SPCD is a novel design. There is
ongoing research. Its pros and cons have not been fully recognized. We would like to shed some insights to help further development in this area, both clinical and statistical.

One clinical question is with regards to mixed patient population associated with SPCD design. Which stage deserves more weight? Stage 1 had the usual patient population, a trial population, and stage 2 has placebo non-responders from stage 1, an enriched population selected to increase signal over noise, but less similar to the patient population for whom drug would be prescribed.

For an effective drug, one would expect to see larger treatment effect in stage 2. However, sample size is smaller. Is there a concern with drop-outs when drop-outs are substantial? Placebo non-responders who stay through end of stage 1 may be intrinsically different from placebo drop-outs.

The second clinical question is with regards to prespecified weight allocation. Is there a clinically sensible weight allocation? If so, what
is sensible? Different weight allocation leads to
different treatment effect and may affect the
strength of statistical significance. And placebo
non-responders are reused so they have more
influence.

Question 3 is with regards to unequal
durations between stages. Is it clinically
relevant to combine estimated treatment effect over
unequal durations between stage 1 and stage 2?
Stage 1 has 5 weeks and stage 2 has 6 weeks in
study 207.

Question 4 is with regards to labeling. If
a drug is to be approved based on SPCD results,
even if neither stage demonstrates efficacy, how
does one describe the estimated treatment effect?

Regarding statistical analysis, the analysis
associated with SPCD tends to be complex because
estimates are from a combination of two stages, and
some patients are used in both stages. It's not
clear whether the correlation between the two
estimates from the two stages can be reasonably
assumed to be zero in any analysis.
To our knowledge, analysis proposed in literature has been used based on zero correlation. This is also an assumption made in applicant's analysis. What are potential impacts of a wrong assumption? If the drug is not effective, the wrong assumption may lead to type 1 error inflation.

If drug is effective, it may lead to bias in estimating treatment effect; that is, the claimed 95 percent confidence interval of treatment effect may not actually provide the 95 percent coverage.

To assess the impact of possible wrong assumption about zero correlation, we performed a bootstrap sampling-based statistical inference without such rigorous assumption of normality. And our conclusion based on the empirical 95 percent confidence interval is consistent with applicant's normal based conclusion. So for this trial, this concern can be dropped. However, in another trial, concern may not be dropped.

The second concern was with regards to substantial drop-outs. Potential impact would be
on type 1 error inflation, still under research. In this application, however, the drop-out rates were not substantial.

Meta-analysis was included in the applicant's package. We should be aware that meta-analysis does not meet the usual standard for substantial evidence. Two adequate and well-controlled trials provide substantiation and protection against false-positive finding. In this application, the meta-analysis was not prospectively planned or agreed as a means to provide evidence of effectiveness.

In FDA's view, combinability of SPCD and non-SPCD studies using meta-analysis does not appear to be sound due to incompatibility of populations in stage 1 and 2. Also, the prespecified primary endpoint was not the same across all studies. Also, the rationale behind combining effects based on primary endpoints that were not specified in SAP or protocol is questionable.

In summary, for study 202, efficacy was not
demonstrated. Study is deemed to be negative study. No prospective plan for multiple testing. Site 124 further undermined strength of evidence. Neutral results for 8/8 dose would undercut any finding of efficacy for the 2/2 dose.

For study 207, efficacy on BUP/SAM 2/2 dose was supported based on MADRS-10 average and not MADRS-10 end of treatment. Acceptability or interpretability of MADRS-10 or average in the user of SPCD and unequal durations between stages is still unclear. Meta-analysis does not meet the usual standard of adequate and well-controlled trials for substantial evidence.

Thank you, and now Danny is coming.

**FDA Presentation - Daniel Lee**

DR. LEE: Good morning, everyone. My name is Daniel Lee, and I'm the clinical reviewer for this application. I'll be presenting the safety portion of the review.

I'd like to start by reiterating Dr. Mathis' earlier statement. Based upon our review of the available safety data, FDA believes that most of
the risk ascribed to other opiates are present to a lesser degree with BUP/SAM. While samidorphan appears to reduce the frequency of adverse events associated with mu opiate receptor agonism, none of the risks inherent to opiates appear to be completely mitigated by its presence.

We analyzed the submitted safety data in several different ways while attempting to determine the best way to analyze the data. As we compared the different analyses, we found that pooling of safety data across the trials appeared to produce misleading results. This was due to the influence of within- and between-trial heterogeneity, differences in randomization ratios, and Simpson's paradox.

When we compared pooled results against results obtained in the individual trials, many of the safety signals observed in the individual trials disappeared. We also noted the emergence of new safety signals, which were not observed in any of the submitted trials. For this reason, all safety data I present today will be from individual
trials.

Overall, BUP/SAM appears to be relatively safe and well tolerated outside of being an opiate. No deaths were reported, and only 11 serious adverse events were reported in the four efficacy trials. While these were divided relatively evenly between drug and placebo, it is important to remember that significantly more individuals were randomized to placebo than randomized to drug.

One of the 11 serious adverse events is notable because it demonstrates that BUP/SAM retains the ability to displace other opiate molecules from the mu opiate receptor and precipitate withdrawal. The last two points on the slide will be the focus of attention for the remainder of the presentation.

As we take a closer look at adverse events associated with opiate-induced increase in gut motility, I'd like to draw your attention to the significant imbalance in population percentages between groups. This imbalance is observed in nearly all adverse events in the highlighted
columns.

Due to the 2:2:9 randomization ratio in stage 1 of most trials, comparisons based on absolute number of cases are not terribly helpful. The same number of cases may represent significantly different population percentages, particularly in the initial stages of each trial.

I'd also like to point out the imbalance of adverse events remains large in stage 2, despite the fact that many of these individuals started out as placebo non-responders in stage 1. These findings largely align with the applicant's presentation.

Nearly everything I've stated about G.I. adverse events also holds true for nervous system adverse events in stage 1. Curiously, the placebo non-responder status does not appear to play a significant role for adverse events occurring in stage 2.

While far fewer cases occur in stage 2, the same trends observed regarding adverse event imbalance remain. These findings also largely
align with the applicant's presentation.

Drop-outs due to adverse events were relatively low in the 4 trials overall. However, most drop-outs reported occurred in the BUP/SAM 2/2 group. The majority of these drop-outs, which are summarized in this table, are attributable to adverse events suggestive of opiate activity in the gut or central nervous system. Once again, these findings largely align with the applicant's presentation.

Our final area of focus pertaining to potential retention of BUP/SAM opiate activity is withdrawal. Withdrawal is measured using the clinical opiate withdrawal scale, otherwise known as the COWS. Unfortunately, conclusions that can be drawn from the COWS data are somewhat limited because only half of the participants in the efficacy trials received a COWS, despite the requirement that it be obtained for all participants.

The COWS data collected in trials 202 and 205 were relatively complete. Missing data were
concentrated in the later trials, and the missing
data appeared to be non-random in distribution
between groups.

For trial 202, COWS data is not particularly
informative. I'm unsure what to make of the two
cases of mild opiate withdrawal noted in the
population exposed to drug in stage 1, then
transitioned to placebo in stage 2. I don't
consider them members of the drug group, nor do I
consider them members of the placebo group. The
single case of mild opiate withdrawal noted in the
2/2 group isn't particularly convincing.

Trial 205 is slightly more persuasive, but
the numbers remain quite small. At visit 13,
BUP/SAM provides 5 cases of mild opiate withdrawal
compared to placebo's 1 case of mild opiate
withdrawal. At visit 14, BUP/SAM provides 9 cases
of mild opiate withdrawal plus 1 case of moderate
opiate withdrawal compared to placebo's 4 cases of
mild opiate withdrawal.

Here is where the problem of missing data
rears its head. Notice that the majority of each
line represents missing data. At visit 12, BUP/SAM provides 2 cases of mild opiate withdrawal compared to placebo's zero cases of mild opiate withdrawal. No cases of opiate withdrawal were noted at visit 13.

Once again, take note of how much data are missing. Despite the missing data, we observe an upswing in mild opiate withdrawal cases in the 2/2 drug group. At visit 13, BUP/SAM provides 5 cases of mild opiate withdrawal compared to placebo's 2 cases of mild opiate withdrawal. No cases of opiate withdrawal were noted at visit 14.

While none of the COWS data is particularly strong on its own, the trends observed in the 50 percent of COWS data submitted suggest that mild opiate withdrawal occurs in a subset of participants exposed to BUP/SAM.

I'd now like to hand the presentation off to Dr. Edward Hawkins. He will be providing the controlled substance portion of FDA's presentation.

FDA Presentation - Edward Hawkins

DR. HAWKINS: Good morning. My name's Ed
Hawkins, and I'm a reviewer with the controlled substance staff. Today, I'll be speaking to you about the abuse potential of buprenorphine/samidorphan, also known as BUP/SAM, and of samidorphan alone.

The applicant is proposing a new drug product that combines the mu partial agonist, buprenorphine, with samidorphan, a drug with mu opioid antagonist properties, formulated as sublingual tablets, also known as BUP/SAM.

In vitro studies indicate that it is possible to separate to some degree the buprenorphine from samidorphan and in BUP/SAM. No studies were conducted to determine the safety or abuse liability of BUP/SAM after manipulation. This includes studies regarding the different methods of administration that are typically associated with abuse, for example intranasal.

Now, as I mentioned, the sponsor is required to assess the abuse potential of samidorphan, which is a new molecular entity that is chemically synthesized from thebaine-derived naltrexone.
Under the Controlled Substances Act, all derivatives of thebaine are Schedule II substances until such time as they are down-scheduled or de-controlled following an abuse potential assessment. Thus, it was necessary to conduct an abuse potential assessment for both samidorphan alone and in combination with buprenorphine.

In vitro receptor binding and functional studies show that samidorphan acts at mu opioid receptors as an antagonist. Although samidorphan is derived from an opioid, thebaine, administration of samidorphan alone does not produce analgesia in animals.

However, it can reduce the analgesic effects of the mu opioid receptor agonist, morphine, and can reverse the cardiac and respiratory depressant effects of the mu opioid agonist, fentanyl. These data demonstrate that in whole animals, samidorphan has activity as a mu opioid antagonist.

A briefing to drugs or introduction to drug discrimination studies; drug discrimination is an experimental method of determining whether a test
drug produces pharmacological effects, which elicit physical and behavioral responses in the animal that are similar to their responses to a training drug, typically a known drug of abuse. Test drugs that produce a response similar to a training drug with known abuse potential are also likely to be abused by humans.

In this study, rats trained to discriminate morphine from vehicle indicated that morphine produced full generalization of 97 percent to the morphine queue. However, samidorphan produced no generalization, 5.4 percent, to the morphine queue.

These data show that samidorphan did not produce sensations that are similar to those produced by morphine. This was expected since samidorphan is a mu opioid antagonist.

Now, a brief introduction to self-administration studies, which is a method that assesses whether a test drug produces rewarding effects that increase the likelihood of behavioral responses in order to obtain additional drug, also known as positive reinforcement.
Drugs that are self-administered by animals are likely to produce rewarding effects in humans. The ability of a test drug to produce self-administration is indicative that the drug has abuse potential.

In this study, rats learned to lever press for intravenous heroin as the training drug, after self-administration of heroin was stable, and also were allowed IV access to the following substances, which produced varying degrees of self-administration measured as infusions per session.

Samidorphan produced 9.2 infusions. Heroin produced double that, at 18.8 infusions, naltrexone produced the same as samidorphan, 8.1, and placebo produced less than 5 infusions.

These data show that samidorphan does not produce rewarding properties that sustain positive reinforcement, similar to naltrexone.

Now, I'll speak about the human abuse potential studies, also known as HAP studies, which are used to evaluate the ability of a test drug to produce positive subjective responses in subjects.
compared to a known drug of abuse or a positive control and to placebo.

Subjects in HAP studies are individuals with a history of recreational drug use, but they are not drug dependent. When the test drug produces consistently large responses on positive subjective scales that are far outside the acceptable placebo range, it is likely that the test drug has abuse potential.

The first of two HAP studies evaluated the oral abuse potential of samidorphan alone at 2.5, 10, and 20 milligrams; oxycodone at 15 and 30 milligrams; and placebo. This study used a randomized double-blind placebo-controlled crossover design in healthy, non-dependent, recreational opioid users.

The primary measure of drug liking using the Visual Analog Scale as a bipolar scale from zero, which is extreme disliking, to 100 of extreme liking, with 50 as neutral, indicated that the positive control drug, oxycodone, at both doses produced statistically significantly higher mean
drug liking scores compared to placebo, which
establishes assay sensitivity and validates the
study.

Samidorphan at all 3 doses; 2.5, 10, and 20
milligrams, produced mean drug liking scores of 57,
58, and 60, respectively, that are within the
placebo range from 40 to 60. The secondary
measures of overall drug liking, high, good drug
effects, and take drug again for oxycodone produced
mean drug scores on each of these positive
subjective measures that were statistically
significantly greater than placebo.

Samidorphan at all 3 doses; 2.5, 10, and
20 milligrams, produced mean scores on each of
these positive subjective measures that were within
the placebo range.

The second of two HAP studies evaluated the
abuse potential of samidorphan alone, again at
higher doses, 10 and 30 milligrams, oxycodone,
pentazocine, naltrexone, and placebo. This study
used a randomized, double-blind, double-dummy
placebo-controlled crossover design in healthy,
non-dependent, recreational opioid users.

Again, the primary measure of drug liking on a bipolar VAS scale from 0 to 100 with 50 as neutral indicated that the positive control drugs, oxycodone and pentazocine, produced statistically significantly higher mean drug scores compared to placebo, which establishes assay sensitivity and validates the study.

Samidorphan at both doses, 10 and 30 milligrams, produced mean drug liking scores of 59 and 61, respectively, that were statistically within the placebo range of 40 to 60. This was also observed for naltrexone, which produced a score of 58.

The secondary measures, measured by using a VAS of overall drug liking, high, good drug effects, and take drug again, indicate that oxycodone and pentazocine produce mean scores on each of these positive subjective measures that were statistically significantly greater than placebo. Samidorphan at both doses produced mean scores on each of these positive subjective
measures that were within the placebo range. This was also observed for naltrexone.

A third HAP study evaluated the abuse potential of sublingual tablets of buprenorphine and samidorphan at 2 and 2 milligrams, 8 and 8 milligrams, and 16 and 16 milligrams, buprenorphine alone at 8 and 16 milligrams, and placebo. This study used a randomized, double-blind, placebo-controlled crossover design in healthy, non-dependent recreational opioid users.

The primary measure of drug liking measured on a bipolar VAS from 0 to 100 with 50 as neutral indicated that the positive control drugs, buprenorphine, at both doses, 8 and 16 milligrams, produce statistically significantly higher mean drug scores of 76 and 82 compared to placebo at 52, which establishes assay sensitivity and validates the study.

Samidorphan plus buprenorphine at all 3 doses; 2 and 2, 8 and 8, and 16 and 16 milligrams; produced mean drug liking scores of 60, 61, and 64, respectively, that were barely outside of the
placebo range.

The secondary measures of overall drug liking, high, good drug effects, and take drug again for buprenorphine produced mean scores on each of these positive subjective measures that were statistically significantly greater than placebo.

Samidorphan plus buprenorphine, again at all 3 doses, 2/2, 8/8, and 16/16 milligrams, produced mean scores on each of these positive subjective measures that were either within the placebo range or barely outside of this range.

As a result, the conclusions from the HAP studies indicate that two single-dose HAP studies conducted with samidorphan indicate that the drug does not produce positive subjective responses. In a third HAP study, the combination of samidorphan and buprenorphine produced positive subjective responses that were much less than those produced by buprenorphine alone. However, they were slightly outside of the placebo range.

Our final conclusions indicate that animal
and human studies consistently show that samidorphan is a mu opioid antagonist with no meaningful abuse potential, similar to the mu opioid antagonist naltrexone.

BUP/SAM produces subjective responses that are lower than those from administration of the same dose of BUP alone, but does retain some clinically significant abuse potential. Overall, the positive subjective responses reported after administration of BUP/SAM showed that the combination has a low potential for abuse.

I will now pass it on to the next speaker, Celeste Mallama.

**FDA Presentation - Celeste Mallama**

DR. MALLAMA: Good morning. My name is Celesta Mallama. I'm a reviewer for the Division of Epidemiology, and I'll be presenting the FDA review of the epidemiologic and surveillance data related to the misuse and abuse of buprenorphine products.

In July of 2017, the National Academies of Sciences, Engineering, and Medicine issued the
report, Pain Management and the Opioid Epidemic: Balancing Societal and Individual Benefits and Risks of Prescription Opioid Use. The NASEM committee's charge was to help the FDA develop a framework for opioid review, approval, and monitoring that balances the individual patient needs with the broader public health consequence of opioid misuse.

The report suggests that such an approach to evaluation of benefit-risk should include assessing evidence of a product's potential for misuse and diversion and predicted risks to family members and society.

The purpose of this presentation is to inform the broad consideration of the benefit-risk balance for buprenorphine samidorphan by addressing the following topics. First, I will present the use, misuse, and abuse of currently marketed buprenorphine and buprenorphine naloxone products, after which Dr. Mark Sullivan will discuss the complex relationships between depression, pain, and substance use disorders.
The goal of this presentation is to provide contextual information for discussion of the overall benefit-risk balance on the product under review. This presentation of real-world buprenorphine abuse patterns is not intended as a prediction of expected abuse of buprenorphine samidorphan or its approval and marketing.

This graph, generated by FDA using IQVIA national prescription audit data shows national projections for the number of prescriptions for buprenorphine-containing oral solid formulations dispensed from U.S. outpatient retail pharmacies. On the Y-axis are number of prescriptions in millions and on the X-axis are years from 2013 to 2017.

These numbers demonstrate that the vast majority of the market for oral solid buprenorphine-containing products consist of products indicated for opioid dependence as part of medication-assisted treatment as opposed to pain, and that dispensings of buprenorphine-naloxone combination products, shown on the graph, in the
solid and dotted black lines, greatly outnumber buprenorphine single-entity products shown on the graph in the solid and dotted gray lines. Buprenorphine transdermal patches, injectable, and implant products are not included in these data.

This slide shows the proportion of respondents reporting past-year misuse of prescription pain relievers in the National Survey on Drug Use and Health. This is a nationally represented population survey of individuals 12 and older in the United States.

Of note, the survey defines misuse as any use other than that as directed by a healthcare provider, and therefore includes non-medical use for a therapeutic purpose such as taking someone else's medication or more than directed to treat pain, anxiety, or withdrawal, or to achieve a high.

The graph on the left shows the percentage misusing each opioid among the total population. As you can see, buprenorphine was less frequently misused in the general population overall than more commonly prescribed opioids like hydrocodone and
oxycodone.

The graph on the right shows the proportion of respondents reporting misuse in people who used that opioid in the last year, including use as directed. Buprenorphine has the second highest proportion of users reporting misuse within the population of people using the specified drug.

An important point to keep in mind when interpreting the data on the previous slide is that the majority of buprenorphine dispensed in the U.S. is for evidence-based treatment for opioid use disorder.

Therefore, compared to the general population who use other prescription opioid analgesics, buprenorphine is disproportionately dispensed to those with existing opioid use disorder who are at elevated risk for misusing and abusing opioids, including through injection or other non-oral routes; or diverting it within social networks with individuals with opioid use disorder.

Therefore, directly comparing levels of
buprenorphine misuse and abuse to other opioids can be challenging and potentially misleading. The same caveat applies to the data on routes of buprenorphine abuse that I will present next.

Several different data sources indicate that abuse of marketed buprenorphine products through non-oral routes is fairly common in cases that come to medical attention among those abusing the products, even when buprenorphine is in combination with naloxone.

This graph, generated by FDA, contains data extracted from the National Poison Data System, a centralized data source maintained by the American Association of Poison Control Centers that captures information on a near real-time basis from a national network of poison centers receiving calls from the public or healthcare workers.

On the Y-axis is the percent of abuse cases reporting exposure by each route and on the X-axis are bars for single-ingredient buprenorphine products and buprenorphine naloxone products. Injection route is shown in black and inhalation
nal route is shown in gray.

About a quarter of the poison control centers' calls from 2013 to 2017 involving buprenorphine abuse reported injection of the drug, and this was similar for single-ingredient and buprenorphine-naloxone combination products. You do also see some inhalation nasal abuse, but it is less frequent than injection.

This graph, also generated by FDA, contains information from the National Electronic Injury Surveillance System, cooperative adverse drug events surveillance, a nationally representative sample of emergency department visits in the U.S., drug from manual abstraction of clinical records by trained coders.

The graph shows percentage of abuse cases reporting injection route on the Y-axis and on the X-axis are bars for single-ingredient buprenorphine and buprenorphine-naloxone combination products.

The NEISS-CADES data source was updated in 2016 to include ED visits related to abuse. In these emergency department data from 2016 to 2017,
more than 40 percent of cases involving abuse of buprenorphine were related to injection of the drug. Many of these were for infectious complications of injection drug use. National estimates for cases involving inhalation abuse did not meet criteria for data precision.

In the 2014 article, Cicero, et al. reported that, in a sample of individuals entering substance use disorder treatment, 34.4 percent of those reporting past-month buprenorphine use to get high indicated that they had injected it in the month prior to treatment.

Among respondents entering treatment centers who indicated buprenorphine injection in the past month, 43.6 percent injected buprenorphine-naloxone tablets. Participants reported a number of simple methods which they believed separated buprenorphine from naloxone, resulting in what they believed to be pure buprenorphine for injection; again, keeping in mind that the population abusing buprenorphine may be particularly enriched with people with advanced substance use disorder who are more likely
to be experienced injection drug users.

To summarize the epidemiological data, the vast majority of the market for oral solid buprenorphine-containing products consist of products indicated for opioid dependence as part of evidence-based medication-assisted treatment as opposed to pain. Abuse of both single-ingredient buprenorphine and buprenorphine-naloxone combination products is common and occurs through both oral and non-oral routes, including injection.

In a survey of patients entering treatment center with previous experience with buprenorphine, respondents reported a number of methods that they believe separated buprenorphine from naloxone, resulting in what they termed to be pure buprenorphine for injection.

However, comparing buprenorphine abuse patterns with those other opioids and using these patterns to try to predict what might happen in patients being treated for depression is challenging due to the high-risk nature of the populations being treated with buprenorphine for
medication-assisted treatment.

The buprenorphine samidorphan under review has a new orally bioavailable antagonist and is indicated for a new population. Therefore, it is not known whether similar misuse and abuse patterns will be seen.

Considering the benefit-risk balance of a new opioid-containing product, it is valuable to understand the population in which it is likely to be used and any safety concerns that may arise from its use in a real-world setting.

In the epidemiologic review, in the FDA background package, we included some information from the published literature describing the complex relationships between depression, pain, and substance use disorders. We have asked Dr. Mark Sullivan, an expert in this area, to speak on this topic, again to inform the discussion on the overall benefit-risk balance for the product under review.

Dr. Mark Sullivan is a psychiatrist at the University of Washington Medical Center for Pain
Relief and Regional Heart Center. Dr. Sullivan is a University of Washington professor of psychiatry and behavioral sciences, adjunct professor of anesthesiology and pain medicine, and adjunct professor of bioethics and humanities.

DR. SULLIVAN: Hello This is Mark Sullivan. Can people hear me?

MS. BHATT: Yes, we can hear you, Dr. Sullivan.

Presentation - Mark Sullivan

DR. SULLIVAN: So I will begin my slides.

I've been asked to talk about depression effects on long-term prescription opioid use, abuse, and addiction, and I begin my presentation with a picture here of a PET scan showing mu opioid receptors throughout areas of the brain thought to be primarily related to depression.

Just a few disclosures, I do have some grants from drug companies, from Pfizer to develop a pain self-management support tool; from Purdue to look at the effectiveness of primary care opioid taper plans; and I have a number of federal grants
in the opioid area as well as some consulting with Aetna about pharmacy benefits; Chrono Therapeutics about an opioid taper device; and with the State of Washington about opioid policy.

So let me begin with an introduction about the epidemiology of chronic pain and depression. There's a great deal of overlap between these two conditions with a greater than 50 percent of prevalence of major depression among patients seeking care in chronic pain specialty settings.

Similarly, there's more than 50 percent prevalence of chronic pain in patients seeking care for depression. This has been examined thoroughly over the years. In a multi-site WHO study, the conclusion was that pain and depression, particularly chronic pain, that it's activity limiting and is mutually reinforcing with depression such that each causes the other.

It has been shown that depression, unfortunately, although it increases the likelihood of opioid therapy, which I will show you data about, it actually decreases the responsiveness of
pain to opioids. That's been shown by I.J. Wasan in experimental study 2005 and in a clinical study in 2015.

Pain also decreases responsiveness to antidepressants and even to combined pharmaco and psychodepression therapy, shown by Matt Bair and Steven Thielke. Chronic pain and depression have quite a similar neuroscience. They show similar patterns of neural activation on functional scans such as fMRIs and PET scans, and they respond to many of the same medications, such as antidepressants.

There is interesting interaction also between chronic pain and substance use disorders. We know, from many studies, that chronic pain is common in patients with substance use disorders. This has been reported between 27 to 87 percent prevalence in chronic pain population, and substance use populations will have chronic pain.

Patients with chronic pain are also thought to be 2 to 3 times more likely to develop substance use disorders. There are not as many studies in
this area. I list a couple of studies here.

Recent reviews of the neuroscience of chronic pain and addiction have pointed towards the reward deficiency as a link between chronic pain and substance use disorders.

Chronic pain is characterized by low dopamine turnover in reward centers, the experience of anhedonia or lack of pleasure, and a set of what are called anti-reward adaptations, where normally rewarding activities no longer are rewarding. This results in high salience of pain relief, but low salience for other rewards, and I list a couple of papers here that discuss this hypothesis.

So what about opioid therapy for chronic pain and depression? It's first important to remember that there are no treatment guidelines recommending long-term opioid therapy for pain and depression. There are in fact no RCTs of opioid therapy for chronic pain in depressed patients. In fact, virtually all of the RCTs of opioid therapy for chronic pain have excluded depressed patients, so in fact we have virtually no evidence base for
that practice, which is common.

Opioid therapy for depressed patients is very common. One study that we were involved in showed that, among HMO patients who had recent depression within the past 2 years, opioid therapy was 3 times more common than patients without a history of depression. In addition to being more common, depressed patients received higher daily opioid doses, more days supplied, and were more likely to get Schedule II opioids.

These findings have been replicated in other samples of commercially insured patients, Medicaid patients, and veteran patients. All of these studies have shown that mental health disorders are especially common in high-dose long-term opioid patients.

My interpretation of this data is that there's a pattern of adverse selection rather than careful selection. Opioid treatment guidelines have generally suggested that clinicians carefully select chronic pain patients without substance use and mental health disorders for opioid therapy, but
what's happening in actual practice is the opposite of this.

So what I have called adverse selection is defined as the selection of high-risk patients with substance abuse and mental health disorders for a high-risk opioid regimen. So what we see is that, among these patients with substance abuse and mental health disorders, they have higher rates of opioid use, especially long term. They have higher daily doses of opioids. They tend to have a longer duration of therapy in days or years. They are more likely to receive multiple opioids at the same time, and they are more likely to receive concurrent sedatives such as benzodiazepines or muscle relaxants.

I recently spoke at the NIH Pain Collaboratory meeting, where I argued that this essentially results in a reverse treatment disparity with vulnerable patients overtreated with opioids, and I've written about this in a number of papers.

Other investigators have also identified
patterns of opioid prescription and receipt consistent with adverse selection hypothesis. Karen Seal from UCSF studied 140,000 Iraq or Afghanistan veterans with chronic pain and found that opioids were received by 6 percent of those veterans without mental health disorders, 12 percent of those with non-PTSD mental health disorders, mostly depression, and 18 percent of those with PTSD. So we get a doubling of the opioid use rate in those with depression and then a tripling among those with PTSD.

These mental health groups were also characterized by receiving higher-dose opioids, multiple opioid sedatives, early refills, and a higher rate of adverse events from the opioid. Davis, et al, published a study explaining that the 16 percent of Americans with mental health disorders receive over half of all the opioids prescribed in the U.S. Odds ratio for receiving opioids for mental health patients was about 2.

A very large study done by Quinn, et al. in 2017 showed that among 10 million commercially

A Matter of Record
(301) 890-4188
insured patients, mental health and substance use disorders were associated with opioid therapy. The likelihood of receiving opioids was about double for those with an anxiety or mood disorder, and it was about triple for those with a non-opioid substance use disorder, and a 9-fold increase among those with an opioid substance use disorder.

Some interesting recent studies have revealed adverse selection of patients with depression into long-term high-dose opioid therapy as a process of self-selection. Halbert, et al. showed that while depressed patients initiate opioid therapy for chronic pain, only slightly more often than non-depressed patients with chronic pain, but they are twice as likely to transition from short-term use to long-term use, and they're self-selecting because they don't quit opioids, which is what most patients do.

Jenna Goesling at Michigan has shown that depressed patients appear to continue opioid use at lower pain intensity levels and higher levels of physical function than do non-depressed patients.
Alicia Grattan with me showed that depressed patients tend to overuse opioids because they use them to treat insomnia and stress.

Unfortunately, opioids are not an effective treatment for depression or anxiety. They have been used for this. One must remember that opioids are the first psychotropic medication really ever discovered and used. And as far back as the Greek physician Hippocrates or the Roman physician, Galen, opioids have been used to treat both mania and melancholia. In fact, if you look in 19th and 20th century psychiatric textbooks, opioids are recommended for these purposes. However, as I mentioned before, there are no controlled studies showing lasting relief of opioids for depression or anxiety.

They may provide partial relief from anxiety and insomnia, but particularly in patients with PTSD, they appear to deepen avoidance and deactivation and prolong the course of PTSD.

Now, it's important to note that buprenorphine may be an exception in this regard.
due to kappa antagonism, which has been written about as a potential antidepressant mechanism by a variety of investigators that I list here.

It's important to also know that mu opioids may also cause depression. Generally, it's been shown that opioid therapy that extends beyond 90 days or has a daily dose over 50 milligrams of morphine or equivalent may increase the risk of depression, according to a series of retrospective cohort studies.

In general, rapid dose increase seems to have the highest depression risk. Most of these studies have been done by Scherrer and colleagues, and they've also shown that recurrent incident and treatment-resistant depression are all more likely in opioid-treated patients.

Other investigators have shown that the most vulnerable patients to opioid-induced depression are those with low-pain self-efficacy; that is, low confidence to continue with daily activities; poor social support, and who began using their opioids at a younger age.
Depression also appears to increase the rates at which patients proceed from use to abuse and addiction. Depression or a psychological behavior called catastrophizing, characterized by hopeless and helpless cognition, increases the risk of misuse, non-medical use, and abuse of prescription opioids among adults and adolescents.

Among adults with non-medical use, depression doubles the risk of progressing to full-fledged opioid use disorder. This in fact may be the path by which depression increases the risk of opioid use disorder among patients with chronic pain, as it pushes them down this path from use to misuse to abuse to addiction or opioid use disorder.

Adverse selection is prominent among opioid-treated patients, but it's even more marked among those patients receiving both opioids and benzodiazepines. We know from national data that opioid prescribing has declined since 2012, but not the prescribing of opioids with benzos.

Between 2001 and 2013, concurrent opioid-
benzo use doubled among privately insured patients, and its estimated increased risk of opioid overdose is between 2 and 10, depending on how well confounding is controlled for. We know that about 10 percent of patients initiating antidepressants also initiate benzodiazepines. That's sometimes due to soothe anxiety or insomnia before the antidepressants kick in. However, about 12 percent of these people who initiate benzodiazepines with their antidepressants become long-term benzodiazepine users.

As we look for recent trends, there's increasing rate of simultaneous antidepressant and benzodiazepine use. Six percent of antidepressant starts included benzodiazepine, also in 2001 up to 12 percent by 2012.

There is evidence of opioid dysregulation in major depression. We see reduced opioid receptor availability in major depressive disorder, and this is particularly marked after sadness induction, as mentioned by these two studies. This reduced opioid receptor availability is associated with
reduced response to SSRI antidepressants and
increased adrenocorticotropic hormone levels.

Greater opioid release after social
rejection is associated with resiliency and lower
affect, so that's a basic release that's adaptive,
and what we see in major depression is a higher
tonic level of opioids, and that's what seems to be
associated with opioid problems.

If we look at the opioid receptor gene,
particularly the G-allele, that is associated with
the kind of psychiatric problems such as higher
neuroticism or a tendency to experience negative
emotional states and higher rates of depression.

We see in these patients with the G-allele,
they have greater reactivity to social rejection,
and this is manifested in altered responsivity of
the anterior cingulate cortex and the insula
cortex, as summarized by Pecina in a recent paper.

So there are some important and relevant
non-pain functions of exogenous opioids. Here, I'm
specifically speaking of mu opioid effects. We've
known for quite some time that mu opioids don't
simply address physical pain, but suppress suppression distress, separation distress, and that's found widely in mammals, whether they be rodents, primates, or human.

We've known and confirmed recently that opioids reduce the affective, emotional, or affective aspect of pain more than the sensory aspect. And as I mentioned, in terms of opioid use among pain patients, it provides more relief of stress, anxiety, and insomnia than mood problems, which may worsen.

Interestingly, in our recent opioid taper study, a number of patients reported that they no longer felt like zombies, and their spouses confirmed they have returned to their pre-opioid personalities, and there have been other evidence that opioids, specifically mu opioids, impair the human capacity for emotion perception and social inferences. This is one Australian study that I cite here.

So in conclusion, the relationship between opioids and depression is close, complex, and
multifaceted. Depression is associated with endogenous opioid dysfunction. Depression is associated with opioid misuse, non-medical use, abuse, and opioid use disorder, and endogenous mu opioids prescribed long term may increase the risk of depression and social cognition deficits.

Thank you. That's the end of my remarks.

**FDA Presentation - Sonya Dunn**

DR. DUNN: Thank you, Dr. Sullivan.

My name is Somya Dunn, and good morning. I'm from the Division of Risk Management. I'm going to present a discussion on risk management for buprenorphine samidorphan.

First, I'm going to present the background on risk evaluation and mitigation strategies or REMS. I'll present a summary of the REMS for buprenorphine-containing products, potential safety concerns associated with the use of buprenorphine samidorphan or BUP/SAM, and possible risk management strategies for BUP/SAM.

I will begin with REMS. A REMS is a drug safety program that can be required by the FDA for
A REMS is designed to mitigate risks associated with drug use and includes strategies beyond labeling to ensure benefits outweigh the risks of the drug.

The FDA Amendments Act of 2007 gave the FDA authorization to require applicants and application holders to develop and comply with REMS program if determined necessary. The FDA has the authority to require a REMS pre- or post-approval.

A REMS can include a number of components such as a medication guide, a communication plan, elements to assure safe use or ETASU, an implementation system, and must include a time table for submission of assessments.

If determined a necessary component of a REMS, the elements to assure safe use can include the following: certification and/or specialized training of healthcare providers that prescribe the drug; certification of pharmacies or other dispensers of the drug; limited settings for dispensing or administration of the drug; having each patient using the drug subject to certain...
monitoring;, the drug is dispensed, administered only with evidence of safe use conditions, for example a pregnancy test or liver function test or enrollment of treated patients in a registry.

Additionally, an ETASU must align with the serious risks listed in the labeling. They cannot cause undue burden on patient access to the drug, considering in particular patients with serious or life-threatening diseases or conditions and patients who have difficulty accessing healthcare.

I will continue on to a summary of relevant REMS. All the buprenorphine-containing products approved for either treatment of opioid dependence or pain and that are intended for use in an outpatient setting are covered by a REMS.

These are the four types of REMS for the buprenorphine-containing products. The first REMS is actually two different programs due to two different sponsor groups, however, the program requirements are the same. The four types of programs are the BTODs, suboxone/Subutex program, the probuphine program, the sublocade program, and
the opioid analgesic REMS program.

The first three REMS are for products indicated for opioid dependence. These products have required SAMHSA training, which I will discuss shortly. This training is not part of the REMS. The suboxone/Subutex and BTOD REMS as well as the opioid analgesic REMS, the circled programs, mitigate risks associated with opioids such as accidental overdose, misuse, abuse, and addiction.

The middle two programs, probuphine and the sublocade REMS, are designed to mitigate risks associated with the formulation of the product. The circled programs are most relevant to our discussion today.

The suboxone, Subutex, and BTOD REMS consist of provider and pharmacy educational materials and an appropriate use checklist. Training is not required as part of the REMS. However, there is SAMHSA training required for providers prescribing for treatment of opioid use disorder.

Buprenorphine-containing products used for opioid use disorder can be prescribed outside of an
opioid treatment program, or OTP, if the provider obtains a DATA 2000 waiver. The Substance Abuse and Mental Health Services Administration, or SAMHSA, manages the Drug Addiction Treatment Act of 2000. They set eligibility and certification requirements for the DATA 2000 waiver, which includes required training. Physicians that hold the DATA 2000 waiver can prescribe and/or dispense buprenorphine products for opioid use disorder in settings other than OTPs.

Buprenorphine-containing products approved for treatment of pain are covered in the opioid analgesic REMS. This REMS requires manufacturers to make training available to healthcare providers involved in the management of patients with pain. Training is not required to prescribe or dispense, and the training focuses on pain management, identifying risk factors for abuse and addiction, how to counsel patients and their families on the safe use of opioids, and fundamentals of addiction medicine.

I will move on to the potential safety
concerns of BUP/SAM. The agency has concerns regarding BUP/SAM for depression. In the U.S., the majority of patients with depression are treated by primary care clinicians. It is important for prescribers to understand that BUP/SAM contains an opioid. Use of BUP/SAM for depression will be chronic, and this could lead to addiction, dependence, or withdrawal.

In addition, there are safety concerns regarding concomitant opioid and benzodiazepine use. Mental illness, including depression, is associated with comorbidities, including pain conditions, opioid use disorder, and anxiety. Concomitant use of an opioid agonist or benzodiazepine could put patients at risk for respiratory depression. Labeling for approved opioid products does address this risk.

Mental illness, including depression, is associated with misuse and abuse of prescription opioids and higher doses of opioids may be needed for analgesia if patients are taking BUP/SAM.

There are safety concerns in women of
childbearing potential. The prevalence of moderate to severe major depressive disorder is 9.3 percent in women ages 18 to 39. Women who are pregnant may be putting their unborn infants at risk for development of neonatal opiate withdrawal syndrome or NOWS. This was not evaluated in the clinical program for BUP/SAM. Labeling for approved buprenorphine provides addresses the risk of NOWS.

The agency is also considering that patients often struggle with adherence to therapy for many chronic conditions, including depression. Given that BUP is a partial agonist and SAM is an opioid antagonist, it is unclear what impact lack of adherence to BUP/SAM could have if concomitant opioid agonists are used.

There is a possibility that changing mu agonist and antagonist effects from inconsistent use of BUP/SAM could increase the risk of withdrawal or potentially increase the risk of overdose and respiratory depression, particularly if higher doses of opioids are used to overcome antagonist effects.
Now, I will discuss risk management. The applicant's proposed REMS mitigated the risks of misuse and accidental exposure. Training is made available to healthcare providers. Training is not required to prescribe or dispense.

The agency is concerned about risks for BUP/SAM, including those associated with opioids as well as other potential safety concerns. Prescribers of BUP/SAM for depression will not be required to take training under DATA 2000 because it is not indicated for opioid use disorder.

In conclusion, if approved, FDA will likely require a REMS for BUP/SAM and is considering how REMS can best address the potential concerns with BUP/SAM in the indicated population and whether training for potential prescribers is necessary as part of the REMS.

Clarifying Questions to FDA and Guest Speaker

DR. NARENDRAIN: We'll now move on to clarifying questions. If people could specify if they have clarifying questions for the agency or
for Dr. Sullivan, that would be good. And before you speak, please state your name for the record. And if you can and you have a specific presenter, please address that person as well.

I'm going to start with Dr. de Wit.

DR. DE WIT: This is Harriet de Wit, University of Chicago. I have a question about the FDA presentation. On slide number 8, this is the abuse liability, could you just clarify?

It looks like the percent of abuse cases reporting injection route are the same for buprenorphine with and without naloxone. Is that correct? So the naloxone is not serving as a deterrent for intravenous use?

Do you need me to tell you who I'm addressing?

DR. NARENDRAN: Who to address it to.

DR. MALLAMA: I can address the question. That's correct. We saw, in the percentage of abuse cases reporting, injection route, that it was similar for buprenorphine, single ingredient, and buprenorphine-naloxone, in the NEISS-CADES data and
the emergency department that's being displayed.

DR. DE WIT: Does that raise an issue with the combination of buprenorphine and samidorphan if it were to be used intravenously, that maybe the samidorphan wouldn't block the mu agonist effect?

DR. MALLAMA: So this presentation on buprenorphine-naloxone combination products can't necessarily speak to what will be seen with samidorphan. It's just to give context of the real-world setting of what we're seeing in terms of buprenorphine abuse, although they are similar in terms of the fact that they both are combination products with an antagonist included.

DR. STAFFA: This is Judy Staffa. I'd also just like to add to that. What we were trying to do was to show that in the real world, many people assume that products that have naloxone added, that that is a deterrent to non-oral routes of abuse, but that's not what we're seeing.

So the other study I think from Cicero suggests that people have developed methods for separating, but we really don't understand fully
what's happening. We just wanted to introduce that
as something to consider. We really don't know
whether the same thing would happen with this
product.

DR. CHIAPPEROINO: Hi. This is Dominic
Chiapperino, controlled substance staff. So
samidorphan would still have an antagonist effect
by the IV route. We don't have data by the IN
route.

DR. NARENDRAH: Dr. Kulldorff?

DR. KULLDORFF: Thank you, Martin Kulldorff,
biostatistician at Harvard Medical School. I had
one additional question for the applicant that I
didn't get through last time, and it's for Dr.
Schindler.

It was stated that the clinically important
effect size was somewhere between 1.5- and 2-point
reduction on the scale that you're using. But if
one looks at the point estimates and the confidence
intervals for the three studies, 205, 206, and 207,
it's clear that the effect size that the study was
powered for -- and I assume you did a power
calculation beforehand, is around 3 or so.

So if a clinically important effect size is somewhere between 1.5 and 2, why did you power the study for an effect size around 3?

DR. NARENDRAN: Does the sponsor want to answer that question?

DR. SCHINDLER: I can just say you're correct that the studies were powered for an effect size of about 3, but we actually observed a smaller effect that showed a significant difference. So the power is really, as you know, our ability to detect a difference if it's there, and we were able to detect a difference.

The second part of your question, I think, is also what is a clinically meaningful difference, and I would defer to my clinical colleagues to discuss what a clinically meaningful difference would be. Maybe one of my statistical colleagues has some additional comment to add.

DR. NARENDRAN: Very brief, very brief.

DR. MEMISOGLU: So you asked about a rationale behind powering it based on a delta of 3,
so that was based on our experience in the 202 study and what we observed from MADRS-10.

DR. NARENDRAN: I think that addresses it.

I think I'm going to go back to the agency's questions because it's their time.

Dr. Jain next. One second. If anybody else has a question for Dr. Sullivan, he's going to have to run in 15 minutes. He's on the phone. So does anybody have a question for Dr. Sullivan of all the lists I have here?

(No response.)

DR. NARENDRAN: No? Okay. Thank you.

Dr. Jain?

DR. JAIN: Felipe Jain, Harvard Medical School. One of the two trials, the 207 trial that this applicant has presented as one of their pivotal efficacy trials, utilized a late change in the primary outcome measure prior to unblinding.

How unprecedented is a request for such a change? Is the FDA aware of other trials within the context of treatment-resistant depression that have requested a change such as this. And did the
applicant present any published literature
supporting their use of the change to this measure?

DR. TEMPLE: This is Bob Temple. I will try
to answer. I don't think we keep a list of these
things, but sometimes, other data, other studies,
and things like that convince people that the
endpoint they were using isn't the smartest one to
use. And as long as we're absolutely positive
there's been no breakdown of the data and nobody
knows what it is, we don't have any systematic
objection to that. It's not common in trials.
Trials are ongoing. You don't usually change the
endpoint.

I wish I could think of them, but there have
been other cases where people have changed the
endpoint. I don't know if -- Tiffany, you're
looking like you know some.

DR. FARCHIONE: Not that I know something
about a specific example, but I think that the
concern here is that it's not an issue that the
change was late because, again, it was made while
the study was still blinded, but with it being that
late, there really wasn't an opportunity for us to provide feedback on those changes.

    DR. TEMPLE: That's certainly another critical question. You plan the study. We've usually seen the protocol. We've said something about the endpoint, and how they change it; they don't know what we're going to think. But that's a somewhat different question. And we worry tremendously about whether there was leakage of the interim data. That would invalidate any change that anybody wanted to make, and it's not always easy to know.

    DR. NARENDRAN: I've been told Dr. Sullivan may have left.

    DR. CELLA: I do have a question for Dr. Sullivan. Are you still there?

          (No response.)

    DR. NARENDRAN: Sounds like he's not available. I don't know if the agency -

    DR. CELLA: Maybe I'll ask it anyway.

          (Laughter.)

    DR. CELLA: Well, it's a question that maybe
the sponsor can answer or maybe FDA can answer. So
David Cella from Northwestern University in
Chicago. As far as I know -- is that Dr. Sullivan?
No. As far as I know, we don't know about any
exclusion criteria related to pain, and pain was
not assessed in the trials.

If that's correct, I was going to ask
Dr. Sullivan to what extent he thought this
combination might be treating pain and not
depression. But maybe other reviewers could
comment.

DR. NARENDRAN: I don't know if the agency
can comment or applicant. Was pain addressed?

DR. VON MOLTKE: We can comment on
inclusion/exclusion. Go ahead, Dr. Stanford.

DR. STANFORD: Pain and pain disorders were
not exclusionary to be enrolled in the program.

DR. NARENDRAN: Thank you.

FEMALE VOICE: But concomitant opioid use
was, right?

DR. VON MOLTKE: Yes.

DR. NARENDRAN: Next question, Dr. Dunn?
DR. DUNN: Walter Dunn, psychiatry, UCLA.

Two questions about the actual trial, so for study 202, coming back to the question of potentially throwing out that site 124 patient, if the sponsors had prespecified the primary outcome as only stage 2 outcomes, is that something the FDA would have possibly considered? And I'm assuming that if that patient was thrown out, that wouldn't have affected just purely stage 2 outcomes.

Second question regarding study 207, my understanding is that the longer duration of the second stage made the analysis somewhat complicated. So was there an analysis done just on week 5 for 207, and would the outcomes have been the same if we looked at 5 weeks for stage 1 and 5 weeks for stage 2?

Then the third question is regarding the withdrawals. So for the assessment of the COWS, it was a little unclear to me what the time frame was. So my understanding was that the patients were tapered for about a week, and then the COWS were assessed at the end of 2 weeks.
Is that correct? And then is there any sense of how long the withdrawal persisted or maybe these mild withdrawals persisted, or is it only a single time point?

DR. OGBAGABER: So for study 202, the patient that was singled out and excluded, and we did sensitivity analyses, he was or she was only in the first stage and was not enrolled in the second stage because he was dosed to the 2/2 arm. And only placebo non-responders would go on to the second stage.

DR. DUNN: If it was prespecified that the primary endpoint was only stage 2 participants, is that something the FDA would have possibly considered as a valid endpoint or would it have to be a combination of both stage 1 and stage 2?

DR. OGBAGABER: Yes. I mean, the SPCD would have to combine and weight the stage to kind of make inference.

DR. DUNN: Regarding study 207, maybe redoing the analysis with just 5 weeks of stage 2?

DR. OGBAGABER: For the MADRS average, I
think they still went, but I'm not sure. The
sponsor can speak on this, if they did the 5/5 for
study 207 for on MADRS endpoint.

DR. VON MOLTKE: We'll address that. Go
ahead, Dr. Memisoglu.

DR. MEMISOGLU: Asli Memisoglu,
biostatistician, Alkermes. Can I get E-107? So
the question was whether having an equal duration
in the number of weeks included in the average gave
a consistent result, and it did.

Can I have the slide up, please? So what
you see here is the MADRS-10 analysis. The top
line is the average endpoint to end of treatment
that we've already presented, and the bottom line
is just limiting the analysis to weeks 3 to week 5
in both stages, and you can see that the results
are consistent.

DR. FARCHIONE: I think he was also asking,
though, about if you use week 5 as the end of
treatment. Right?

DR. DUNN: Actually, yes, they addressed
that question, yes. It's just limiting the
averages to 3 and 5 for both stage 1 and stage 2.

DR. FARCHIONE: So you only wanted to know about the average, not if they just chose week 5 as end of treatment?

DR. DUNN: Actually, that's an excellent point. Yes. If you just ended it at week 5 for stage 2, what would that have looked like?

DR. MEMISOGLU: I'd also like to comment on the first part of your question about the stage 2 and study 202. And in that case, the stage 2 analysis was positive on its own. I think the p-value was --

DR. DUNN: Right. And then in terms of just the endpoint, at week 5 for stage 2, for 207, not the average, but actually just the endpoint?

DR. MEMISOGLU: Right, and that did not meet the threshold.

DR. DUNN: That did not meet criteria. And the final question was about the withdrawal and the COWS assessments, 1 week of taper and then withdrawal assessment at 2 weeks. Is that correct?

DR. LEE: I would invite the applicant
DR. VON MOLTKE: For the sponsor, I'll ask Dr. Stanford to come on up.

DR. STANFORD: To answer your question, the taper only existed in 202, so the main analysis for withdrawal was in the pooled 205, 206, and 207 studies, where there was no taper and it was abrupt discontinuation. And then for 205, there was one week follow-up, and in 206 and 207, it was a 2-week follow-up.

We also did the same analyses in our long-term data, where there was three time points; 1 day, 1 week, and 2 weeks after discontinuation of study drug. In that long-term study, the exposure was up to a year.

DR. NARENDRAN: Next question, Dr. Joniak-Grant?

MS. JONIAK-GRANT: Elizabeth Joniak-Grant. Was a dossier for the validity of using the MADRS-6 submitted as was recommended by the FDA?

DR. FARCHIONE: Yes, it was. I noted that it was submitted shortly after the breakthrough
therapy advice.

MS. JONIAK-GRANT: Can you give a little more detail as to why it was still deemed, in the FDA's view, as not valid?

DR. FARCHIONE: The main reason was because of the four missing items, so without assessing four key symptoms of depression, we didn't feel like it accurately reflected an improvement in depression overall. And I think, yes, we have Yeh Fong Chen here from our clinical outcomes staff, so he can speak more to that.

DR. CHEN: They submit a proposal to use MADRS-6. And we said, well, please submit the evidence to support MADRS-6 is as we have [indiscernible] MADRS-10. And what they submitted is actually quantitative study of effective analyses in some correlations, basically to show -- the main argument is that MADRS-6, the 6 items is unidimensional and the 10 items is not unidimensional.

So based on the unidimensional assumption, then the total score for the 6 is actually more
appropriate in that they show -- actually the
result that is submitted does not confirm whether
it's unidimensional. It confirms MADRS-6 is
unidimensional, but did neither confirm or not
confirm MADRS-10 is unidimensional or not
dimensional.

But as Dr. Farchione mentioned, the most
important thing is lack of four items, and that is
very important in terms of diagnosis of depression
and then also the recovery of depression symptoms,
and it's also in the DSM V.

So we evaluate the content validity first.
We don't think that MADRS-6 covers all the content
that is important to the patients. And then given
the quantitative [indiscernible], it did not fully
support that MADRS-6, and MADRS-10 is not
unidimensional.

So basically, it's just that it might be
actually all the symptoms. The 10 symptoms is a
continuance of the severity of depressions, and
then it lacks the more severe -- the MADRS-6 lacks
the more severe side of symptoms.
Then we see a lot of studies -- actually they found MADRS-10 is unidimensional and some MADRS-10 is not unidimensional. And I think that's because the inclusion criteria of whether the patient have more severe or less severe. When they have less severe patients, there is less report the appetite, the sleep problem, the suicidal thoughts, and so it doesn't show up as unidimensional, but I think it's a continuum.

DR. NARENDRAN: I do want to reiterate one person, one question. Let's go ahead.

Dr. Conley? We still have a lot of questions to go.

DR. CONLEY: Well, actually, I do have a couple of concerns about this. It's more than one question, so sorry, but I'll do one.

DR. NARENDRAN: Maybe we can come back in line after that --

DR. CONLEY: At the end.

DR. NARENDRAN: -- if we have more time.

DR. CONLEY: So I'll start with the one that I think is probably the most important to the
spon.

In Dr. Dunn's presentation, you gave a good presentation about, just in general, what REMS are, but not really anything that I picked up because the sponsor did propose a REMS of whether what they were proposing was acceptable to the FDA or what your thoughts were about it, so I assume that's going to be a committee charge.

DR. DUNN: Can you guys hear me? This is a little tall for me. We did meet several times and discuss the REMS and what type of program we would be putting together or having the sponsor put together. But at this point, we still are not decided exactly what the program would consist of.

So I don't think we're putting it out to you as an exact question. It's more of like a discussion point of what do you think. But where we were focused was these programs that have required training versus programs that don't. I was trying to make that point in my presentation.

So that's something that, in particular, we want you to think about, because of the risks that we know of with opioids or opioid products, in this
particular product, are some of those concerns more concerning?

That was the point that I was trying to make, such as women of childbearing potential, and NOWS, and concurrent benzodiazepine use, just things that we already know about, but are we more concerned in this particular population, and if we are, then do we want to require training or not. It's just something to think about in terms of the program.

DR. NARENDRAN: Ms. Witczak?

MS. WITCZAK: Kim Witczak, consumer rep. This is in line with REMS. What is the agency in terms of direct to consumer advertising? What are the guidelines there, as well as communicating out to the patient organizations like National Alliance for Mentally Ill, who could be out there as promoting this as a possible treatment for their patients.

Is that going to be part of it or is that part of the discussion? Because as somebody has spent her entire career in advertising and
communications, I saw in the presentation about J.D. being the ideal candidate, I could see all kinds of things from a marketing point of view, that I would be out, saying you're not going to gain weight; all of those things that are going to be very attractive to people who may not want to go on a antipsychotic.

DR. DUNN: I'm sorry. Specifically what's the question?

MS. WITCZAK: Well, with direct-to-consumer advertising, are there any kind of guidelines or parameters around communicating to patient organizations that will be out there as well?

DR. FARCHIONE: That wouldn't be part of that. The REMS would be more related to the physicians and the prescribers. We do have a separate group that reviews the marketing materials to make sure that there's nothing false and misleading in there.

MS. WITCZAK: So I guess my question is, can advertising happen right away when the drug got approved? Can it be out there, being communicated
to the general public at the same time that there's still the REMS that's going to the prescribers? So that might be more for you guys internally.

DR. FARCHIONE: The company would submit their initial marketing proposal to us for review.

MS. WITCZAK: Okay. So because it has a REMS, you can't advertise it for a year? So those are some of the things that I think are important to the general public.

DR. DUNN: The REMS programs have nothing to do with advertising. The communications that are under the REMS programs are the letters that we send to healthcare providers or healthcare provider organizations to give them information about the REMS program specifically and that discusses the risks and requirements of the program or whatever training is available through the program.

That's all the communication materials that would be under the REMS, and for patients as well.

DR. UNGER: This is Ellis Unger. There's no difference between a drug with and without a REMS in terms of advertising if that's your question.
DR. NARENDRAN: Next question, Dr. Ruha?

DR. RUHA: Michelle Ruha. My question is for the FDA. I just want to clarify something about the MADRS-10 average. I think I understood from the sponsor that using the average at each assessment is more reflective of possibly real-world treatment because there's really no end of treatment. It's really the end of the study, but in the real-world, it would be the end of the treatment.

I believe the FDA did not necessarily agree with using the MADRS-10 average, and I just want to clarify that and clarify why. Is it because it wasn't previously used in any studies, or it's not validated, or did I misunderstand that?

DR. FARCHIONE: There were a few concerns that came up during the course of discussions with the applicant. One of the issues is we've never used an average before. Now, it's not completely out of the question that that might be a way to increase power when you have variability and all of those things. But every other drug that's been
approved for this indication has managed to win at the end of treatment, so there was that concern.

Then there was also the issue that with this being an opioid, is there some tolerance that develops, and then you have a decrease in efficacy that would be masked by an averaging strategy.

They did show the data on the one side where the patients who did continue on the combination continued to be okay, so that relieved our concern on that issue a little bit. But the difficult thing here is that we've got a study design that's intended to reduce the impact of the placebo response on the study, and then you have this still difficulty separating, and then you throw in this averaging strategy, which now also kind of mitigates things.

There are just so many different ways of slicing up the data. We wanted to minimize the number of bells and whistles, I guess you could say.

DR. NARENDRAN: I'll just ask my related question. In the other medications that have been
approved for adjunct treatment, what was the primary outcome measure? How long did those trials last?

DR. FARCHIONE: Six to eight. It was either 6 or 8. I can double check.

DR. NARENDRAN: It's fine. You can let me know. You can let us know later. The next question is Dr. Iyengar?

DR. IYENGAR: Satish Iyengar from the University of Pittsburgh. The FDA has brought up two rather general issues of how to deal with missing data and also the possible correlations. What I also saw, though, in the presentation was that the bootstrap verification and also the careful depiction of the drop-outs indicated that although there may be a violation of assumptions, it was not a serious issue.

To what extent are these general issues applicable to the studies that these people have done?

DR. OGBAGABER: Regarding the MAR assumption for study 207, actually, if we can go to backup
slide 13, 13 and 14, what we did was a graphical
depiction of these spaghetti plots for the drop-out
reasons and not really analytical.
And there is no method to check MAR using an
analytical method as you might know. It's all
graphical.

From here, drop-outs who behave -- within
each treatment group, we are assuming that they
behaved the same; each completer in the treatment
groups as if they hadn't dropped out. That's the
assumption and it's reasonable.

What was the second question?

DR. IYENGAR: The second question dealt with
the assumptions that you were checking with the
bootstrap. I think it was differences in variances
and also the correlation issue.

DR. OGBAGABER: Right. The bootstrap, slide
number 9 and 10, number 9. So for the bootstrap,
because the assumption for the applicant that they
use, using the method mentioned, was based on
normality assumption. So the only way we could
think of was a bootstrap to kind of check and go
over if that assumption of the covariance between
the two treatment stage effects being none, zero,
we used bootstrap, and we didn't make any of the
normal assumptions, and the conclusion remained the
same.

DR. NARENDRAN: Next question, Dr. Besco?

DR. BESCO: Thank you. Kelly Besco, Ohio
Health. Knowing that the current buprenorphine
products are being misused as often injected, it's
possible to separate these two products with
certain solvents. And in order to evaluate if this
product is safe, I think it would be helpful to
know what those solvents are and if they're common
household products.

I recognize that this is a public forum and
that we probably can't disclose that because of the
current buprenorphine products that are on the
market, but I just wanted to note for the record, I
think that would be helpful for members of the
committee to know what those products are for our
evaluation.

DR. CHIAPPERINO: This is Dominic
Chiapperino. You are correct that we can't speak to those issues during the committee meeting.

Thanks.

DR. NARENDRAN: Next question, Dr. Crawford?

DR. CRAWFORD: Thank you. Stephanie Crawford from the University of Illinois at Chicago. This question is directed to Dr. Ogbagaber.

In your presentation, because I think it's pretty salient for our discussion later, going back to your slide 8 regarding study 202, the single subject from site 124 who had more extreme results, the FDA Office of Scientific Investigations questions the subject's eligibility from your slide based on incomplete and/or contradictory source information.

Is it possible for you to share more to clarify for us what does that contradictory source information and/or is it something very major that was incomplete data?

DR. OGBAGABER: A staff from OSI is coming to get to that question.
DR. KRONSTEIN: My name is Phillip Kronstein from the Office of Scientific Investigations, letting you know that we conducted an inspection of that site, and yet incomplete means that for this subject, there was no collateral information that was not required by protocol. But if you have no collateral information, then you are relying simply on what the subject reports.

The subject in one part reported that they'd been -- according to the sub-investigator who saw the subject, reported the subject had been on fluoxetine for 10 weeks prior to enrollment in the study.

However, as part of the study, it was called the failsafe, a safer interview was done. This safe interview, as many may be aware, is a totally independent interview, I think out of MGH, if I remember correctly, and we reviewed the safer interview. On two places on the safer interview, it is written that the subject was only on fluoxetine for 1 week.

Then there was an addendum on the safe
interview. This addendum was written 2 days later, where we said, spoke to the clinic, and the subject was on fluoxetine for 10 weeks.

We presumed that the clinic was not the subject's psychiatric clinic, but rather the clinic of the investigator. And the reason we presumed that is because the subject indicated on a form, they did not want their primary care physician or their psychiatrist, if any, they didn't indicate, contacted.

So there was no explanation for this discrepancy. In fact, the safe interview is supposed to be an independent interview. And if indeed they did contact the clinical investigator site, it would no longer be an independent interview.

So this raised concerns about whether the subject was truly eligible. And the fact is, based on, again, these contradictions, you could not determine basically two possibilities. Number one, the subject was eligible; number two, the subject was on fluoxetine for only 1 week. If the subject
had been on fluoxetine for only 1 week, of course
some of the results you may see may have been due
to fluoxetine and not the product.

So there's no additional information, and we
say we cannot determine whether the subject was
eligible.

DR. NARENDRAN: We have three more questions
for the agency that I want to get, and then after
that, we'll start the open public hearing at 1:15,
and maybe give extra time before the charge is
issued to respond to this.

Is that okay?

DR. VON MOLTKE: May I speak to the chair
after to be recognized later?

DR. NARENDRAN: After the open public
hearing, we can bring you guys back to address
that.

DR. VON MOLTKE: Okay, because we really
would like to clarify this since we've sent
information to the agency on this. Thank you.

DR. NARENDRAN: Thank you. Dr. Acri, your
question?
DR. ACRI: Thank you. Jane Acri from NIDA.

And my question has to do with the degree to which the effects of buprenorphine are blocked by samidorphan. In the briefing materials, there was one graph that showed that the effects on pupillary constriction was completely blocked by an equal dose of samidorphan with a dose of buprenorphine. And in the human abuse liability study, also, the 1 to 1 ratio blocked the liking effects of buprenorphine, which was important.

But then there was something in the briefing materials also that showed an in vitro functional study showed that maximal stimulation of mu opioid receptors produced by buprenorphine was only halved by co-administration or by samidorphan. It wouldn't be administration in vitro study. So it looked like the effects of buprenorphine were not completely blocked by samidorphan as shown in the other two figures.

In addition to that, I think we haven't really discussed the effects of this active metabolite that is a mu agonist, that shows
accumulation over a period of 7 days. We've looked
at the abuse liability study that appeared to have
been done with acute dosing, but it seems to me
that with chronic dosing, you might get
accumulation of this metabolite, plus you have
incomplete blockade of the μ receptor.

So I'm just wondering if we've done a full
enough evaluation of the μ opioid effects?

DR. CHIAPPERINO: Those are excellent
questions. Regarding the studies that we have, we
don't have studies looking at buprenorphine/
samidorphan effects in ratios other than 1 to 1,
and that's an important point, particularly with
Dr. Meisel's points earlier about the potential to
manipulate the product.

As far as the μ opioid agonist metabolite,
we don't have data speaking directly to that other
than the adverse event profile in the chronic
studies. Of course, those are done in the patient
population.

So we don't see a higher incidence of
euphoric events in those studies. We did not see
them in the phase 1 studies to any noticeable degree. So whatever the implications of there being a metabolite that is a mu agonist, it's not showing up in pharmacodynamic effects. So there's an absence of data that may make us more comfortable for sure.

DR. ACRI: Right. I'm also wondering about its contribution to the withdrawal effects. The COWS data was not gathered in a consistent way from one day to the next, so I think it was the FDA that presented the fact that it was sort of sporadic. You couldn't see how much withdrawal people were having over time. And the fact that you're getting any withdrawal at all is kind of a concern and makes me think there's more mu receptor stimulation there than meets the eye.

DR. CHIAPPHERINO: That's a good possibility.

DR. NARENDRAN: Next question, Dr. Hernandez-Diaz?

DR. HERNANDEZ-DIAZ: Sonia Hernandez-Diaz from Harvard Chan School of Public Health. I have a question about how the FDA interprets this
stage 2 part of the study or any other study that
requires taking a placebo for some weeks, and among
those who do not respond, then go ahead and
randomize them.

With the interpretation of those results for
labeling or for reporting the results, would you
say then that the results will affect a population
that after being exposed to something in the
placebo or sugar pill do not respond, those are
ineligible for that medication?

It seems to me that is such a selective
group that the results from that part of the trial
or any trial, that that reflects a completely
different population.

So first, how do you then report and
translate that into clinical practice when
recommending the medication? And also, how are you
going to combine that stage 2 with stage 1 when
there are two different populations? We would not
pool that data. It's like they are coming from two
different populations, the second one being
selected in post-randomization.
So I don't understand why we are spending time talking about the type 1 error in that design when it's a different question, a different population, what we are doing in stage 2.

So I guess two questions. Why are we discussing type 1 error rather than questioning whether they can be combined, and two, how could you translate the results from stage 2 type of population into recommendations for treatment to a population that after being exposed for weeks doesn't respond to placebo?

DR. TEMPLE: It's a good question, but I would distinguish two aspects of it. One is whether the drug has an effect in somebody, which is a very important part of our conclusions about whether a drug should be approved. The second is how do you label it accurately?

Just to give an even worse example, from your point of view, probably, we accept the idea of a randomized withdrawal study; that is, you take people, put them on drug. You then take the responders and randomize to continued therapy or
placebo. We've described that in a number of places. It's an enrichment strategy, if you like. But then, once you've established that the drug works by that, how do you decide how to label it? Who's it for? What can you expect?

The answer is not easy, but at least you know the drug works. So we tend to live with some of those uncertainties. Usually, there's more than one trial, not just the randomized withdrawal study.

But in our labeling, we would always say how the study was done. We would say only a hundred people were given the drug; only 20 appeared to respond, and the randomized withdrawal study was done in those. So you'd have some idea of what the likely response rate is.

But your question is a good one. If you use any of these maneuvers, how do you label it exactly in a way that tells people what would happen in ordinary life? And my short answer is, we live with that because establishing that something works is really important even if you don't exactly know
how to tell people how to use it.

Is that enough?

DR. HERNANDEZ-DIAZ: Yes. Thank you.

DR. NARENDRAN: Thank you. Last question,

Dr. Meisel?

DR. MEISEL: Steve Meisel from Fairview in Minneapolis. We've spent an awful lot of time
today talking about studies 202, 205, 206, and 207,
but not study 208. 208 is the open-label long-term
study, and I realize it's open label. But I wonder
if the agency would comment on the findings and any
conclusions that we can draw from 208.

My look at this, half of the patients
dropped out over the course of the year. And of
those, there was a 60 percent response rate,
meaning that there was only 30 percent. So how do
we take the short 5-week study and then look at the
52-week open-label study, and make something of
that? Are there any conclusions we can take from
that study 208?

DR. FARCHIONE: Typically, what we are using
those longer term open-label studies for is just
for long-term safety data. We're not really looking to them for support of efficacy at all.

DR. MEISEL: So we should take no conclusions whatsoever about efficacy based on 208?

DR. FARCHIONE: We aren't taking conclusions on efficacy from 208.

DR. MEISEL: So all we have are the 5- or 6-week trials. So the corollary to that; are the 5-week trials sufficient for an antidepressant?

DR. FARCHIONE: So I did go back to look at the other three that are approved for adjunctive treatment of depression, and all three of them were 6-week trials.

DR. MATHIS: This is Mitch Mathis. We don't think that short-term trials adequately describe any of the diseases that we approve medications for, so we always as a post-marketing commitment will then ask for some longer term data.

We in the past have taken this issue to a different body and AC, and asked them should we have maintenance data before we approve a drug acutely, and the answer was no; it will take too
long. We need new therapies, et cetera.

So our tradition is to do the short term, get the approval based on short-term data, and then we have some long-term safety data that goes with that. That's what this 208 is. And then look for a randomized withdrawal trial post-marketing, should the drug be approved.

DR. NARENDRAN: I think we could stop here. I'd like to do a 40-minute break if that's okay, shorten the break. After the open public hearing, if it's okay with the division, people who have extra questions can probably ask it before the charge. I'll let the division decide on that, if they're agreeable to that. So 40 minutes from now, we'll meet. That's 1:35.

We'll now break for lunch. We'll reconvene in this room 40 minutes from now at 1:35. Please take any personal belongings you may want with you at this time. Panel members, please remember that there should be no discussion of the meeting topic during lunch amongst yourselves or amongst members of the audience. Thanks.
(Whereupon, at 12:53 p.m., a lunch recess was taken.)
AFTERNOON SESSION

(1:35 p.m.)

Open Public Hearing

DR. NARENDTRAN: We're going to start again.

Both the FDA and the public believe in a transparent process for information gathering and decision making. To ensure such transparency at the open public hearing session of the advisory committee meeting, FDA believes that it is important to understand the context of an individual's presentation.

For this reason, FDA encourages you, the open public hearing speaker, at the beginning of your written or oral statement, to advise the committee of any financial relationship that you may have with a sponsor, its product, and if known, its direct competitors.

For example, this financial information may include the sponsor's payment of your travel, lodging, and any other expenses in connection with your attendance at the meeting. Likewise, FDA encourages you, at the beginning of your statement,
to advise the committee if you do not have any financial relationships as well. If you choose not to address this issue of financial relationships at the beginning of your statement, it will not preclude you from speaking.

The FDA and this committee place great importance in the open public hearing process. The insights and comments provided can help the agency and this committee in their consideration of the issues before them.

That said, in many instances and for many topics, there will be a variety of opinions. One of our goals today is for this open public hearing to be conducted in a fair and open way, where every participant is listened to carefully, and treated with dignity, courtesy, and respect. Therefore, please speak only when recognized by the chairperson. Thank you for your cooperation.

Will speaker number 1 step up to the podium and introduce yourself? Please state your name and any organization you are representing for the record.

A Matter of Record
(301) 890-4188
DR. FAVA: I am Maurizio Fava. I am the vice chair of the Department of Psychiatry at Mass General, director of the Division of Clinical Research at the MGH Research Institute, Director of Clinical Trials Network and Institute, and associate dean of clinical translational science at Harvard Medical School.

I have been involved in the research that has been presented today. I've helped Alkermes design their studies, but I don't have any equity on Alkermes. I have never received any personal honoraria for my work. And I've always worked on behalf of Mass General. And I paid for my own travel to come here to make a comment on this.

My conflict of interest is that I'm the co-inventor with David Schoenfeld from Harvard School of Public Health and Mass General of the Sequential Parallel Comparison Design.

My first statement is that I think, as clinicians, I'm a clinician, we struggle with treatment-resistant depression. We have over 40,000 people who commit suicide each year, and a
substantial proportion -- I think someone was calculating STAR*D. We calculated that, at the end of the four stages or levels of STAR*D, about 40 percent of the patients, 35 to 40, had not really achieved remission despite multiple shots on goal.

So treatment resistance is a big problem, and what we have thus far clearly does not address a significant portion of our patients, so new tools, new therapies are critical.

Dr. Sullivan mentioned that opioid dysregulation happens in depression. And depression is very heterogeneous. Not everybody will have an opioid dysregulation, but certainly, a significant proportion of patients do have opioid dysregulations.

That is why, in my opinion, depression is a risk factor for opioid use disorder. Any risk factor for opioid overdose is intentional and unintentional. So if you add the 40,000 suicides with the 40,000 lethal overdoses, this is a big problem that we're facing as clinicians.
One of the first patients that our depression program lost to suicide was someone with treatment-resistant depression, was not responding to monoamine therapies, and became an opioid addict and committed suicide. So we know these patients exist.

Dr. Sullivan mentioned something that is not quite correct. There is tianeptine, which is a mu agonist. It has been proven extensively by Rennehan and others. It is a proven antidepressant in a number of European countries and Asian countries.

So the opioid mechanism has been exploited. The problem with tianeptine as a mu agonist is the potential for abuse. There have been many cases of abuses in eastern Europe of tianeptine. So this drug with the combination of a mu antagonist really creates a deterrent for that, to the point that I like to point out that the one patient who had opioid use disorder had went into a significant withdrawal the moment the patient was started on the study, suggesting the fact it is clearly a
I think Dr. Dunn's observation is absolutely correct with respect to the phase 2 study, that even if you remove the subject in that academic side -- and by the way, someone said -- because we're responsible for the independent interviews.

When we look at patient-reported history, patients are sometimes not great historians, so we do call the site to get documentation information about it. So it's still independent, but we want to get additional information.

But in the phase 2 study, I like to point out that there was a 4-point difference in stage 2, a very robust 4-point difference in stage 2 without that subject. So I think it was an absolutely correct observation, and the study remains positive whether or not -- in fact the statistical elimination of the worst and the best patient still maintains statistical significance.

So in my mind, I feel that the evidence is there. You have two positive studies. And yes, with innovations embraced by alchemists, the
sequential study, the averaging -- by the way, the averaging of multiple visits is common in pain studies.

So I think there were innovations, but I think that embracing innovation is better than having a 50 percent failure rate, which is common in CNS trials. Thank you.

DR. NARENDRAN: Thank you. Will speaker number 2 step up to the podium and introduce yourself? Please state your name and any organization you are representing for the record.

MR. COUNTS: Hi. Nathaniel Counts, Mental Health America, and Mental Health America paid for my time and travel to be here today.

Good morning. I'd really like to thank the committee for its time and effort in considering this for this full day. We really appreciate it. I'm the senior policy director at Mental Health America, which is the nation's leading community-based nonprofit dedicated to addressing the needs of those living with mental illness and promoting overall mental health of Americans.
Our work is driven by our commitment to promote mental health as a critical part of overall wellness, including prevention services for all early identification and intervention for those at risk, integrated services, and support for those who need it with recovery as the goal.

I was hoping to really underscore some of the need because I know you've heard lots of prevalence data today, but I wanted to kind of dive a little bit deeper on that. So as just alluded to, depression is one of the leading causes of disability and a key contributor to two ongoing epidemics, opioids and suicide.

I think the thing we wanted to underscore is even though there is lots of therapies out there already, lots of antidepressants, and options, and research, these problems do persist. So our view from MHA is that it isn't going to be just about identifying needs as early as possible and connecting people to the medications that already exist, but there is actually, we think, a real need for additional innovation.
Well, I'll kind of dive into it. I know we've heard about STAR*D this morning, but the Lancet in 2018 published the large network meta-analysis of antidepressants and their effectiveness. And it was exciting because it finds odds ratios going from 1.37 to 2.13 across all antidepressants.

So all of them improved depressive symptoms, which I honestly always as a lay person have trouble interpreting odds ratios. But a BMJ article then goes on to say -- and I'll quote this so I don't mess this up as well -- "All antidepressants in the meta-analysis worked. They all significantly increase the odds ratio of a HAM-D score reduction by 50 percent, and across all of them, the odds ratio was about 1.66, reducing it by 50 percent."

To put the odds ratios, though, in a more clinically relevant context, we need to know what proportion get better in the placebo group. This information was not provided, but other researchers suggest that 30 to 40 percent of the placebo group
participants report improvement in remission.

I actually saw in Psychological Medicine there was a meta-analysis suggesting about 53 percent, even as high as -- they're using this typical placebo response; when you convert the odds ratio of 1.6 to 10, it ends up being 10 to 12 percent more people in the treatment group benefit compared to the placebo group.

So this raises like a very real issue with the effectiveness of existing monotherapies and the need for additional innovation in this space. And I felt this from personal experience as well. I struggle with the depression of taking multiple antidepressants at all different doses, and levels, and everything to absolutely no effect.

So I'm one of the other 50 percent for which I didn't achieve spontaneous remission and for which medications weren't effective, and this can be really hard to have ongoing depression. But I'm extremely lucky and have way more protective factors than most.

I've actually never shared this before.
This is like a totally bizarre context to be sharing these things, and especially after hearing all this epidemiologic data about the different constructs involved in depression. So not everyone's as lucky as me, and I think having additional information will really be critical.

I think 10 to 12 percent just can't be acceptable for how effective our existing medication is. For every patient population, we can't just have 10 to 12 percent of people achieving response of existing therapies.

I think through methods of existing adjunctive therapies and pairing them with other things, we can probably get -- let's see if we even double or triple it, and that will still be 20-30 percent of additional effectiveness.

But I think we're really far from where we need to end up being. And I think it's even more horrifying when you really think about what depression is as a condition. The diagnostic criteria itself are feelings of worthlessness or guilt almost every day, marked diminished interest.
for pleasure in almost all activities nearly every day, and recurring thoughts of death or suicide.

    I mean, thinking about what that's like to experience that level of pain every day is extremely intense. It's even worse if you think about 16 million adults every year do experience this.

    I think the most tragic thing that I've been seeing is CDC data from BRFSS is showing that it's trending in the wrong direction for adolescents. All these risk factors are going down. Depression is increasing over the past 3 years, though, and suicidal ideation is increasing. And these adolescents are going to become adults, and then we're going to have a larger population of depressed adults.

    So I think that we do have a need for innovation and really thinking about how can the pipeline lead to better, newer therapies that do help people and begin to alleviate the crisis we're facing. So I really appreciate the time all of you are taking today to think this through. Thank you
so much.

DR. NARENDRAN: Thank you. Will speaker number 3 step up to the podium and introduce yourself? Please state your name and any organization.

MS. WALKER: Hi. My name is Ali Walker. I am here on behalf of the American Foundation for Suicide Prevention as a volunteer. My time is not being paid for, and I have no financial disclosures to provide to you.

I am by trade a critical care physician assistant. I'm also the chair for the National Capital Area Chapters local chapter of the American Foundation for Suicide Prevention. And most importantly, I'm a suicide loss survivor.

As of 8 years this past Sunday, I lost my friend, Roma, to suicide. And I hope to elucidate that I agree with what's been said so far about the extreme importance of innovation as far as treating major depressive disorder because it is a huge risk factor for suicide.

Also, on a systemic level, worldwide, it is
the 11th leading cause of disability in the world.

In the United States, it's the second leading cause of disability, which I found to be quite unfortunate. Not only is it common, but it's persistent. Over a 2-year period, more than 40 percent risk of recurrence for those who have a single major depressive episode. And for those who have a second episode, their risk of recurrence for subsequent major depressive episodes is greater than 75 percent.

I did some research on the history of antidepressant development over the years, and I was quite fascinated to find that the first antidepressant medication was actually invented by mistake in the late 1950s, upon the discovery of euphoric side effects of iproniazid, which is being used to treat tuberculosis.

Unfortunately, given the side effects of a lot of drugs in this class, interactions with tyramine, which is found in delicious things like wine and cheese, those medications really were not ideal, so there was an ongoing pursuit to develop
more antidepressant medications.

We saw a lot of innovations coming out from the 1980s, late 1980s into the 1990s. In 1987, fluoxetine, the first SSRI, was released, and really was ground breaking at the time. Unfortunately, though, since the early 2000s, really, it seems there has not been a whole lot of new discoveries. We've seen a number of drugs come on the market that are enantiomers or have different salt compounds in the composition of those medications, but ultimately are not significantly different from things that are already existing.

Also interesting to me was the efficacy pattern over the year. So the early iproniazid, drug makers actually reported about a 75 percent or 70 percent response rate for patients taking that for major depressive disorder.

I heard the mention of the STAR*D trial. I actually did look at that study by the NIMH from 2001 to 2006, and only saw about 30 percent response rate in first rounds of treatment for 8 to
12 weeks for standard therapy, which is awful, not acceptable. As a clinician, for all of the other health situations that we treat, we would never be satisfied with that. Even after multiple rounds of therapy, the most resistant patients, only about 67 percent, achieved a significant response.

So to put this in a more personal context, I wanted you to consider a scenario for a moment. The debilitating nature of depression is very real, and I'll give you the example of my grandmother, who over the summer was faced with the realization of her own mortality after my grandfather was admitted to the neuro ICU under the assumption that he had weeks to live from a leptomeningeal disease.

In fact, he just had a very unusual presentation of Lyme disease. Fortunately, he was very easily treated. But after his release from the hospital, my grandmother still persisted to have significant mood changes. She was by all definition depressed, lost interest in engaging with her friends. She's a person that's out on the golf course every day, organizes fundraisers, very
active individual, and she was having a hard time getting out of bed in the morning, was not taking phone calls when I would call her, which is highly unusual for her.

Eventually, with some encouragement, we got her to go see a physician, who recommended that she start an antidepressant. So they did take her and start her on mirtazapine, but after a couple of weeks, actually, into this medication, she was noticing a lot of abdominal cramping, diarrhea, vomiting, the typical G.I. symptoms that are very commonly associated with antidepressants.

So she got fed up because her mood had not improved and now she has all of these uncomfortable side effects and tossed the bottle in the trash, and that was it. And that's not an effective way to manage a condition that has completely turned around your life and interfered with your ability to enjoy the things that you do.

This scenario is very common, and I have many friends and family members who struggle with depression. I myself have struggled with
depression upon the loss of my friend to suicide. And it's unfortunate to me that I have to, on a regular basis, really fight to push people to comply with medications being prescribed to them in the hopes that they might improve on antidepressant medications, knowing that they'll wait 6 to 12 weeks on average to have a response in their mood, and knowing that in that time frame, they're subject to a number of side effects from G.I. discomforts and sexual dysfunction, and other things that really do not contribute to positive living.

So I will leave you with that, but I just urge you to recognize that there is so much need for innovation in the field of mental health and in treatment of major depressive disorder.

DR. NARENDRARAN: Thank you. Will speaker number 4 step up to the podium and introduce yourself? Please state your name and affiliation for the record.

MR. CORSETTI: Thank you. My name is Gordon Corsetti. I'm here on behalf of the American
Foundation for Suicide Prevention, and I have no financial relationships to disclose.

I'm a depressive. I suffered for years until I learned how to attack my depressed thoughts and gain mastery over them. I deal with anxiety. I'm naturally fearful in unstructured social situations and very bright strobing lights, not these, can trigger panic attacks that paralyze my body and terrify my mind.

I live with suicidal thoughts, but I do not act on them. I remain alive despite my brain's best attempts to kill me. But here's the rub. I never would have learned how to attack my depressed thoughts, how to breathe through my anxiety, and how to be okay with a suicidal thought without my antidepressants.

I live with my mental illnesses now, but years ago, I truly suffered. My first experience with an antidepressant came during my freshman year of college in 2006. An emergency room doctor prescribed me the lowest dose of Paxil. And I swallowed that pill every morning, and 2 weeks
later while walking across campus, my thoughts just stopped.

Standing in the middle of campus, I waited for an intrusive thought to pop into my head, but no thought came. I didn't have a single horrible thought in my head. And this wonderful sensation of peace settled over me, and for the first time, I walked forward without fear in years.

Since that day, I've been on SSRIs, SNRIs, atypical antidepressants, and even a souped-up version of vitamin D called Deplin. All of these medications came with various side effects. My weight went from 150 pounds to 195 pounds and back. My desire for food and intimacy were drastically reduced. Worst of all, though, was a persistent sense of emotional numbness.

I accepted these side effects because the net gain that I received gave me a much greater quality of life. Still, I hated needing these pills to feel or act normal. So I did what many others have done. I stopped taking my medication. I thought I could manage my symptoms through sheer
willpower.

Over the next several years, I started and stopped many different medication regimens, and each time I stopped, I spiraled into a progressively worse depression. Three times, I tried to take my own life.

Now, some would argue that antidepressants cause me to attempt suicide. That is an incorrect argument. Withdrawal from antidepressants contributed to my attempts to kill myself. I stand here today and proudly declare that while I have experienced some of the worst consequences of antidepressants, through their proper use, I have experienced more joy than I ever thought possible.

These medications gave my mind space to learn positive coping strategies, to build a strong support network of family and friends, and ultimately to gain agency over my illness. If several hundred years ago, I went to a medical professional and told him my symptoms, he would sit me down and drill a hole into my head to release the obvious demons from their imprisonment within
my skull. Two-hundred years ago, I would have been strapped to a table so that the black bile from my veins could be let out through bloodletting, and of course, I would feel better as a result of that.

In a psychiatric hospital maybe just 60 years ago, I might have been lobotomized or given this extreme electroshock therapy if only to make me more compliant with the staff.

In 1987, a year before I was born, SSRIs were introduced. SNRIs followed just a few years later. Today, as a 30-year-old man, I am being treated with medications that were discovered while I was wearing diapers.

Now, I'm grateful that I live in an era that favors medications over drilling into a skull to treat the mentally ill, but my hope is that a few years from now, we'll look back on the last 30 years of antidepressant use and think, this is how we treated depression. We have so many more options now.

It is time that we take the next step in treating mental illness in this country. It is
time to invest in research to widen available treatments. It is time for a new generation of antidepressants. Thank you.

   DR. NARENDRAN: Thank you. Will speaker number 5 step up to the podium and introduce yourself? Please state your name and organization for the record.

   DR. BODKIN: Hi there, all. This has been fairly inaudible up until now. I hope that's audible. I'm Alec Bodkin, J. Alexander Bodkin. I'm a psychiatrist on staff at McLean Hospital. I run the clinical psychopharmacology research program and probably know more about the use of buprenorphine as an antidepressant than any living person, and we'll get to that.

   I'm here on my own dime, but have consulted Alkermes. In fact, I gratis-consulted to them as they were first thinking about this possible project because it seemed like an important way of making safer a very effective treatment that had big problems associated with it.

   But in any case, I've been prescribing to
appropriate patients, i.e., very treatment-refractory patients for whom ordinary monoaminergic antidepressants just don't work, buprenorphine, and other clever things, but buprenorphine specifically for the past 30 years, when the dear departed Jonathan Cole and I put together the first American study of buprenorphine in depression.

Since then, I have consulted on countless cases. I've always had a handful of buprenorphine-treated patients in my own practice. And it has serious problems, but it also has awesome efficacy when it is needed.

One of the things that's been utterly missed in this discussion, that's always missed, is that there are responders who do extraordinarily well and then there are ordinary patients who maybe do so-so, but there may be side effects. They're all blended together as though they have the same problem, but depression is a multifaceted, many different syndrome thing, really requiring many different alternative treatments.

In any case, I'd like to get through my
written little things. I put so much work into it and, if there's any time at the end, I'll speak my mind beyond that.

Recently, I had the good fortune of being able to treat 7 subjects in Alkermes' very long-term, year-long study of safety, efficacy, and tolerability at 2 milligrams of -- , I used to call it Alks 5461, but going forward, it'll be BUP/SAM.

These were not very treatment-refractory patients, as I customarily had treated with buprenorphine, but they were treatment refractory enough that SSRIs and related medicines didn't really do the job adequately or didn't do it at all.

So of our N of 7, 3 of them didn't do particularly well. I mean, that's the way it goes, and 4 of them did splendidly. So 57 percent of my little N experienced remarkable degrees of recovery, the kind that one rarely sees, that persisted after treatment was completed.

Anyway, all of them, I will note, were astonished at how well they felt, not at the
beginning of treatment. Unlike buprenorphine, which causes a little bit of euphoria with dose number 1, BUP/SAM causes nothing at dose number 1. I mean, it's a little nausea sometimes, but it's invisible. It creeps up on you.

But over time -- and it creeps up rather slowly, faster than SSRIs, faster than MAOIs, faster than tricyclics, but it creeps up slowly so that it really wasn't until about week 6 that my patients began to note that they felt better than they ever had in their lives. And this really persisted the remainder of the year.

We had one guy, a very troubled man, an artist, who really didn't get well for 6 months, but at that point was transformed. His life is completely changed. His art has risen to high new levels. He's gotten married. And he feels he could not have done that without this help.

So anyway, when this stuff was stopped, nobody had a withdrawal syndrome, unlike buprenorphine, where that's the terrible problem with buprenorphine. You can't get off it. If
you've been well for a year and a half, you want to
get off it, you get so miserable as you get down to
that last quarter milligram that you just say,
"Forget it; ill just stay on it," and that's what
people do.

There's no such problem with this compound
at all. It's invisible. You stop it and nothing
changes. You don't even lose the mood benefit.
And in fact, I spoke to my patients within the last
week getting ready for this and people still feel a
year later like their lives are simply better than
they had been, worth noting.

Anyway, let me briefly spell out some
important clinical properties of the Alks 5461 just
for the assembled. The therapeutic effects of the
drug in a significant portion of treatment-
refractory patients, or resistant patients with
depression, a significant proportion have extremely
robust benefits that persist long after the drug
treatment has been concluded.

This compound causes no acute euphoria. It
really causes no euphoria, and it has no more abuse
potential than an SSRI, and that is no
exaggeration. It has no immediate effects on mood
at all, and in fact, everyone had to have first
patients come in for their first -- well, they came
in for this anyway, but we stayed with them until
it had taken effect, and nobody felt anything other
than a little nausea sometimes.

This stuff differs from buprenorphine, which
causes euphoria, though it fades soon. And the
only real problem with that is that, in
irresponsible patients, they want to hang on to
that fading euphoria and they push the dose and
they push the dose.

So you have to be very careful who you bring
into treatment with buprenorphine and stay away
from people who have a tendency toward drug abuse.
That's just a requirement. People who are going to
pursue that buzz will not do well, nor will their
providers.

But unlike buprenorphine, its effect rose
over months. As I said, it takes about 6 weeks;
one patient took 6 months. And this would
certainly make the drug unattractive to the potential drug abuser who wants to feel a quick lift, because there is no quick lift.

What else have we got here? Just that it is extremely important, about this drug, that it causes no withdrawal syndrome at all, even after an entire year of daily use at 2 milligrams, which is a non-trivial dose, and there was no dosage tapering in the study, and wellness continued indefinitely.

So the great limitation on buprenorphine, apart from its abuse liability, is the chain that it puts around people who are on it can't get off it, which is not the end of the world if you need it, but it is a burden that this compound [indiscernible].

But I will only say in closing -- don't want to take up too much time here -- is that we really need to attend to the reality that in depressive illness, there are robust responders to different treatments. Occasionally, people will flower on Prozac, not very frequently. Occasionally, people
will flower on Nardil, on the MAO inhibitors, which were felt for a long time not to be very good because most people didn't flower on them, but people who got well got extremely well.

Well, people who get well on this stuff get extremely well, and it's important to keep that in mind. It's quite harmless. If it isn't going to have any street value, buprenorphine is cheap stuff, so no one's going to spend a lot of money and time trying to extract a little bit from this. It's just widely available. It would be silly and a waste of money.

So this is not going to add to the drug problems of our nation, and it really will give people who otherwise can't get well, because their problem relates to basically the dynorphin system created by kappa receptors.

DR. NARENDRAN: I think your time's up.

DR. BODKIN: By fighting those receptors, people get well. Anyway, I should go. Thank you all so much.

DR. NARENDRAN: Thank you. Will speaker
number 6 step up to the podium? Please state your name and institution for the record.

MS. WICKMAN: My name is Kathryn Wickman, and I live with depression. I'm speaking today on behalf of the Depression and Bipolar Support Alliance. I have not received any compensation from DSBA or Alkermes.

You've heard a lot of science and data so far about mental illness and possible treatments. I would like to put a face to those statistics. I am here to talk about the anguish and emotional pain that is my depression and why we need more treatment options.

I have been living with the disease of depression for 36 years from the age of 14. On my worst days, depression meant lonely hours, deciding if I had the strength to continue living, with the seemingly never-ending unbearable burden of anguish and hopelessness. No one should need to live with a pain so strong that death seems to be the only escape.

Other times, depression meant crying with
gasping tears in the stairwell at work, then
scraping myself together to get to my desk, where I
would lie and say my red eyes were due to
allergies, not tears. I would tell people the
reason I couldn't think clearly enough to function
at work was due to the allergy medications, not due
to my depression and the antidepressants that were
clearly not working.

My brain was so foggy, I could accomplish
nothing more than watch the clock count down to the
end of the day, when I could finally crawl back
into bed. I needed then and need now reliable,
affordable treatments to prevent these down days.

In my 30s, it took 5 tormented years to find
a treatment that would subdue my depression. I am
the poster child for treatment-resistant
depression. During that time, I tried over
40 medications and supplements. Not only did they
not work, but I had to live with unacceptable side
effects. I needed more treatment options.

Finally, at the age of 38, after living with
depression for 24 years, I found a medication that
worked. I had my Wizard of Oz moment, where the movie starts in black and white and switches to color. I realized everything had been black and white in the past. I did not know what it meant to live without the burden of depression.

With a medication, suddenly, I saw color. Though the depression was not completely gone, I then had about 11 years of roughly stable moods. Then in 2017, I was diagnosed with breast cancer. My oncologist told me, on top of everything else, I might have to change one of my antidepressants because it conflicted with a cancer treatment. I warned her I was scared of what her treatments might do to my moods. I was not afraid of the cancer that might kill me. I was terrified of the depression that might make me want to kill myself.

As I feared, my depression came back during the cancer treatments. Ultimately, I had about 10 months of a severe depression, 3 months of which were debilitating. At the start of this depression, I thought, at least this time it will be easier to find a solution. This time, they must
know more about mental illness. They must have
more treatment options for me now.

I was disappointed to learn this was not the
case. Between having already tried and failed with
most of the medications the decade earlier and now
needing treatments that were acceptable to both my
psychiatrist and oncologist, I had very few
options.

My psychiatrist finally said we were down to
my last alternative. I needed to do
electroconvulsive therapy. I was indignant. This
was 12 years after my last major depression. That
was 12 years to do research, 12 years for testing,
but I was in the exact same place of no viable
treatments.

The best news my psychiatrist could offer me
was that the methodology of pulsing electric
current through my brain was now less severe than
it used to be.

We needed better treatment options when I
was living with severe depression from the ages of
14 to 38, and we needed better treatment options
8 months ago, when I had another incapacitating depression as a result of the cancer treatments. Unfortunately, I suspect I will have at least one more debilitating depression in my life. I ask that you work to make sure there are better options when that happens. I must have better treatment options by them because the failure of mental illness treatments could mean life or death to me. Thank you.

DR. NARENDRAN: Thank you. Will speaker number 7 step up to the podium? Please state your name and organization for the record.

MR. SCHARF: Good afternoon. My name's Eric Scharf. I am the advocacy advisor for the Depression and Bipolar Alliance. However, I am providing testimony on behalf of an individual who was planning to be here today and could not attend at the last minute. His name is Allan Sweet. He lives with depression. I am speaking on behalf of the Depression and Bipolar Alliance and Alliance of Disabled Veterans. And I am speaking on behalf of him as not received any compensation from DBSA or
Alkermes for his appearance.

"Over the years, I have faced many challenges for my attaining and maintaining wellness, as I have worked on being stable with my major depressive disorder. I have a variety of symptoms, but I believe that these 7 have plagued me the most: low self-esteem, disturbed sleep, irritability, anxiety, guilt, lack of motivation, and suicidal thoughts.

"Being a veteran has allowed me access to mental health services that have been helpful most of the time. I have been prescribed various medications over the years for my depression and sleep. Though I was unclear when the depression was first onset, I do know that after a parotidectomy I had in 2007 at the Milwaukee VAMC, my depression symptoms increased.

"Since that time, I have participated in support groups, seen a number of mental health professionals, one-on-one counseling with a social worker, and, again, a variety of medications. During this time of year, I'm also plagued with
seasonal affective disorder, where the days seem dark most of the time and I find myself irritable and full of guilt.

"My irritation comes from the lack of good sleep, which leads to guilt, where I believe I am not fulfilling my responsibilities. Though I attempt to practice good sleep hygiene, eat nutritionally, and exercise, I sometimes lack motivation fulfilling as part of my wellness routine. I try to see my mental health professional every 90 days and often cancel, only to reschedule appointments again due to lack of motivation and not practicing my wellness routine.

"During this time of year, I use a sun lamp in the morning for 30 minutes or so as suggested by the social worker I was seeing. And though this helps, I am not consistent, leading me to entertain suicidal thoughts primarily surrounding my self-worth and, again, self-esteem.

"Depression impacts my daily living, keeping me from making good decisions in the area of employment and following a wellness routine. I
find myself withdrawn and isolating myself from others. The medication helps others, but I have yet to use the right combination that helps me get started in the morning, which leads to long days and even longer nights from lack of sleep and following my wellness routine. Living with depression is a challenge and I will not give up, as I know the person I can be and become."

Thank you on behalf of Mr. Sweet.

DR. NARENDRA: Thank you. Will speaker number 8 step up to the podium? Please state your name and organization for the record.

MR. SPERLING: Good afternoon. My name is Andrew Sperling. I'm the director of legislative advocacy for the National Alliance on Mental Illness, and as an employee of NAMI, they are paying my time to be here today. I have no other disclosures.

First, let me start with NAMI's perspective on depression. The story you just heard from Kathleen [sic], I'm sad to tell you, is rather the rule rather than the exception for people living
with treatment-resistant depression. It is enormously frustrating to be told to take a medication and deal with monotherapy over and over, and not see yourself get better.

It's an enormous challenge. We know that 16 million Americans experienced depression within the past 12 months, 7 percent of our nation's population. You've heard already about the public health burden. I'm going to particularly focus on the burden associated with suicide.

In the United States today, mortality from suicide now exceeds that of both breast cancer and prostate cancer. Those lines crossed just a few years ago. You look at the advances we've made in early identification and more effective treatments for breast cancer and prostate cancer, but yet suicide continues to go along, like over $40,000 a year.

With treatment-resistant depression in particular, it's very, very challenging. As the STAR*D study demonstrated nearly a decade ago, only one-third of patients actually get better on
monotherapy. So what happens the other two-thirds? They need adjunctive therapy. Unfortunately, only three medications with an on-label indication for adjunctive therapy are antipsychotics.

We know about the side effects associated with these medications, including weight gain, metabolic syndrome, are extremely challenging and make it very, very difficult for patients. Monotherapy is futile for the majority of patients living with treatment-resistant depression, so we need better therapies.

I also want to note for the record that I was on the patient stakeholder advisory group for the last three prescription drug user-free agreements; 4, 5, and 6. As part of that, we work with many of our colleagues in the patient advocate community across all disease states to try and bring some changes to the FDA around patient-focused drug development, to get this agency to look more at surrogate endpoints, adaptive clinical trial design, other things to get the voice of the patient into the process of scrutinizing new drug
applications in the work that the FDA does.

Unfortunately, that doesn't seem to have really seeped in here in the psychiatric drug division office, and it's quite frankly telling in some of the back and forth between the sponsor and the agency that's taken place here today.

In August, NAMI submitted comments to the agency on draft guidance for industry on major depression, and we had some challenges there, and I'd be happy to share that testimony with the committee, particularly around some things that FDA needs to be doing to address placebo effect.

You've heard some of this here today, why we have a higher incidence of placebo effect with all psychiatric disorders, but particularly with depression, and investing and adopting things such as Sequential Parallel Comparison Design, and other types of trial designs that can actually address placebo effect would be an advance forward.

Just as we did in our submission on the comments on the guidance for industry, we will again urge the agency to look at ways to use
adaptive clinical trial design to get around placebo effect, to really spur development of new medications, and also surrogate endpoints.

We need to get beyond depression scales that have existed for the better part of 60, 70 years and have better -- look at clinical endpoints, not only that matter to patients, but modernize the way we're measuring depression and measuring the symptoms of depression.

We at NAMI strongly support innovation. The treatments we have available, particularly for treatment-resistant depression, simply aren't good enough. Patients are demanding better treatments and the opportunity to live a full life in the community. Thank you very much.

DR. NARENDRAN: Thank you. Will speaker number 9 step up to the podium? Please state your name and organization for the record.

MR. POLLOCK: My name is Michael Pollock. I serve as the chief executive officer for the Depression and Bipolar Support Alliance. DBSA has received funding from Alkermes to support our
education and outreach programs. DBSA is the leading peer-directed national organization focusing on mood disorders, depression and bipolar.

Unlike any other organization of its kind, DBSA is created for and led by individuals who themselves have a mood disorder diagnosis. This first-person lived experience informs everything we do.

In the 60 years that have passed since the first antidepressant medication was approved by the FDA, there have been significant advances in the scientific understanding of depression, yet treatment options that support individual definitions of wellness remain elusive for many. Information has been incremental. People electing such treatment are consequently frustrated by and losing hope of a pharmacologic solution.

The first priority for treatment is ensuring that a person living with depression or bipolar disorder is provided a pathway out of crisis and onto stability. However, all too often, this baseline stability is also the end game established
for successful long-term care.

In a DBSA-distributed survey this past August that realized over 6,000 responses, nearly a third reported having 10 or more discrete periods of severe depression, and 36 percent indicated that its impact is persistent.

DBSA believes that every person deserves the opportunity not just to survive, but to thrive. And to do that, we need to ensure true wellness as the end goal for mental health treatment.

Additionally, the idea of wellness cannot be embraced without considering the whole health of the individual. The comorbidities associated with depression are not insignificant. The prevalence of depression among individuals living with heart disease, diabetes, polycystic ovary syndrome, CPOD, movement disorders, and Alzheimer's, just to name a few, is well known, and the effect depression can have on the positive outcomes of comorbid conditions is significant.

Choosing between effective treatment for a comorbidity in mental health is counterproductive.
Individuals living with mental health conditions on average die 25 years sooner, and as a result of co-occurring conditions that can be either exacerbated by depression or could exacerbate the depression, the cost of settling for reduced symptoms is simply too great. And for many, it can be a matter of life and death.

Even more challenging than understanding the whole health ramifications of pharmacologic interventions associated with comorbidity is the realization that no one medication typically provides an entire range of symptom relief for depression.

Additionally, these interventions have differing risk-benefit tolerances for each individual. Further, the considerations around medication risks and benefits can often be different. The prescriber may approach the challenge from a clinical perspective, symptom relief, while the patient on the other hand, may be seeking well-being outcomes.

When seeking solutions, patients weigh the
risks and benefits of that intervention against
symptom relief, functional considerations, and
those well-being outcomes. Added to this decision
is the fact that an intervention may not be
consistent in both its symptom relief and side
effects among the patient population.

This often results in a frustrating trial
and error period for both prescribers, who want to
help their patients, and the patient who's looking
for improvement. Unfortunately, during the trial
and error period, many patients reach a point where
they abandon hope and not just a pharmacological
intervention, but in any type of treatment.

The mood disorder community feels abandoned.
Many researchers consider depression a problem
solved due to the number of pharmaceutical
interventions currently available, but nothing
could be further from the truth.

If I've communicated anything today to this
committee, it's first that patients' voices count.
Patients want and need solutions that support a
pathway to wellness. One size does not fit all.
Solutions are complex as the individuals seeking them. And individuals will evaluate the risks and benefits of those solutions based on their own life circumstances.

When considering this application, I urge this committee to not abandon the patient community. Thank you.

DR. NARENDRA: Thank you. Will speaker number 10 step up to the podium and state your name and organization for the record?

MR. MADIGAN: Good afternoon. My name is John Madigan. I'm the senior vice president and chief public policy officer for the American Foundation for Suicide Prevention, which paid my way here today. And I've lived in D.C. for 41 years, and my Uber driver took me around the Beltway on a trip that I could not repeat, but that's a fun note.

I've been in my position now for 10 years. My work is both professional and personal. I lost my sister to suicide 21 years ago. As many of you know, AFSP funds research to improve interventions,
train clinicians in suicide prevention, and 
advocates for policy that will save lives.

AFSP is creating a culture that's smart 
about mental health. AFSP brings people who have 
been affected by suicide out of the darkness and 
give them opportunities to help others like many of 
my volunteers here today.

Many of you know there's no single cause for 
suicide. Suicide most often occurs when stressors 
and health issue coverage create an experience of 
hopelessness and despair. Depression is the most 
common condition associated with suicide, and it's 
often undiagnosed and untreated.

Conditions like depression, anxiety, and 
substance problems, especially when unaddressed, 
increased risk for suicide, yet it's important to 
note that most people who actively manage their 
mental health conditions go on to a full and 
engaging life.

This is why I'm here today. The National 
Center for Health Statistics issued a report 
February 13th of this year, indicating that 1 in 12
U.S. adults reporting having depression. Fifty percent report some degree of difficulty with work, home, or social activities; 30 percent reported moderate or extreme difficulty.

The World Health Organization reports that depression has risen to be the current leading cause of medical disability worldwide. This is likely the result of several facts because of the disabling nature and symptoms of this serious condition because depression is so highly prevalent, because stigma keeps more than half the people with depression from pursuing treatment, and importantly -- this is really important -- because treatment options are currently limited in effectiveness and in options.

As the former lead lobbyist for the American Cancer Society in the 1980s, I was pleased to read FDA Commissioner Scott Gottlieb's public comments on September 13, 2018 at a Friends of Cancer research conference. Much like cancer, a diagnosis of depression can for some people present a grave situation. For many patients with depression,
their mental suffering is as painful, and in some cases, more painful than physical health conditions.

At that September conference, Dr. Gottlieb said, I quote, "We at FDA want to take every chance we have to foster and maximize the kinds of innovations that will make these opportunities available to patients." He was of course referring to cancer.

I ask today, why not apply the same opportunities for innovations and opportunities for patients suffering with depression or other mental health conditions. The FDA should do everything possible to create an environment where companies are willing to make important investments in a new generation of antidepressants that are safe and effective.

In 1970, the American people made clear their desire to battle cancer deaths in the United States. President Nixon responded during his January 1971 State of the Union address, "The time has come in America when the same kind of
concentrated effort that split the atom and took man to the moon should be turned towards conquering this dreaded disease. Let us make a total national commitment to achieve this goal."

The Food and Drug Administration in 2018 and beyond can lead a battle to create new opportunities and innovations for depression drug treatments that just might help us win the war against suicide and mental illness. Thank you.

DR. NARENDRAN: Thank you. Will speaker number 11 step up to the podium? Please state your name and organization for the record.

MS. KENNEY: Good afternoon. My name is Lauren Kenney, and I'm here on behalf of the American Foundation for Suicide Prevention, and I have no financial benefits to disclose.

I have been surrounded by mental illness and depression as long as I can remember. I lost my dad to suicide when I was just 9 years old. In the 20 years since he has been gone, I have struggled with my own mental health, and in 2016, I was given an official diagnosis of moderately severe
depression.

After meeting with my treatment team, it was decided that medication was going to be a good option for me. I had tried talk therapy, and although it was helping, I needed additional assistance in order to function on a daily basis.

Finding the right medication to help me manage my depression symptoms was a trial and error process. It took me and my doctor time to figure out the right medications as well as the proper dosages. This process came with side effects and many, many frustrations.

That begin said, we were able to find something that worked for me, and slowly, the relief started to come. After being at a point where I was unable to sleep, unable to get out of bed, and at times unable to physically move, shower, or get dressed, the medication allowed the fog in my head to clear. Instead of constantly being on the verge of tears, my eyes cleared and I was able to function. Eventually, I was even able to function again at a high level.
Earlier this year, I changed my full-time career and started working for a mental health awareness nonprofit. Every day, I hear from people who are living successful, happy lives while living with mental illness. They share their stories with me, and many of their stories include their treatment plan, which more often than not include medication to help manage their symptoms. It's clear to me that proper medication is crucial in the management and treatment of depression.

For me personally and many folks I know that have shared their stories, medications have helped them live a more fulfilling satisfying life. However, I also know that I'm a lucky one. I was able to find medication that worked for me, and the results outweighed the side effects.

Along with all the people who have shared their story of a successful treatment plan, there are those who have shared their story of all the different medications they have tried, that have not given them relief from their depression. And there are even more stories that we have yet to
here because right now, their struggle is too much; their pain is too debilitating.

These folks are some of the strongest people I know because despite finding little to no reprieve, they keep fighting. It is for those people who are continuously fighting to find relief for their indescribable pain and depression, that I ask you to vote yes to new medications to treat depression. They deserve a chance to live a life with depression that is manageable and not all consuming.

We all deserve to live a fulfilling life without having to battle the agony of depression, and I know that new medication could help make that possible for those who continue to suffer.

As a suicide loss survivor, I know how devastating mental illness can be and how it can affect family and friends of those we lose. We owe it to those fighting through the pain every minute of every day and to their family and friends who do everything they can to help them to allocate our resources to properly treat depression. Not only
could new medications alleviate the daily struggle for those with depression, but new medication could also help us save lives.

Over the years, we have developed many new ground-breaking treatments for other diseases, and now it is time we focused our research, energy, and resources to fighting depression.

My struggle matters. The struggle of those who are unable to speak to you today matter. My depression matters, and the depression of those who have been unable to find the proper treatment matters. And I along with others who are brave enough and strong enough to live with depression every day should have the opportunity to live better with the help of new, more innovative medications. Thank you.

DR. NARENDRAN: Thank you. Will speaker number 12 step up to the podium? Please state your name organization for the record.

DR. SAMBUNARIS: I'm Dr. Angelo Sambunaris. I'm a psychiatrist. I practice clinical medicine in Atlanta, Georgia. I'm also a clinical
investigator for various pharmaceutical clinical trials and also investigator-initiated trials with universities, currently Duke University, Medical College of Georgia, studies that we design ourselves. And I'm an adjunct professor at the Mercer University College of Pharmacy.

My disclosures, I'm not a compensated speaker here today. I am not a consultant or advisor to Alkermes. I am currently investigating studies with Allergan, Alkermes, Sunovion, and Tonix Pharmaceuticals. And in full disclosure, I was notified yesterday morning that I was permitted to show up here today, and Alkermes did help me with my travel expenses. Airlines didn't think this was urgent enough to give me a discount.

The FDA is tasked with two questions. One, is this compound safe? The data says yes. The rest of the discussion is speculation and what-ifs. Second question is, is this compound efficacious, and the data tells us yes, but because of issues of placebo response going on around the world, especially in areas of anxiety and depression, a
new study design was introduced, and that has
clouded that picture of whether this drug is
efficacious.

We have a unique new chemical entity here.
It addresses a societal need. It addresses
depression symptoms. If we look at patients with
depression, over 90 percent experience symptoms of
anxiety. This medication is anxiolytic. It's the
first thing that I see as a clinical investigator,
as I'm evaluating these patients.

Patients with depression, over 90 percent
have sleep disturbance. This medication is
sedating and it helps with sleep, again helping
individuals with that impact on quality of life.

Something that we saw in our clinical
investigation over the last few years in the open-
label treatment studies, when I worked in the
pharmaceutical industry many years ago, our
statisticians told us that 1-year studies typically
resulted in 50 percent of the patients dropping
out. They dropped out for various reasons; the
burden of participating in a clinical trial or even
that they had improved and decided that they didn't need to be in the study anymore, only to drop out and find out later that they needed to get back into that study, but were not permitted to do so.

When we were running the 208 study, we had over 80 percent patients in retention, and in addition, knowing that we have depression as a chronic illness with episodes of remission and then a return, a recurrence.

At our clinic in Atlanta, we have decided to focus on how to improve the methodology of medical research. One of the ways we do that is by not enticing people to participate for financial compensation. We actually give them free after-study care.

Typically, in a study, when there's improvement, after the study is over, we will see in augmentation trials that within a matter of weeks, the patient will, on monotherapy, lose their effect. Once they lose that effect, we need to try and figure out what augmentation strategy to use.

With the Alkermes compound, we did not see
that. We actually saw that some type of modulation was occurring in the system where, a month out, 2 months out, 6 months out, they still were improved on monotherapy, and the longest that we went with a patient was 9 months. And we decided to follow him just to see how long before there was a return of symptoms; 9 months, and at that point, he said, "I need something new to be added to my treatment."

Current treatments are atypical antipsychotics, and you've all heard about the side effect profiles with them. There are many off-label treatments out there and insurance does not cover them. So patients are relegated to paying out of pocket, which is a burden to them, and unacceptable. We don't do that in any other therapeutic area.

I think that the information from all the speakers that went before me, those that have personally experienced the symptoms, is compelling enough to say, "Let's really consider this medication with a vote of yes." I believe that the
advocacy groups have also done the same thing. They've told you that there's a societal need. And I believe that scientifically and clinically, and from the experience of an investigator that has seen these patients and treats these patients that we have, again, there should be a yes vote on this compound. Thank you.

DR. NARENDRAN: Thank you. Will speaker number 13 step up to the podium please? Please state your name and organization for the record.

DR. FOX-RAWLINGS: Thank you for the chance to speak today on behalf of the National Center for Health Research. I am Dr. Stephanie Fox-Rawlings. Our center analyzes scientific and medical data to provide objective health information to patients, health professionals, and policymakers. We do not accept funding from drug and medical device companies, so I have no conflicts of interest.

As we've heard today, depression is a serious health issue. However, FDA requires proof that new drugs are safe and effective. New drugs have to work in order to actually help patients.
I want to focus on efficacy first. The evidence presented for BUP/SAM does not provide adequate evidence that it reduces depression more than placebo. Two of the three trials designed to provide evidence of efficacy did not find any statistical benefit of the drug compared to placebo.

The third trial, study 207, had a statistically significant reduction in depression on the MADRS for the 2/2 dose, however, the trial had shortcomings. For example, it used the MADRS-6, which lacks important aspects of depression and therefore can't prove efficacy.

Using the full MADRS, the only time that the drug was more effective than placebo was for just a few weeks in the middle of the short trial. By the end of the trial, the placebo group was doing just as well as the treatment group. This does not demonstrate a meaningful benefit for patients, and though statistically significant for that short time, the treated patients improved less than 2 points more than placebo on a 60-point scale. This
small difference seems too small to be clinically meaningful for patients.

This test has clear standards for improvement. Responders are defined as those whose symptoms improve by at least 50 percent, and remission is defined by scores that are less than 10 or below.

Patients taking the drug were not more likely to be a responder or to go into remission than the patients taking placebo for any of these efficacy trials. This again raises questions about whether the result in study 207 was clinically meaningful compared to other studies of antidepressants. These trials may have just been too short to demonstrate this, but there needs to be some sort of confirmatory evidence for this result.

Study 202 was designed as a proof-of-concept study, and FDA points out that it can't prove efficacy because it lacks the statistical controls to make sure that the difference did not occur by chance. Also, the relatively high drop-out rate
for patients in the drug arms could have had a large effect on the results.

We agree with the FDA reviewers that the lack of a dose response also raises red flags. If the drug is effective, the higher dose should at least show a statistical trend towards significance, but it didn't.

As we know, in studies of depression, the placebo groups often do quite well. Placebo controls are essential because they help control for the natural ebb and flow of depression episodes. The sponsor tries to eliminate evidence of a placebo effect. However, without a long-term comparison of the placebo drug to drug arms, it is not possible to determine whether natural fluctuations in depression or treatment are affecting the results.

There are also concerns about safety. The clinical trials included very few older patients, and older patients metabolize drugs more slowly and are more likely to have adverse reactions. They're more likely to be taking other drugs that can
interact with this drug.

Like all opioids, even ones that are designed to be abuse deterrent, this drug has the potential for misuse and abuse. This is a major concern because depressed patients are more likely to have substance abuse disorders and are at increased risk for opioid overdose.

In summary, the clinical trial does not provide adequate evidence that this drug reduces symptoms of depression. There are concerns about the potential for long-term harms to patients and others who might use or abuse it.

This drug needs to provide strong evidence of efficacy before approval. Although refractory depression is a serious condition, prescribing new treatment with unproven benefits and unknown risks is dangerous. As you know, new drugs for depression tend to be more widely prescribed than the narrower indications that FDA approves. Approving the drug for even treatment-resistant depression could easily contribute to the opioid epidemic. Thank you.
Clarifying Questions (continued)

DR. NARENDRAN: Thank you.

The open public hearing portion of this meeting is now concluded and we will no longer take any comments from the audience. Before we go to the next section, I do want to provide the sponsor an opportunity to answer the question that was raised in the morning about the single subject.

DR. VON MOLTKE: Thank you. We wanted to address the patient from study 202. This is the patient that there was a request from FDA to exclude. This patient is a 56-year-old African-American male who had a very strong response to the medication, and we've investigated, and there's no reason at all to exclude this gentleman's response.

I'm going to ask Dr. Pathak to very quickly and succinctly just list to you the pieces of evidence that we made copies of and sent back to the agency. So they have these.

DR. PATHAK: Sanjeev Pathak, psychiatrist at Alkermes. We went to the site, and we confirmed that the patient was eligible, had long-standing
depression, had cycled through many therapies, and
the fluoxetine was for an adequate dose and
adequate duration. Slide up, please.

We provided copies of 5 original documents
substantiating that the patient was indeed the kind
of patient who would be a candidate for a therapy
like this. There was a clinical note signed by the
physician listing the prior treatment as well as
the initiation of fluoxetine.

There was also the ATRQ done at the site.
There was the independent safer interview with an
additional note. There was a concomitant
medication log. So in totality, the patient was
appropriate for this study.

DR. VON MOLTKE: May I also ask if we can
clarify a comment made on the COWS scores that were
missing?

DR. NARENDREN: Sure.

DR. VON MOLTKE: Thank you. I'll ask
Dr. Stanford to comment. On a number of the
trials, there was no requirement that the drug be
stopped if they were going to roll over into the
208 long-term study. So in that case, there was no
COWS because patients didn't stop.

So maybe you can just address that.

DR. STANFORD: Yes. So for the 202 study,
patients completed participation and were done.
However, for the 205, 206, and 207 studies,
patients either had a follow-up visit or were
allowed to roll over into the long-term safety
study. For 205, they were required to complete the
follow-up visit, and that's where the majority of
COWS data came from.

For 206 and 207, there was no reason to stop
medication if they were going to roll over directly
into the long-term study. So there was no missing
data. This was all per protocol. And I just
wanted to clarify that 266 patients were in the
placebo-controlled dataset and 839 in the long-term
safety study.

DR. NARENDRAN: Thank you. I don't know if
the agency has a comment on COWS now.

I think we'll move to the next section. Dr.
Mathis will provide us with a charge to the
committee and present the questions for discussion.

**Charge to Committee - Mitchell Mathis**

DR. MATHIS: I know I took too much time this morning, so I'll give it back to the group this afternoon. The charge to the committee is pretty simple. It's to think about these data, think about the questions we have to answer from a regulatory point of view if you can do that.

The next set of questions we have to answer -- and we need your help with this -- are have substantial evidence as we defined it earlier of efficacy been provided for this drug product? If it has, then, or even if it hasn't, you can consider the next question about has the safety been adequately characterized?

Then we will ask you, once those questions have been answered, to do what we do with these type of data, decide the synthesis of that, do the benefits outweigh the risks. And I know that's a lot to ask, and I'd like to spend the rest of the afternoon, the time that we have this afternoon, to answer your questions so that you can get to a
Questions to the Committee and Discussion

DR. NARENDran: We will now proceed with the questions to the committee and the panel discussions. I would like to remind public observers that while this meeting is open for public observation, public attendees may not participate except at the specific request of the panel.

Question number 1, has substantial evidence been presented by the applicant to support the effectiveness of buprenorphine/samidorphan for the adjunctive treatment of major depressive disorder? This is a voting question.

DR. Mathis: We have some time built into the schedule -- this is Mitch Mathis -- to discuss this, and I'm sure there's some leftover issues that you had had.

DR. NARENDran: So if anybody has any questions, this is the opportunity to raise.

Dr. de Wit?
DR. DE WIT: I have a leftover issue, and that concerns the trial 206, that we all accepted was negative. It had quite a large number of subjects, I think 280, if I'm not mistaken. And in some of our materials, it was said that it might have been negative because of a large placebo response, but judging by the figures that are presented here, describing the two graphs for the placebo versus drug, it doesn't look like the placebo response is much greater than it was in the other three trials.

So I wonder if someone could comment on the negative trial and also if the FDA could say to what extent can you -- what's your level of accepting 2 out of 4 trials or 3 out of 4? How do you judge what's acceptable in this kind of case?

DR. FARCHIONE: I think that this is a good opportunity to talk a little bit about -- to just acknowledge the idea that the rate of placebo response in antidepressant trials is a problem. I think we heard that a lot from the public comments period as well. And it's increasing over time. So
that means that it's only getting more difficult to see a difference between drug and placebo in clinical trials.

So in the best-case scenario, we would have better drugs with bigger treatment effects, and then it'd be easier to see the difference between drug and placebo, but that has proven difficult to say the least. So it's not really unreasonable to look at trial designs as a way to manage the impact of placebo on the trials.

So earlier, somebody had asked about looking at just stage 2 in study 202, and you guys, the applicant, noted that just looking at stage 2 alone would have been positive, and that's great. So because none of the patients treated with the buprenorphine/samidorphan combination in stage 1 or that study had continued on drug in stage 2, it's an enriched population. It's essentially similar to a placebo lead-in, which is similar to study 206.

So even if we were to say look at that now, and say, yes, so stage 2 of that study looks like
maybe you've got a positive response, it's still fraught with all those issues of the post hoc comparisons and whatnot.

So we would still want some substantiation of that result, which then we would look at study 206, and study 206 was negative, so you don't get that substantiation there.

I think it's also a good chance for me to mention that with regard to SPCD, conceptually the design is appealing. From a clinical perspective, it makes sense to me that you would be able to try to mitigate the placebo response in your trials with a design like that.

The stats are admittedly very complicated, but if we could agree on what the analyses should look like and everything, it's not unreasonable to say that maybe this could help with this problem that we have that's very troublesome.

But with SPCD, you'd expect the drug placebo difference to be amplified in stage 2. So it should be easier to achieve statistical significance at the end of treatment, which didn't
happen here, so that pretty much leaves one of two possible conclusions. Either SPCD doesn't work or the drug doesn't work.

We don't know in this case what the right answer there is because we don't have just your basic standard, like drug versus placebo from the get-go, design, whatever.

With regard to the number of negative trials, that actually has come up in front of the committee before. The consensus among the committee at that particular meeting was, well, you've got two positive trials. The negative trials were not as impressive. You had the two positive trials. It's hard to get two positive trials, and so they said you've got two; that should count.

DR. NARENDRAN: Next question, Dr. Hernandez-Diaz?

DR. TEMPLE: Can I just add something? As people have told you, something like 50 percent of the apparently well-designed trials of antidepressants fail. So if you do four trials,
it's not unusual to have two failures. The question here is whether they have positive trials. That's the question.

DR. UNGER: This is Ellis Unger. Just in the interests of being completely transparent, in study 202, let's assume that the patient from site 124, there's no question about their qualification for the study. Nevertheless, we get very concerned when a study result is basically hinging on a single patient.

Now, you don't have to be a statistician to understand that if your p-value is 0.049 and you remove a patient with a good response, that you're now over 0.05, and you've taken a study from a positive study to a negative study. We understand that.

But here, the result was so extreme, we find it unsettling. And I'm saying that to be transparent. I want the committee to think about it. You can convince us that we shouldn't be concerned about it, but we are concerned about it. And I want to let you know that we're concerned
about it and give you the opportunity to disagree if you wish. But it does make us very concerned.

DR. VON MOLTKE: May the sponsor add anything?

DR. NARENDRA: Sure.

DR. VON MOLTKE: Dr. Anderson?

DR. ANDERSON: Aparna Anderson. I'm a biostatistics consultant, and I've been compensated for my time and travel. So if you take that question of influential data points in the analysis and you do a sensitivity analysis to probe that, the methodologically sound approach to take is to look at it in a symmetric way, not to just take the best data point from the xml group and see what happens.

If you want to look at that, you have to treat it symmetrically because you're affecting the distribution that way. And so that's why the sponsor conducted an analysis that trimmed both sides of the distribution and replaced it.

So it's not just trimming because you lose variability just by virtue of reducing your sample
size. You also have to replace those data and do it in a rational way.

The sponsor did it in two ways. They replaced it with respect to the adjacent data point. They also did a multiple imputational analysis where they drew on the overall data observed from the rest of the trial.

I think that's the methodologically sound way to do it. Otherwise, it's a very, very one-sided, lop-sided way to do it.

DR. NARENDRAN: Thank you. I think we get the disagreement there.

Move to the next question. Dr. Hernandez-Diaz?

DR. HERNANDEZ-DIAZ: Sonia Hernandez-Diaz from Harvard Chan School of Public Health. My question is to understand the efficacy to answer that question. And through the discussion, very interesting points from the participants.

I think we are getting the idea that depression is complex, that maybe there are many things going on, that maybe the treatments can
target specific things that can work for some
patients and not from others.

We also got an answer from this morning that
we are looking at our average. So I guess there
might be patients in the trials that might be
responding very well and others that might not be
responding.

So I wonder if, through the trials, it would
be possible to look at the data in a way that we
can identify where there are some patients that are
really dropping 20 points in the scale and others
are not, or is everybody dropping 1 or 2 points,
because that would I think point to the possibility
that maybe there may be a way to identify patients
that might benefit from this adjuvant therapy a
lot. And that would be, I think, a good outcome
from the results.

So I don't know if you have looked at the
data in a way to identify whether there is
potentially a group. I understand it's not the
primary endpoint.

DR. VON MOLTKE: That's a great question,
and we have looked at it that way. Dr. Schindler
will go through that with you.

DR. NARENDRA: The agency wants to comment
back on that?

DR. OGBAGABER: Can you go to slide 2 of
back-up slides? We just want to show the spaghetti
plot of all the patients in study treatment group
2/2 for study 202. And this is the subject we are
talking about, how it defers from the rest of the
patients. And clearly, it's an outlier, and that's
why we did the sensitivity analysis.

There is no one good robust sensitivity
analysis method, but what the sponsor showed us, we
don't disagree with. It's not uncommon for us to
look at extreme results or extreme subjects and
re-do the analysis.

DR. VON MOLTKE: We would just say he's a
great responder, but anyway, here's Dr. Schindler.

DR. SCHINDLER: Yes. I can show another
view of the data from study 202, if we have slide
up. This is the individual patient data for
stage 1 and stage 2, split by BUP/SAM and placebo.
Green is BUP/SAM. Gray is placebo. And you see in stage 1, the responder that we're talking about is that individual; each line is an individual patient.

So that patient at the far right-hand side of the top left is the patient that was removed. So it was the largest responder, the best responder. And what we've heard over the course of the day is that some patients respond more and some patients respond less.

You can see, when you look at the placebo groups, that there's some placebo responders, too. So the patient was the best responder, and you know, in any analysis, when you take out the best responder, you're going to have an effect on the analysis.

You see that there is intermediate response over the course of stage 1. When you look at stage 2, you also see a very different picture that there's a response in stage 2, but these are the placebo patients who have rolled over onto stage 2. So this shows the individual patient responses.
And you can see that that big responder is not really such an outlier, but just the best responder in the dataset.

DR. NARENDRARAN: I'll let the agency respond.

DR. UNGER: This is Dr. Unger. It almost looks like a cumulative distribution. We did a histogram. Unfortunately, it was too late to get in the slide set, but if you simply put patients in bins by their response, this patient is way off the chart. And basically, the spaghetti plot showed that he started out with the worst HAM-D in the study and ended up with the best HAM-D in the study. So his delta is huge compared to anyone else. This I think is a little bit misleading.

DR. NARENDRARAN: Thank you. We'll move to the next question. Dr. Kulldorff?

DR. KULLDORFF: Thank you. My name is Martin Kulldorff. I'm a biostatistician at Harvard Medical School. The key two questions is, is this drug really working or not? And then if it is working, what's the clinical amount or clinical effectiveness of it. How much does it work? I
wanted to try to synthesize these questions from a statistical point of view.

I'm not concerned about this one patient. I'm not concerned about SPCD. I think this has a very marginal effect on the big picture. But one thing that is concerning is the issue of multiple testing and not using necessarily the lack of clarity on what was the pre-defined outcomes to measure.

So the first study, 202, was an exploratory study, so I think the interesting thing is to look at the 205, 206, 207, and I do think that all of them are relevant for our evaluation, the effect sizes.

Two of them were not statistically significant. One was or was not depending on which outcome we use, so let's say we used the one where it was statistically significant. We then have point estimates of a benefit of 1.9, 1.7 -- sorry, 1.9, 1.8, and 0.2. So the question is, then that together, that amount, enough evidence to see that it's effective?
Now, I like meta-analysis, but I don't like at all the way it was done by the applicant here because, when they did the meta-analysis, they didn't use the prespecified outcomes. They used other outcomes. So they found an effect size in the meta-analysis was 2.0, 1.8, but those are actually higher than the average of these three studies that was not exploratory.

So if we take the average -- approximately based on the sample sizes and the point estimates. But if we take the 205, 206, and 207 together, their average effect size, using the average for the 207 rather than the 5-week, we get a combined of 1.3 effect size. If we also include the 202, the effect size will be about 1.5.

So my assessment is that the point estimates for the combined effect size is approximately 1.3 and definitely less than 1.5 rather than in the 2 range.

So then the second question is, is this an important clinical effect size that is important clinically to be able to provide the patient or is
it too small for that? And we heard that 1.5 to 2 was clinically important according to the applicant, but they had powered the study for 3.

So let's say 1.5. Is that clinically important effect size? Is it useful for a patient with a 1.5 reduction in this score? If the answer is no, then, well, that's pretty much what this study says, 1.5. So then there's no point in approving the drug because it doesn't have sufficient efficacy.

If the question is yes, that 1.5 is useful clinical -- as a biostatistician, I can't judge that. But if the answer is yes, then what should have been done was to do a study that's powered for 1.5 instead of 3.

It's a very common mistake done both in terms of companies, but certainly also in academia, that one does a preliminary study, kind of a small sample size, and one gets an effect size, and then one powers the subsequent true study with that observed effect size from the sort of trial study. And that was sort of what's done here because the
trial study had an effect size of 2.8 and then they powered it for 3.

Now, that's very, very dangerous to do that because, by chance, the true will be half the time more than 2.8 and half the time less than 2.8. So by doing that process of powering a study, you're almost guaranteeing a very large risk of having an inconclusive study.

In this case, it's even worse because there were actually two doses in the exploratory study, and one was 0.5 and one was 2.8 benefit. So then if you pick the biggest one, well, that's more likely to have been sort of biased towards the upper side.

So the reason why we have a very inconclusive study here, it's because, if 1.5 is truly important clinically, then the power should have been studied for that.

So if that is the case, then I think it would make sense for the applicant to do a study which actually has that type of power to be able to, with 80 percent or 90 percent power, detect the
clinically important effect size of 1.5.

That will answer the question if this drug is actually useful or not for the patients. We have heard that it's important to have new drugs, but obviously we don't want to give drugs to patients if it's not working. That's just wasting precious time and increasing suffering.

So if the applicant does have a study, then the question is what's the design. One of the issues that were discussed was weighted over 3 weeks or the end of time point. And I think there's a lot of good reasons for doing a weighted one because there's a lot of variability in the measurements because these are surveys.

So by doing the average of three measurements instead of taking only one, one reduces the variability, increases the power, and reduces sample size needed.

So I think that's a good call to do a weighted one. The only thing that wasn't quite clear here is it should be agreed upon by the applicant and the FDA ahead of time, that that will
be the main outcome to the study. So I would certainly encourage to do that weighted outcome.

With the SPCD, there are pros and cons, but I think there are a lot of good things with the SPCD. So I would not have any objections to doing that in a follow-up study as long as the power is sufficient to detect what's clinically important.

The third one is in terms of the outcome measure, it's sort of strange. At least in one of these SPCDs, the outcome was different for stage 1 and stage 2. I think there was week 5 in one and week 6 in the other, or some other thing like that. And to me, that doesn't make any sense. You might as well use the same one. That will be much more clear and that wouldn't have any drawback for using the same one.

So that's basically my sort of synthesis. It's unclear at this point whether this drug has efficacy. If it does, the point estimate is 1.5, and whether that is clinically important or not, I cannot judge, but if it is, I hope there will be a new randomized study to be conducted to find out if
we truly have such an effect from this drug. Thank you.

DR. NARENDRAN: Thank you. Those are good comments.

DR. VON MOLTKE: A comment from the sponsor regarding the clinical efficacy of the effects seen?

DR. NARENDRAN: You can comment very briefly, but I think it was more of a synthesis of his data, I mean, his read on the data.

DR. VON MOLTKE: I also just wanted to caution. I continue to hear the concern about the 8 milligrams of buprenorphine-containing being bigger than the 2-milligram dose in the exploratory. And I would really caution, with a complex pharmacology, of something like buprenorphine, there is a lot of literature out there that that is not the case to be expected. So I would just caution on making that inference around 8 being better than 2.

DR. NARENDRAN: I'd like to take the opportunity -- I mean, one of the things that you
do mention,. there is a lot of pharmacology, but we haven't seen a lot of pharmacology of what in vivo occupancy is.

You threw out nociceptor and you mentioned kappa opiate receptors, but we don't even know if this compound really has what level of occupancy. Looking at the in vitro occupancy, it's about 0.08 nanomolar at the mu opiate receptors, and it's about 1.5 for kappa in rodents, and it's about 4 for kappa in primates.

So that would suggest you have very modest occupancy at 2 milligrams. And if you were at 8 milligrams, I would think you would fully saturate and block mu opioid receptor occupancy, and you would have more kappa antagonism, but you don't see that.

So to me, the pharmacology of it is still not fully defined. You guys talked a lot about SPCD and placebo lead-in, but you're still in sort of '70s and '80s kind of understanding of the molecule. That also leads into some of the opioid withdrawal effects that the doctor from NIDA, a
person mentioned. And again, I feel there might be some metabolite accumulation and mu opiate saturation that's overwriting samidorphan's effects, and maybe that's what you're seeing.

So I think those would also be important studies to do to understand the pharmacology, and then power your study accordingly, is my thought.

[Inaudible - off mic].

DR. VON MOLTKE: Sorry about that.

One thing you are correct about is that the program really did have its genesis sort of backwards from the clinic, starting from an observation, and then again empirically adding samidorphan and bringing it down until we saw, with the visual analog scales, et cetera, that we were extinguishing drug liking and all those effects.

So from that standpoint, that is true. I can tell you, though, that back at the time when the 202 study was done, even at that time, there would not have been the presumption that the 8 dose would have been better than the 2/2 dose, just given the history that we had, and that we were
also still in the process of adding that antagonist and trying to figure out what ratio we were going to end up with to take forward. And we did end up with our anchor dose out of that study.

So I would agree with you it's not as nice and clean as maybe having had the PET imaging and all of those things, but it was empirically determined in a way that made sense and got us a dose that worked in phase 3.

DR. NARENDRAN: The Jordan Carp [ph] data, I think which is done with just the buprenorphine, I think it was effective at 0.5 milligrams without samidorphan, which again highlights to me that a very minimal mu opioid receptor occupancy is what you're going in terms of therapeutic effect. And maybe your 2 and 2 is just allowing for that, so maybe 0.25 of buprenorphine in itself might be effective.

DR. VON MOLTKE: Right. There's certainly been discussion about that. I think the company really had as its goal something that was not pure buprenorphine. Really, the goal was to put out
something that had buprenorphine with something to mitigate it. And while I've heard some conversation about naloxone not fully preventing misuse in terms of parenteral use, certainly samidorphan is quite a bit different than naloxone, which is notoriously short lived, for one thing. Samidorphan is also more potent.

So we really wanted to have something that was not pure buprenorphine.

DR. NARENDRAN: Thank you. Next question, Dr. Cella?

DR. CELLA: Thank you. Mr. Chairman, can I first clarify something with Professor Kulldorff, and then make a comment and question?

DR. NARENDRAN: Sure

DR. CELLA: I just want to get the term "effect size" because that's the core of my concern.

When you were using that term "effect size" I think you were referring to it like a difference in the score on the MADRS, as an effect?

DR. KULLDORFF: The difference in the score
on the MADRS between the placebo and the --

DR. CELLA: Right, the 1.53, 2.0. But the effect size that the sponsor showed overall averaged around 0.2 between groups, and that's an expression of the difference between groups over the variant, the standard deviation or some measure of variability; very small, maybe meaningful, but very small.

So that's what I'm struggling with, is that they had a range across the different studies of zero to around 0.6, but most of them in the 0.2 to 0.3 range. And we're being asked this question, is this substantial evidence?

I'm not as concerned with the one patient or with the design, and I was very moved by the public comment period, very impressed. The distinguished clinicians that spoke talked about responding patients.

If you put up the sponsor's slide E-177, which actually gets at what I'm struggling with, the last speaker, who does the systematic reviews, made the comment that there were no differences in
response rate between groups, between placebo and the combined therapy.

Dr. Jain asked a question earlier that I don't think actually was really answered. You asked about responders in placebo versus -- so I'd like to give the sponsor another opportunity to answer the question of the proportion of benefitting individual patients -- which is shown on this waterfall plot, and to me looks pretty similar. In fact, if anything, it kind of looks like the placebo patients in some ways might do better; at least, there's a lot of them that do quite well.

With an effect size of 0.2 across a mixture of negative and positive studies, I find myself really interested in knowing how many people as individuals are really getting a big benefit relative to those on placebo?

I'm impressed by the clinicians who spoke in the public session that it may be that this drug, in the right hands, with patient selection carefully done and monitoring done very well, that
you can see this benefit that doesn't exist if you just gave a placebo. But I don't see it in this waterfall plot, and I'd like to give the sponsors another chance to answer Dr. Jain's earlier question.

DR. PAPAKOSTAS: Thank you. So let me try again to address the excellent question of effect size. We're looking at numbers, and we want to know what does it look like in the real world.

The best way that explains it to me personally as a clinician, as a researcher, is to compare it with the precedent; to standardize it, compare it with a precedent. Slide up.

This is a standardized effect size for the positive and negative studies of BUP/SAM pooled, as well as the positive and negative studies of the most recently approved adjunct brexipiprazole pooled, and they map on to each other. So in the studies, BUP/SAM has the same efficacy as the most recent approval, which is brexipiprazole. And brexipiprazole, as we know -- I talked a lot about the limitations of the atypicals, but they are also
very effective augmentation strategies.

Ten years ago, when we did not have atypicals, we were in an even more difficult situation. We need to keep moving forward, replicating the efficacy of the atypicals, as done here, and trying to mitigate all the side effect profiles. So that's how I view it personally.

Thank you.

DR. VON MOLTKE: We'll have one of our statisticians talk to you about more outlying responses.

DR. MEMISOGLU: Hi. Asli Memisoglu, biostatistician, Alkermes. Can I have E-178?

Slide up.

This is a similar waterfall plot, but rather than MADRS change from baseline, this is showing percent change from baseline, which gives you a clearer view of subjects that meet the criteria of response.

This isn't a cumulative distribution. Each bar represents an individual patient. And the dotted vertical line that's towards the middle of
the graph shows the point at which patients meet
the criteria of response. And you can see that in
both stage 1 and stage 2, there's a larger
proportion of subjects who were given BUP/SAM 2/2
that met the criteria of response.

But there are a couple of other interesting
points I want to make on this presentation. If you
look at stage 2, those bars that occur above the
zero line are individuals that are getting worse,
and you can see that we have a large proportion of
patients that were given placebo that are getting
worse. You can also see that the depth of the
individuals who exhibit response, so they would be
in the lower right-hand corner on BUP/SAM, is
greater than those that were on placebo.

You can see by the difference of the
vertical lines that the greater proportion of
subjects on BUP/SAM 2/2 showed a good response
compared to those on placebo and provide support
for what we're observing with the primary efficacy
analysis. Thank you.

DR. CELLA: Could you just clarify? It's a
minor thing, but it says HAM-D 17 in the title, but MADRS-10 on the side. Is this HAM-D because it's 202?

DR. MEMISOGLU: It is HAM-D. I apologize for that.

DR. NARENDRAN: Thank you. Dr. Conley?

DR. CONLEY: Thanks. Two things, and partly about the fairness to industry comment here. I'm really going to talk off of Dr. Farchione and Dr. Ogbagaber's presentation, but just as a general thing. I'm stuck a little bit on the refuse-to-file on the reversal.

What I've heard so far today is the thorough presentation by the FDA about your concerns about the data, and I do hear that. But I'm struck on the sentence that the applicant clarified analyses, and then that was okay to submit.

I realize all of these things are hard whenever you wind up saying no, but I was also struck by everyone's comments and the efficiency of our whole process here. And I'm wondering if we've gone through the most efficient process, if you're
as concerned as you really are, should this have been accepted. Because what I saw from Dr. Ogbagaber's presentation is, actually, you disagreed on all the studies because, in the slide number 2, you disagreed on the two that they thought were positive. And we heard from the sponsor that they feel that there's positive data from 205, which you said you agreed was negative, so that's actually also a disagreement.

That seems like a lot of disagreement, and it feels like I'm just a little concerned about that process.

The second point, maybe to help this some, about SPCD. On Dr. Ogbagaber's presentation, slides 20 through 22, there's actually a lot of good thoughts about what might help and what your concerns are about this, but I wonder if it's time -- and I know it's hard to issue guidance, but if there's some method to come out, because we can't argue that this is new. I mean, this SPCD was published in, I think, 2003. So it's unacceptable to say 15 years is innovative stuff
anymore.

Should sponsors think that we need a special protocol assessment or something if you want this or what might be a good method? So those are my two things, a bit about the process, about acceptance, and the other about how to do SPCD.

DR. FARCHIONE: So part of the issue with the process, like I mentioned, the back and forth during the development program, there was a lot of very tight overlap between getting things and giving comments and whatnot.

At the end of the day, we got the NDA application, when we sent the letter back, the way that we described the analyses for 207 was actually wrong. That was the factual error, so we had the teleconference. And there was the question of -- we had said that we disagreed with the MADRS-6 and we disagreed with the averaging strategy, and yet, there it is; it's still in the plan.

But in the strictest sense, they had their prespecified primary analysis because it was
prespecified prior to unblinding, so then technically it becomes a review issue at that point, which is why we ultimately ended up filing it.

As far as SPCD, one of the conundrums that we keep beating our heads against the wall at this point, a lot of the proposals for analyses are based on simulations. And we had this big meeting, 2, 3 years ago or something, where we talked about SPCD in just a dedicated forum just for that, and tried to encourage sponsors, like just send us your data. If you've done an SPCD trial, send us your data so that way, we can start to try to look at those assumptions and the simulation to see if the assumptions in the simulations line up so that we can actually settle on what the best analyses are.

Honestly, this is really the first real-world set of data that we've seen. So that is why a lot of it is still unsettled and up in the air. We just don't have the information in order to answer to the questions.

DR. CONLEY: Thanks for both.
DR. OGBAGABER: Semhar here. So with regards to SPCD design, we've always kept an open mind about it, and also this is the only applicant that sent this SPCD design to DPP, and DPP has never received a SPCD design before. We are open to learning and we are actually encouraging applicants to submit more SPCD designs so we learn more and collect data.

So with regards to some of the simulation-based analysis that's been done with regards to weights, weights have been selected from simulations basically, and we can learn more about what weights to pick for stage 1 and stage 2 if we had real data from trials.

DR. CONLEY: One thing just in a quick response here -- I know with timing to be very quick -- I hear you, but it is important that if sponsors are actually in the process of trying to get something approved and with a clock ticking and timing, it would be challenging to send in data just at risk to let you collect more data to figure out where your heads are at.
So that's why I'm saying, I don't know if guidance can answer it. I realize this is not an easy thing to answer, but I'm putting it up right now. It is a real challenge to give back to the sponsor, to just give us more data when your actual program's at risk when you give that data out.

DR. HARRIS: Could I just make two quick points about this? The first was, at our end of phase 2 meeting, we asked about the acceptability of SPCD for our trial designs, and we were told that from a clinical perspective, they were acceptable. We also submitted the 202 data multiple years ago, so they've been at FDA for several years.

DR. NARENDRAN: Dr. Iyengar, do you want to comment, being a statistician?

DR. IYENGAR: Yes. This is in part in response to your question about SPCD. When I first saw the two treatment-resistant depression and SPCD, I thought something didn't quite match because you're interested in SPCD when there's a large placebo response, and I thought, how could
you have a large placebo response with something that's resistant? What does resistance mean in that case?

The other thing concerning the weighs, that bothered me because the weights were always pitched as pre-determined, determined a priori, and the 0.6 and 0.4 are very common. So I did a little bit of Google searching, and it's only last year, 2017, that a PhD thesis this was written at UNC, which dealt with adaptive determination of the weights, based on the data using a particular model.

So although the procedure is very old -- it's about 15 years old now -- there's not a whole lot that's known about it in terms of its properties, not only theoretically, but also how it works in practice. It's not like linear regression, where we have hundreds of years of experience. It's quite limited.

DR. NARENDRA: Dr. Dunn?

DR. DUNN: Walter Dunn, psychiatry, UCLA. Some questions and comments on two issues regarding the safety and abuse potential for the product and
then also the proposed risk mitigation strategy.

So I agree with Raj about the pharmacology in terms of our understanding of buprenorphine and samidorphan and how they relate to abuse potential and withdrawal. It's not entirely clear based off of what I've heard today.

So I'm wondering if we can maybe look at one of the other studies. It was mentioned that 208 was not designed to look at efficacy, but it was mentioned that it had an usually high retention rate after a year.

So I'm wondering from the FDA's perspective, in terms of all the studies that you have, for these kind of one-year type studies, what is the average retention rate, and is this an outlier. Because I'm wondering -- there are a couple possibilities.

One, the patient decided to stay in because there are minimal side effects from the medication, which is great. Alternatively, there could be some type of mild euphoria that's not being picked up by the visual analog scale, or there is some slight
withdrawal effect when they come off the medication for a couple days, that we know encourages them to remain in the study so they can get the drug.

I'm asking this from the perspective of a clinician, because in the real world, patients miss doses. Patients will double or triple up on doses depending on how they feel, and all these questions are not addressed in these studies here. These patients are pretty much taking it every day, and we're looking at withdrawal after a certain amount of time.

So I'm wondering, again -- and again, this is not a question we can answer, but is there something unusual about 208 in terms of a number of patients that were retained?

So that really speaks to the question of how safe is this drug, are we really treating depression, or are we putting patients on this mild dose of a mu opioid agonist, which actually I believe is probably where most of the, quote/unquote, "antidepressant activity is coming from."
Are we really treating the depression or are we just providing them a mild amount of euphoria? Depending on how you define antidepressant effect, is getting them better than their pre-treatment baseline?

In terms of the REMS mitigation strategy, I'm wondering if the question of safety is not entirely clear for us, if there's a more restrictive requirement after approval. And I'd be open to ideas of course, but what comes off the top of my head is kind of what's seen for clozapine at this point. Patients are getting a blood draw every month. They get a 30-day supply at most.

Is this something that the FDA is considering, where perhaps patients are getting U-toxes to make sure they are taking the medication and that's not being diverted. And if they test positive, they get another 30-day supply, but patients are not given anything more than a 30-day supply.

DR. NARENDRAN: We should let the agency comment.
DR. FARCHIONE: I think that we've sort of bled into a little bit voting question 2 and discussion -- question number 3. But just with regard to the question about is it an unusual rate of retention in the studies, I don't know off the top of my head what the typical rate of retention is in some of the long-term studies. I know we do see a little bit of attrition over time.

But the one thing to remember is that, by the time you get to the long-term, open-label study, those are usually all patients who have responded to the drug and who have just continued on just for the safety evaluation. So they're motivated to stay in because, if they've already responded, then they want to keep taking the drug.

DR. MATHIS: This is Mitch Mathis. That's right. I think for a drug that is a controlled substance with an NME that you can't get outside of a clinical trial, you can retain the patients easier than you can a supplement or something already approved, for instance.

So that would explain it until proven to be
something else, to me, just the opportunity to stay on a drug that has already worked for you and that you've already tolerated.

DR. VON MOLTKE: We do actually have some data compared to other studies if that would be helpful.

So Dr. Pathak, be quick.

DR. PATHAK: Sanjeev Pathak, psychiatrist, Alkermes. Their attention is consistent with other recent adjunct programs. Slide up, please.

You can see it for yourself over a period of a year, brexpiprazole, from the publication, 51 percent, and also BUP/SAM, 51 percent.

DR. NARENDRAN: Dr. Jain?

DR. JAIN: Felipe Jain, Harvard Medical School. I think a number of members of the committee are struggling with this question of whether the drug is effective, in part, because we don't have standards of what constitutes effectiveness aside from the historical controls that I think the company is appropriately presenting.
We've heard very eloquently from several patient groups, from NAMI, from DBSA, that they really want new treatments that will be effective for individuals. And yet, there's a disconnect between that and the analysis results that the company has submitted and that FDA has requested, I think based on that prior literature, in which we're looking at group averages in symptom reduction that, on their face, are not all that substantial.

Within trial 202, the HAM-D reduction in score was just under a 3, where moving that one outlier brings it down to just over a 2. Now, if I'm seeing a patient in my clinic and saying I'm going to give you this medication for your treatment-resistant depression that is going to improve your depression by a little over a 2, if that's about the median if you are analyzing the data in that way, that doesn't sound on its face to me like an effective strategy.

Now, if I had the numbers on response and remission across the trials and I could compare...
those between placebo and BUP/SAM, following on what Dr. Bodkin was saying about his results as an individual provider, that would help me to understand a little bit more about how to look at this in terms of its effectiveness in clinical practice.

So it seems to me that the public is asking for a different set of analyses than the sponsor is submitting, and the sponsor is submitting those on the basis of the prior literature, looking at treatment-resistant depression. And the effects of it, they're just really hard to grapple with a 1 to 2-point change on the MADRS on average. If I'm dealing with an individual patient, that's really hard to grapple with.

That's one of the reasons that I'm struggling so much with answering this question. I know it's an active area, but it would be really useful to know from your perspectives why response and remission data are not required in this kind of situation, and just your overall thoughts on individual level of effectiveness.
DR. TEMPLE: We are very interested in the distribution of effects. We're always interested. In the typical depression trials, however, a typical result would be, you start out at a HAM-D of 24, 25. The placebo group will change by 10 or 12 and the drug group will change by 14. That's the difference. But there are huge changes in both groups.

Whether if we looked -- and I'm not sure we have -- whether if we looked to define some kind of remission rate or something like that, you'd see a statistically significant difference between the groups. I'm not sure. I don't think we've done that, and we've relied on the mean change.

But there is obviously a distribution of results, and we're getting increasingly interested in all areas in showing what the distribution of results is.

DR. JAIN: And thank you for that clarification. Obviously, it's not a 1- to 2-point reduction in absolute change from baseline. It is relative to placebo.
DR. TEMPLE: It's the difference between the two. Right.

DR. JAIN: It's the difference.

DR. UNGER: Yes. I'll speak on behalf of the applicant here. This is Ellis Unger. If the mean change is 2 points, but actually a third of people respond and no one else responds, those third of people are getting a 6-point difference.

So we've been very interested in the distribution. And in the last, I don't know, 5 or 6 drugs we've approved for neurology indications, we've shown histograms that show the distribution of response, and we intend to keep doing that. I think we've done that for a couple psychiatry drugs also. It's very important.

DR. TEMPLE: We used to show cumulative distributions, but it turns out nobody can read those. So what we like to show is bar graphs divided into change of 5, change of 6, change of 7, and we are very interested in that.

DR. VON MOLTKE: So the sponsor does have some more data with regard to comparative effects
if that would be helpful, to just quickly see the slide. Dr. Pathak?

DR. PATHAK: Sanjeev Pathak. As regards to comparative effectiveness or difference versus placebo, the results are very consistent with other approved adjunct therapies. And, importantly, I think it's important to recognize that these are patients who have gone through several failures, have been suffering from symptoms for a substantial period of time.

Slide up, please. These are deltas on MADRS scales, MADRS end of treatment. At the top are the three studies from BUP/SAM, 202, 205, and 207. And you can see the point estimate and the confidence intervals. And at the bottom are the three adjunct antipsychotics and their studies. And this information comes from their labels. Brex is the most contemporaneous or most recently approved antidepressants.

What you can see is that the deltas are comparable and these are clinically meaningful differences. And Dr. Papakostas would like to add
something.

DR. PAPAKOSTAS: Yes. Just briefly let me add; I don't want to take too much time up, but having done one of the first studies of the atypicals, the most cited meta-analysis on the atypical -- I just want to share, having published a meta-analysis on atypicals in the American Journal of Psychiatry, highly cited, this pretty much maps up to the efficacy of the atypicals in terms of standardized effect sizes.

So you see here the distribution to the three studies, the two positive studies, as well as the remaining studies for the other agents. That's not where you see a difference. The difference in my mind is in the risks that patients do not have to take with this medication. Thank you.

DR. NARENDRAN: Agency comments? No.

DR. FARCHIONE: The only other thing I was going to say with regard to response and remission is, in fairness, that's not something we usually ask for as a primary, in part, because trying to settle on what's the best definition for that is a
little bit challenging in a clinical study.

Do you have a numerical cutoff for your
definition or do you define it by percent
improvement? If you could define it by percent
improvement, sometimes people who improve by
50 percent, if they were really sick at baseline,
they're still sick enough to get into another
trial.

So it's very sticky to try to figure out,
which is why we've just settled on the change from
baseline in a score. But we encourage looking at
that as secondary exploratory analyses, less than
secondary.

But I think part of what it feels like
you're asking about is, is there a way that you can
look at the patients who do respond and try to
figure out what's unique about them versus the
people who don't. That we would love to have for
any of the medications that we have approved, or
unapproved, or whatever. And it would be terrific
if there was some way that you could look at those
patients who did better and tell us what was
different about them that made them more likely to respond, but we just don't have that.

If we ever get that at some point, that's going to be a real revolution in psychiatry, but we're just not there at the moment.

DR. NARENDRAN: Dr. Riley?

DR. RILEY: So I don't have concerns about the outlier. I don't think we ought to exclude participants from clinical trials purely on statistical grounds alone. And I'm not really that concerned about the SPCD. I know we have debates that it's a relatively new approach, and some of the parameters, and how they're set, and those types of things. But I think, by and large, there's enough acceptance, enough studies of that, that I'm not too concerned about that.

I am still trying to wrap my head around the shift to the MADRS-6 as the primary endpoint. And the reason I'm particularly concerned about that is this question has to do with major depressive disorder, and the public comments were very clear about not only mood and affect being an important
component of this, but suicidal ideation, and sleep disturbance, and appetite disturbance, et cetera, et cetera, none of which are measured by the items that were excluded from the MADRS-10 to produce a MADRS-6. So basically, you're only measuring one facet.

I don't believe that a measure has to measure every single facet of a construct to be content valid or valid in general, so I think you can do it with a subset of items. I think that's a pretty reasonable assumption. But I think there's a need here -- and I don't know if the sponsor has it or if others have it, but a need to be able to show that the MADRS-6 does map reasonably well onto clinical indicators of major depressive disorder and other measures of clinical depression. And if we have that data, I would feel a little more comfortable with using the MADRS-6 as a primary endpoint. But I'm still trying to figure out why the shift to the MADRS-6 as a primary endpoint.

DR. VON MOLTFKE: May the sponsor comment on that?
DR. NARENDRAN: Briefly.

DR. VON MOLTKE: We did have interest based on the literature, but I will say that, once we heard FDA with regard to the fact that this really was not at a point yet where they were ready, we backed down to the second in the hierarchy, which was the MADRS-10. And it is on the MADRS-10 that the 207 is being asked to be regarded as positive. So we continue to be interested in MADRS-6, but for this application, it's MADRS-10 for the 207 study.

DR. NARENDRAN: Thank you. That's clear. Dr. Crawford?

DR. CRAWFORD: Thank you, Mr. chairman. My comment was very much like Dr. Riley's, but perhaps with a slightly different twist at the end. And I agree with everything you've stated as well as we heard the totality of everything presented from the applicant, sponsor, the agency, and the powerful statements from the public speakers.

But if we could show the question again, posed to the committee, I would like clarification from the agency, is this a binary yes or no.
question? Because as I'm looking at the DSM IV criteria, if one were to accept the design and the data presented with that caveat, then the evidence presented would meet at least the 5 minimum criteria for the diagnosis of MDD.

I struggle a bit more if it is truly a binary question in terms of if it were approved, would the labeling be limited to mild depressive episodes, or -- because the data with the MADRS-6 average, with all respect to the applicant, that is the basis for concluding efficacy, not the MADRS-10 average. But we don't see what we did here about the suicidal thoughts. We don't hear about a sleep disturbance as a concentration and the appetite.

So Mr. Chair, I guess I'm asking for clarification about how we should interpret this question.

DR. MATHIS: This is Mitch Mathis. So the way we usually do this is traditionally we have a couple of endpoints that will measure this thing we call major depressive disorder. And the Hamilton-D and the MADRS both will suffice, and there are some
newer ones that we've been evaluating.

The sponsor in terms of meeting the qualification of adequate and well-controlled for substantial evidence must identify that ahead of time and tell us what they expect to change before they enroll in randomized patients, so that when we get the results, we can compare the results with some competence, that we're measuring the same thing in both groups and any differences are related to drug. That's how it works.

As they pointed out, they picked MADRS-6. They did that faster than we could agree to it or not agree to it, but the second was a MADRS-10 average. And they went to the MADRS-10 average and they won. We hadn't agreed to that, but we've said that we think averaging might have some utility at some point. We were going to analyze that.

Then at MADRS end of treatment, which was the third endpoint, the usual endpoint, they did not make it. So for us, it is binary. You either have substantial evidence or you don't, and that's the decision that we're asking you for. And we've
never written a label where we said maybe the four
or five parts of the MADRS-10 changed, so we'll
label it for that and not for everything else.
It's major depressive disorder or not.

DR. NARENDRAN: I think we've had a lot of
discussion on question 1. I know there's a couple
people who have questions unless you have something
really burning that you wanted to make.
Dr. Griffin, real succinct, quick?

DR. GRIFFIN: Pardon?

DR. NARENDRAN: Succinct discussion.

DR. GRIFFIN: Sure. I just want to say
that, as a primary care physician, treating people
with these disorders, I understand the need for
more drugs or better drugs, but I also appreciate
FDA's charge with substantial evidence because I
think it's really important to get the information
before the drug gets out there because I think this
drug was tested in about a thousand people, but it
is going to be used in many tens or hundreds of
thousands of people, over many, many years.

So I think we want to be really, really
confident that this is an effective drug before it gets licensed and used. And I think it's a very interesting drug and we saw some very interesting data that's intriguing, but I think the question is do we have substantial evidence, and I think that's really important.

DR. NARENDRAN: Dr. Joniak-Grant?

MS. JONIAK-GRANT: Elizabeth Joniak-Grant. Just a quick comment, really. I'm concerned that the study population was about 75 percent whites and the remaining was largely blacks, leaving out Asians and Hispanics, Latinx, particularly because when you look at some of the results for the study 202 and 207, there are differences in effectiveness by race. And that just sort of disappears in the discussion, and I think that's something that's worth further investigation.

DR. NARENDRAN: Good point. Thank you.

I think we could vote on the question and make our comments later. I think we've had a lot of discussion on this particular -- Dr. Cella, go ahead.
DR. CELLA: I really apologize,
Mr. Chairman, but this question affects my vote.
This is a question to probably Dr. Mathis, or
Unger, or Temple. The sponsor is suggesting, by
what they've presented, that we should consider
recent approvals and the effect size in those
recent approvals in the neighborhood of what I saw
as 0.2 in these typical antipsychotics, as a basis
for saying that that's a reasonable effect size to
approve a drug on.

Do you agree with that or should this be an
individual decision? There's a precedent they put
on a slide. I'm wanting to know if you agree with
that as precedent.

DR. UNGER: We base approvals on whether or
not there's substantial evidence of effectiveness.
We don't get into what's a clinically meaningful
effect size.

DR. TEMPLE: I don't think that's true. We
have to believe that whatever the effect size was
matters. There's a court precedent for saying
that, Warner-Lambert v. Heckler. So we have to
believe it's in the ballpark for that. It is in
the ballpark for what previous approvals have been,
maybe a little less, but it's not way different.

Then, as I always like to point out, the
mean isn't the whole deal. There's a distribution
of results, and you've got to think about that,
too.

DR. FARCHIONE: But also remember that for
those trials, they did have two adequate and
well-controlled studies with positive results in
each of those studies individually, and they were
able to demonstrate it within the trial. It
doesn't have anything to do with comparison across
trials or between different drugs. It's whether or
not you have the adequate and well-controlled
positive studies for your drug.

DR. TEMPLE: Where they showed an effect.
It's not so much whether it was 2.2 or 2.3. They
showed an effect it was nominally significant in
each of two trials.

DR. FARCHIONE: Right. We've got to get to
the p-value before we can start talking about
whether the effect is clinically meaningful or not.

   DR. MATHIS: Does that answer your question?

   (Laughter.)

   DR. NARENDRAN: That was three different answers, but very helpful.

   (Laughter.)

   DR. NARENDRAN: I assume there's no further discussion on this question, and we will now begin the voting process. So please press the button on your microphone that corresponds to your vote. You will have approximately 20 seconds to vote. Please press the button firmly. After you have made your selection, the light may continue to flash.

   If you are unsure of your vote or you wish to change your vote, please press the corresponding button before the vote is closed, and that should change it.

   I'll read the question. Has substantial evidence been presented by the applicant to support the effectiveness of buprenorphine/samidorphan for the adjunctive treatment of major depressive disorder?
(Voting.)

MS. BHATT: Voting results,. 3 yes; no 20;
abstain zero; no voting zero.

DR. NARENDRAN: I think, now that the voting
is complete, we'll just go around the table and
have everyone who voted state their name, vote, and
if you want to, state a reason why you voted and
the reason you voted. We'll start from here.

DR. CELLA: David Cella. I voted yes, a
very reluctant yes based upon that interesting
answer to the question, but expected that will
travel to question 3.

DR. RILEY: A reluctant no from me.

DR. CRAWFORD: Stephanie Crawford, no for
the reasons stated, concerns about the MADRS-6
average.

DR. MARSHALL: Brandon Marshall. I voted no
because of the inconsistency of the outcomes across
studies 202 and 207.

DR. FLYNN: Kathryn Flynn. I voted no
because the definition of substantial evidence,
meaning at least two positive studies, and I didn't
see that convincingly presented.

DR. JENSEN: Roxanne Jensen. I voted no for some of the similar reasons mentioned, and I felt that using the average endpoint for 207 versus end of treatment did not meet the threshold.

DR. DE WIT: I voted no for the same reasons, that the magnitude of the effect didn't meet the standards of clinically significant to me.

DR. ACRI: I voted a very reluctant yes because it all kind of hinged on what the definition of substantial was. And if two positive trials is considered substantial, then I felt like they had that, but obviously the effect size was fairly small and there was some controversy about the methods and the analysis.

DR. MEISEL: Steve Meisel. I voted no, a number of reasons, most of which have been discussed here. We don't know the mechanism of action of this drug. We have no idea, just some postulates about this, and so we're just throwing something we think might have some good, but we don't know why even if it did.
We don't know if the impact that we are seeing, which is modest, is that of an antidepressant or of a narcotic euphoric effect. The impact is mild. And I'm no statistician, but when I took statistics back in college, I was told that you can put out a graph or a statistical model to prove anything you want.

The controversy we've heard today about the stats and the fact that the sponsor changed their definitions and changed their presentation all along the way, despite recommendations by the FDA along the way, to me is very concerning.

I think instead of starting off with a model and let's see what this drug does, yes or no, there's a likelihood here that we tried to retrofit a model to prove the point. And there's no evidence of that, but I think all of the controversy, and the discussion, and the history of the discussions with the FDA make me very concerned.

DR. GRIFFIN: This is Marie Griffin. I voted no for lack of substantial evidence. I also
didn't like including 202, which the FDA persistently said was not eligible and was a pilot study.

DR. BESCO: Kelly Besco. I also voted no and just want to also commend and say that I appreciate the members of the public that did share their personal experiences today and just want it known that I do also support the need for additional new innovations around treatment for resistant treatment depression.

But I did vote no because I think there is additional need for more confirmatory evidence.

MS. JONIAK-GRANT: Elizabeth Joniak-Grant. I voted no. I really was excited that there was some attempts to try and look at depression, treating it in an innovative way. It obviously is a need that needs to be filled.

Unfortunately, though, with this one today, I'm not convinced of the efficacy. I think there were just too many changes sort of all along the way with this study to what qualified as effectiveness. I also have concerns about the
homogeneity of the population, especially when whites were showing in certain things that they were showing better effectiveness.

MS. WITCZAK: Kim Witczak, and I voted no. And I think I voted no because there was so much controversy and things changing. And as a consumer, it's really hard because I know that people are desperate for treatments. But with that much controversy, and if the public was actually here to hear this debate, I think it would be really eye opening. And at the end of the day, we pay the ultimately price because we're the guinea pigs.

DR. FIEDOROWICZ: This is Jess Fiedorowicz. I very much appreciate the need for innovative and better treatments for this horrible disease. In reviewing the trials, I agreed with the FDA that 202 should be considered an exploratory study. I had concerns about risk of type 1 error in 207 with the weight change in prior outcome and when it was only significant with the late change and one that was not approved by the FDA.
Even looking at the sponsor-selected analyses, the effect size was very small, and I thought could even be overestimated because there is uncertainty about the integrity of the blind when you have a medication that did have some immediate effect that was discernible on drug liking, and the inclusion of patients with pain, which I didn't see measurement of.

So when you put that, at best, a small effect against the risks and these concerns for type 1 error, I voted no.

DR. NARENDRAN: Raj Narendran. I voted no. Dr. Meisel crystallized my comments, but the pharmacology to me was a question. I'm still concerned about the 8-milligram didn't show effect and the 2-milligram did.

202 in general raised questions. I was a little concerned about the single subject. It at least gives me pause that that subject drives the results so strongly.

I was also concerned about the late change for 207 of the outcome measure. And I just feel
like there's probably more that needs to be done.
A lot was learned on the fly and adjusted for 207
to get it to be a positive trial perhaps, but maybe
another trial would give me more confidence that
this is effective or substantial effectiveness has
been shown.

DR. JAIN: Felipe Jain. I'd like to
underline the reasons given by each of the members
of the committee who voted no, and I agree with all
of them, which I find myself surprised by.

As someone who has prescribed opiates to
patients with treatment-resistant depression and
feels that they may have some benefit within our
armamentarium, I still did not feel that the
evidence presented by the sponsor was convincing
of substantial effectiveness.

In addition to what has already been said,
I'd also like to underline the lack of transparency
regarding the individual subject-level data.

DR. DUNN: Walter Dunn, psychiatrist, UCLA.
I voted yes with a qualification. But in terms of
the yes vote, I was willing to accept the inclusion
of that patient from site 124, also impressed with
the fact that even with the removal of that
patient, the stage 2 data was still significant.

    In terms of the change for 207, I was
    convinced that the analysis using the average had
    clinical relevance and that it remained relevant
even if you looked at weeks 2 through 5, 3 through
7, 4 through 5.

    The qualification with the yes vote would
    have been the recommendation that a much more
    restrictive REMS be applied. I certainly
    appreciate the observation that the overall
    clinical effect size was minimal to moderate.
    However, as we all know, there are going to be some
    patients who are fantastic responders, and we need
to balance that out with potential risk for misuse
    or abuse for this particular medication.

    So it would have been with the
    recommendation that, yes, we approve. However,
    that clinical effectiveness was demonstrated, but
    with a much more restricted REMS to what I spoke
    before about perhaps limiting 30-day supplies,
coming in monthly for U-toxes.

DR. IYENGAR: Satish Iyengar. I also voted no, in part, because I didn't think there was substantial evidence. I also just saw too many methodological issues that swayed me this way.

DR. HERNANDEZ-DIAZ: Sonia Hernandez-Diaz. I voted no, and I didn't have a problem with type 1 error with exclusion or inclusion of the patient. My methodological problem was more with the cherry-picking of what studies to present, within each study, which dose you choose to present, or changes in the outcomes to present from the protocols during the trials.

Despite this cherry-picking, if the effect was huge, I would be voting yes, but I think, despite that, I don't see a substantial and clinically meaningful difference with placebo. We could wait one more week for the placebo in some of the trials, and we would be in the same room. So for that reason, I voted no.

But having said that, I think that at the same time, there is no evidence of lack of effect
of pursuing this potential effective therapy. So I hope that we keep doing research and try to find an effective therapy in this area of the opioids.

DR. HABEL: I'm Laurie Habel. I voted no for many of the reasons that other members did. I thought that three of the four studies did have results that were suggestive of efficacy, but they just weren't robust enough; 2 out of the 3 were only statistically significant with certain outcome measures and not with all the outcome measures. And so that inconsistency bothered me.

DR. KULLDORFF: I'm Martin Kulldorff. I voted no. I don't think there is evidence that this drug works, and I think it's unethical to give patients a drug if we don't know that it works. If 1.5 was the point estimate for the effect size comparing the drug with placebo, if that's the point estimate, if that's a clinically important effect size, then I hope that the applicant will conduct a study with appropriate sample size and appropriate power.

If there truly is an effect size of 1.5,
then you will find out that, and then we will know that it is both effective and that it is clinically important if 1.5 is that. If 1.5 is not a clinically important effect size, then I hope you will continue your innovative thinking about these drugs to find a different way to help these patients.

In either case, I would like to thank the applicant for their strong efforts of trying to find something that helps patients. Thank you.

DR. WARHOLAK: I'm Terri Warholak, and I voted no. I recognize that major depression is an issue. And from the public comments, I was very moved by the unmet need and that the public is asking for innovative solutions. And I commend the applicant for giving this a try, but I don't feel that the evidence presented at us was substantial in order to show efficacy. I think there were a lot of methodological issues, and I think Dr. Meisel's comments really captured what I wanted to say.

I'm very concerned about the last-minute
changes, especially with the things that were changed at the last minute being significantly different. And I'm not convinced that there's a big enough treatment effect to make a clinical difference, either. But I'd like to see efforts in this area to continue.

DR. RUHA: Michelle Ruha. I voted no. My no vote really hinged on the FDA's definition of substantial evidence. I think by looking at all of the evidence the sponsor presented that the drug probably does work, but maybe only in a specific population of patients that we haven't identified yet. However, that one outlier in study 202 actually bothered me, too, because I don't feel very comfortable with any study where you eliminate one subject and it changes the significance of the study. So I did not feel like it met the definition of substantial evidence.

DR. NARENDRA N: So just to summarize, it sounds like people had concerns on multiple levels in terms of whether a substantial effectiveness standard has been met. There was a question about
the single subject, questions about the in vivo pharmacology. Maybe there's opiate withdrawal effects or not clearly delineated, concerned about change in outcome measure, how things are done at the late stage, inclusion of people in pain, racial differences across outcome measures.

So with that being said, we'll move to the next question. Question number 2 is also a voting question. Has the applicant adequately characterized the safety profile of buprenorphine/samidorphan for adjunctive treatment of major depressive disorder?

So this is also a voting question. Please press the button on your microphone that corresponds to your vote. You will have 20 seconds to vote. Press the button firmly. After you have made the selection, the light may continue to flash. If you are unsure of your vote or want to change your vote, you can change it.

(Pause.)

DR. NARENDRA: One more person hasn't voted. Everybody, press the button again. If
everybody does it again, that person's vote should register.

(Voting.)

MS. BHATT: Voting results, 13 yes' 10 no; zero abstain; and no voting is zero.

DR. NARENDRAN: We'll go around the table and answer the questions. State your name, your vote, and the reason. We'll start from here, Dr. Ruha.

DR. RUHA: I voted yes. There really weren't any severe safety concerns, and I thought they did a good job showing even that it really wasn't a huge issue with withdrawal or euphoria.

DR. WARHOLAK: This is Terri Warholak and I voted no. I was concerned that we only had studies on abuse for the single dose. I didn't see anything on drug-drug interactions, and I would have loved to have seen some of the studies that the sponsor alluded to about what it took to separate the components for abuse prevention.

DR. KULLDORFF: Martin Kulldorff. I voted yes. One can never be completely sure about safety
before an approval of a drug, but I thought that applicants did a nice, thorough job, as well as is possible to do in a pre-approval process. So thank you.

DR. HABEL: I'm Laurie Habel. I voted no basically for the same reason as Terri did.

DR. HERNANDEZ-DIAZ: Sonia Hernandez-Diaz. I voted yes because I think they did characterize the safety profile. Of course, we want more evidence as it accumulates if it's approved. But they characterize it in the sense that they accept that there is some [indiscernible] opioid effects, like [indiscernible] constipation and so forth. But there is some data to support that there is no huge withdrawal within a large proportion of patients at least.

DR. IYENGAR: Satish Iyengar. I voted yes, in large part because I thought they did a good job of addressing head on the opioid crisis and sort of their approaches to that problem.

DR. DUNN: Walter Dunn. I voted yes based on the very defined time periods that were looked
at in the studies. I voted yes with the comment 
that I think things like long-term use in terms of 
withdrawal or abuse should also be looked at maybe 
perhaps with post-marketing studies and also for 
the potential for drug-drug interactions. We know 
that our patients aren't just taking this alone in 
the real world.

DR. JAIN: Felipe Jain. I voted no, not 
because I thought that this combination drug was 
unsafe relative to other opiates on the 
market -- buprenorphine's already on the 
market -- but because I did not feel that the 
sponsor had adequately addressed its safety in this 
population of major depressive disorder.

Specifically, thinking about some of the 
concerns raised by Dr. Sullivan in his presentation 
and the short-term nature of the placebo-controlled 
trials in clinical practice, treatment-resistant 
depression patients are not going to be treated for 
5 weeks, are not going to be treated for 6 weeks. 
They're going to be treated continuously. And I 
think that their safety does need to be assessed
relative to some control over a more extended period of time.

There was a paradoxical reasoning as well that I found myself trying to grapple with, which was that there was a clear signal for patients to immediately like the drug and for that to be present on a dose-response basis, higher with 8/8 than with 2/2. Yet, that didn't seem to be fully addressed within the analysis.

Some of the individual-level data regarding the clinical opiate withdrawal scale, again, were obscured by grouping people into categories of mild withdrawal symptoms, which span a huge range of symptomatology, moderate withdrawal symptoms.

So although I feel that there's some promise for this type of approach for treatment of treatment-resistant depression, I have more questions regarding its safety in this population than I have answers.

DR. NARENDRAN: I voted no because I'm still concerned that it does work through the mu opioid receptor. I think there is some euphoria effects.
There's probably some withdrawal associated with it. I'd also like to know more about that metabolite and how it accumulates in 7 days; what can that mean?

I also think it would be very difficult to screen patients for methadone or things like that. Buprenorphine's prescribed by other providers, not necessarily by the PCP or the psychiatrist. These people are going to get put on this med, so I think there's a lot of issues that could occur and create problems in the real world, so I voted no.

DR. FIEDOROWICZ: I'm Jess Fiedorowicz. I voted yes with similar caveats that were noted by Dr. Dunn about a postmarketing need for a longer-term controlled study. I was influenced by the FDA's statement that they said they didn't believe short-term trials were adequate and that the tradition is to do the short-term first and require other studies postmarketing. So that's why I voted yes.

MS. WITCZAK: Kim Witczak. I voted no. I voted no, I think, for a couple reasons. We don't
know the long term, so we don't know what we don't know. So that's one thing, ultimately. And then I heard the public speakers talking a lot about the suicide, and really that being a risk, and also thinking a lot about who the ideal patient is, J.D., and then listening to the presentation that we had by our guest outside speaker, talking about that depression or with the opioid use. And I feel like there isn't enough data to say that it's completely safe yet.

MS. JONIAK-GRANT: Elizabeth Joniak-Grant. I voted no for a couple of reasons. First, I couldn't help but wonder sort of what happens if an individual needs acute pain treatment, especially if they can't communicate that they're on this drug. It's like they go into an emergency room or something like that.

I would have liked to have seen the adverse events looked at by subpopulation because this is a population with a number of comorbidities and also looked at it in relationship to what other drugs they may be taking.
I was concerned that there were few patients over 65 and, in fact, in study 202, there was no one included that was over the age of 65. That was exclusion criteria. We see more comorbidities in older populations and more drug usage. That could cause problems. I was concerned that the population, again, was homogeneous with mostly white and mostly female participants.

Really, to echo what Terri said, I would really like to see the data about the ability to separate buprenorphine from the antagonist. I'd like to see the data. I think that would be really interesting because people are really good at doing that. They can do it in all kinds of ways we never really thought of, and I think some more information about how that could operate in the real world is super important.

DR. BESCO: Kelly Besco. I voted no as well for many of the same reasons, also worried about the separation aspect and not knowing what those solvents are, and the fact that they could be under anyone's kitchen sink, for all that we know.
Also, just another real-life practice note, as an inpatient pharmacist, I see more and more patients coming in on medication-assisted therapies for OUD. And they come in and they need treatment for acute pain. And I see physicians -- and hopefully Dr. Griffin won't give me a dirty look -- just throw opioids at these patients trying to counteract that antagonist effect. And sometimes, we get a little aggressive and we start stacking the opioids and have an unintended over-sedation event.

So I agree with Dr. Dunn, if this medication or a medication like it would move forward, we would need a multi-layered robust REMS program around it.

DR. GRIFFIN: Marie Griffin. I voted no for many of the same reasons. I think this is a really vulnerable population with high pain, high use of opioids or opioid use disorder. So I think even though they're not required to have more than 5- or 6-weeks trials, I think really, for opioids, we need to think about if they're going to be used.
long term, maybe we need longer-term trials.

DR. MEISEL: Steve Meisel. I voted no for much of the same reasons. Adequately characterized safety profile, I think the characterization here has been very, very narrow. First of all, the issue of samidorphan, I have never heard of a narcotic antagonist having an active metabolite being an agonist. That seems a little weird. But that has not been quantified or qualified in any way, shape, or form.

The fact that it's got some half-lives that are moderate, 4 to 5 to 7 hours; the situation that Dr. Besco described, a patient coming in with acute pain, they've broken their leg, and now they need narcotics, if they come in on one of these drugs that's used for opiate abuse syndrome, people know it. If they're on naltrexone, they know it. If they come in on one of these or like a drug like Contrave, which is used for weight loss, people don't realize there's a narcotic antagonist in there. So people aren't going to realize that they're trying to overcome the narcotic antagonist
within there, and that's going to increase the risk of some really serious problems.

I think the whole issue of abuse, again, it's a vulnerable population. We already heard that we can get a yield of 23 percent pretty simply. It won't be very long before we get a yield of 60, 70, 80 percent. And I can guarantee you, with the opioid crisis that we have in this country, this will make headlines for people dying from buprenorphine toxicity when that happens, if it's a question of when.

DR. ACRI: I'm Jane Acri. I voted no, partially because of reasons of basic pharmacology. We don't know what the mechanism of action for the antidepressant effect is. We don't know if it's a very low level of mu agonism, if it's kappa antagonism, if it's NOP agonism. We don't know, and I think that really needs to be delineated. I'm concerned about the active metabolite and its possible accumulation and the likelihood of causing physical dependence and withdrawal. I just don't think that's all been adequately characterized. I
think a lot more could have been done to answer those questions.

I also noticed -- and we never discussed this -- that both compounds are metabolized by the cytochrome P4503A4 and a lot of other drugs are metabolized by those CYPs. And the possibility of drug-drug interactions was never really discussed and I didn't see any data to address it.

Then finally, I think it's a very vulnerable population, a population that's highly at risk for drug dependence in the first place, so I think it's been a very risky proposition to introduce an opiate in that type of population if there are other alternatives.

DR. DE WIT: I voted yes. I thought, from the point of view of abuse liability testing, the applicant had done all the tests that we normally do for abuse liability, both in humans and in animals. And in particular, the test was done at the highest possible risk population, that is, opiate abusers, so they typically don't do the abuse liability in the target clinical population,
but rather in a population that's known to abuse that category of drugs, so they're looking at diversion there. And to my mind, I saw no signal for any abuse liability based on the data threshold were presented.

On the other hand, I did wonder a little bit about drug interactions. We didn't hear very much about using the drug in combination with something like alcohol, which would be very widely used by depressed people. And I also didn't hear very much about whether the formulation was tamper proof. But in general, I thought they did a good job with the risk assessment.

DR. JENSEN: Roxanne Jensen. I voted yes. I think, overall, the s applicant did a good job laying out the averse event, but I do think that some of the comments made before we got here are important, especially about what happens in real-world situations, where things aren't as controlled and monitored. Thank you.

DR. FLYNN: Kathryn Flynn. I voted yes, primarily for reasons that Dr. de Wit just stated
very eloquently and I won't repeat.

DR. MARSHALL: Brandon Marshall. I voted yes, primarily because I was convinced that samidorphan is doing what it was intended to do, that it reduces significantly the euphoric effects, and significantly reduces the risk of abuse potential, and that withdrawal was infrequent, and at present, mild, and could be medically managed.

That said, I'd like the caveat that I would like to see more data with polysubstance use, particularly where we see increased risk of overdose with buprenorphine, alcohol, and also prescribed a non-medical use of benzodiazepines as well.

DR. CRAWFORD: Stephanie Crawford. I voted yes. In terms of the question asked, has the applicant adequately characterized the safety profile in treating MDD, I thought they did. That stated, though, I also agree with Dr. de Wit's statements. Without repeating those, if the drug product were approved, however, we know that NDAs do not typically have long-term data, but
postmarketing studies would be looking at certain issues that have been stated by several around the table.

I was also struck by something in Dr. Sullivan's presentation, though it was only based on observational studies, retrospective cohort studies that longer-term greater than 90 days, opioid therapy at certain levels may increase the risk of depression. So it's a question mark, but still I thought the applicant adequately characterized the safety based on the indicated uses. I would have had more concerns about any REMS and potential misuse in that safety because the issue of the precipitated opioid withdrawal scared me.

DR. RILEY: Bill Riley at the NIH. I have to say these are tough calls, I think, on both efficacy and safety. I voted yes here. I understand the real-world concerns about issues regarding comorbidity, and drug interactions, and that sort of thing that I think need to be considered.
I was also struck after listening to the public comments and remembered that we already have a safety issue in the real world, which is people are prescribing buprenorphine for treatment-resistant depression now, without this antagonist for the mu opioid that I think is actually a really thoughtful way to think about how to reduce the safety risks of that use.

So I applaud the sponsors for making the attempt to do that work, and I think that's the way to go forward with this.

DR. CELLA: David Cella, Northwestern. I voted yes. I don't really have anything to add to the comments about the sponsor doing a presumably good job of short-term due diligence, and I was comforted by the obvious need and plan for a REMS program if this is approved.

DR. NARENDRAN: Just to summarize, I heard a wide range of opinions and the sponsor did a really good job in characterizing the compound in, again, the defined time periods, and it didn't have euphoria, it didn't have abuse liability, and
didn't cause withdrawal.

Then I also heard that people were concerned about the fact that they may have some [indiscernible] and mu opioid receptors and it has euphoria. It has withdrawal. I did hear consistently that people were concerned about the separation issue of who voted no. People were concerned about the external speakers' highlighting of depression, pain, being a high vulnerable population for substance use disorders, which could be a concern, and drug-to-drug interactions is obviously a concern with other opioids, sedatives, benzodiazepines. So that led people to vote no.

Hopefully that summarizes it, and we'll move to question number 3, which is a discussion question, but you've discussed everything possible that you can already, but we'll move to this question.

Related to the potential risks associated with the use, misuse, and abuse of buprenorphine/samidorphan in postmarketing settings, we'll start with A, discuss any concerns you have about the
risks of misuse, abuse, addiction, or overdose with buprenorphine/samidorphan in the intended patient population or in others who may have access to the drug.

Anybody who has comments; we'll start with Dr. Besco.

DR. BESCO: Kelly Besco. I know we've talked a lot about the patients themselves who will be taking these medications, but I'd be remiss if I didn't remind everyone that a large percentage of our opioid addictions begin with family members diverting their family members' unused medications or medications perhaps they're taking from their medicine cabinets.

I worry that if a patient would have an unsuccessful trial of this medication, it could remain in the home and certainly be accessed for potential misuse and abuse.

DR. NARENDRAN: Anybody else? Comments? Dr. Ruha?

DR. RUHA: Michelle Ruha. I just wanted to comment that, clinically, I treat people who
overdose on opioids every day, and I just want to remind everybody, I don't mean to completely downplay it. Buprenorphine is an important opioid also. But it is the safest opioid out there, and this dose is extremely low.

So this is 2 milligrams we're talking about, and I realize people could stack pills and have a big cumulative dose, but it's a very small dose of buprenorphine. And I admit, adults with opioid overdose every day and I've never admitted a buprenorphine because it does tend to have a ceiling effect with respiratory depression.

I have admitted children who have gotten into them, so I'm not saying it's risk free, but I don't want us to overestimate the danger of people overdosing on 2-milligram pills of buprenorphine and the risk of it either.

DR. NARENDRAN: Dr. Besco?

DR. BESCO: Kelly Besco. I definitely appreciate that comment and agree with you from a pharmacokinetic standpoint and looking at the kinetics of the medication. But I think we need to
remember, too, that access could be a gateway to something down the line that would be more potent.

DR. NARENDRAN: If there's no other questions, we'll move to -- Dr. Jain.

DR. JAIN: Sorry. No problem. So Felipe Jain. One concern that I have pertains to some of the studies that Dr. Sullivan referenced in post-traumatic stress disorder, which is highly comorbid with MDD, in which opiate treatment over the longer term can be associated with avoidance and can be seen to inhibit recovery.

That's partly where some of my concern about doing a longer-term controlled trial in the pre-marketing window stems from. I think that an opiate for MDD is different from an atypical for MDD. It's different than an SSRI for MDD because of the known abuse potential for opiates as well as this signal from PTSD.

DR. NARENDRAN: Any other comments?

(No response.)

DR. NARENDRAN: We'll move to the next subquestion. Discuss any concerns you have more
generally about approving a product containing buprenorphine and opioid for the first time for the treatment of depression. Dr. Joniax-Grant?

MS. JONIAK-GRANT: Elizabeth Joniax-Grant.

Not a huge thing that I'm like, oh gosh, this is the most horrible thing ever, but something that keeps kind of popping into my mind is that with so many of these patients also having pain, would getting an opioid over time increase their pain issues?

All the latest data suggests that taking opiates long term increases pain and so how would that play out with this population?

DR. NARENDRAN: Any other comments?

(No response.)

DR. NARENDRAN: We'll move on to the next question. Discuss whether you are concerned about the risk of opioid overdose if other opioids are used, misused, or abused concurrently with buprenorphine/samidorphan, perhaps in higher doses intended to overcome buprenorphine/samidorphan's mu antagonistic effects, with potential additive
opioid agonist effects that have not been fully
categorized. Dr. Meisel?

DR. MEISEL: Steve Meisel from Fairview. We
talked about this somewhat in one of the previous
questions. But there is no doubt that if you think
of a hospital setting, an ambulance setting or
something, and somebody's on this and they broke
their leg, or dislocated their hip, or whatever it
may be, and you got this narcotic antagonist on
board, that's going to lead to all sorts of pushing
of all sorts of doses, and then the narcotic
antagonist wears off before the narcotic does. And
then what?

Somebody gets into the operating room, and
they don't realize they're on whatever the brand
name of this thing is. They don't realize that
there's a narcotic antagonist that goes along with
that. And what's the anesthesiologist going to do
and how is that going to affect the care inside the
operating room?

When people are at home, I think the same
sorts of things happen, but in a much more subtle
way. So you go to the dentist, and now, you get your prescription for Vicodin or whatever it may be, but you're on this, and the dentist doesn't know anything about this stuff, and they just think you're on an antidepressant. Then that doesn't work and then what does the person do?

Again, there are all sorts of unintended consequences when you approve a drug like this. When Contrave was approved, nobody thought about this stuff, and we're now seeing all sorts of problems related to that because people don't realize there's naltrexone in Contrave, and nobody knows how to deal with it. There are no guidelines. Nobody's been able to dream up guidelines, and nobody's dreamed up or even thought through the guidelines with this, and it's going to create all sorts of unintended consequences.

DR. NARENDRA: Dr. de Wit?

DR. DE WIT: I'm not so worried about the overdoses with combination with buprenorphine and another opioid. It's a partial agonist, so it's going to be its own limiting factor. If you add
another agonist, then that's going to block its effects, plus the samidorphan is an antagonist, so it just doesn't seem like you're going to overcome the effects of this combination with another opiate.

DR. NARENDRAN: Anybody else? Dr. Meisel again?

DR. MEISEL: Just to respond to that, the real-life situation of that is that they got the narcotic antagonist in their body, and now they've got acute pain. Now, you're trying to treat that acute pain and you can't overcome it because of the narcotic antagonist. I think, to me, that's my real concern.

DR. NARENDRAN: Thanks for that clarification. We'll move to the next subcomponent. What risk reduction strategies could be implemented to decrease risk associated with buprenorphine/samidorphan use? Dr. Dunn?

DR. DUNN: I think, with every pharmacological device that we use in medicine, if we can treat the patient or provide benefit, we can
definitely do harm. And I think the question here is, what is the harm long-term.

Again, as I mentioned before, I think the company did a pretty good job in terms of short-term safety data, but in terms of long-term safety data, that's still kind of up in the air, and then hence, the postmarketing recommendations were for long-term follow-up.

I'm assuming that if this were to get approved, the label would be for acute treatment of a major depressive episode. But as my colleague, Dr. Jain, mentioned, for these patients, they're going to be on it longer for than 6 weeks. They're going to be on it for longer than 12 weeks, especially if it works. They're going to be on this thing potentially for a lifetime, and the safety aspect of that is a little bit unclear.

As I mentioned before, a more robust REMS strategy would, I think, placate some of those concerns, limiting to 30-day supplies, frequent U-toxes, or perhaps even before the postmarketing results came out, perhaps you could only limit it
to a 3-month or 6-month treatment, that patients
who were treated on this, if they don't respond in
that time frame, that's the one chance that you
got.

But if you responded, perhaps that's a
discussion that you and your clinician can have
about maintaining this long term. But I think that
clinicians need to also take responsibility in
terms of adequately educating the patient about the
risks and benefits.

I think the tricky part about this
medication is that the potential risks, if they do
arise, they're subtle. All right? They're
long-term risks. There's not something that the
patient is going to acutely notice. And I think
those types of risks are difficult to convey to the
patient because our brains are built to look at
kind of acute effects.

The long-term effects, long-term
consequences, they're harder to see. That's why we
have opiate epidemic. That's why we have
alcoholism, because those long-term effects are
harder for us to see.

So I think, in the beginning, a very robust restrictive REMS strategy, and some of the postmarketing results came out showing that the risks were not as great as we feared, perhaps easing those restrictions.

DR. NARENDRAN: Dr. Meisel?

DR. MEISEL: Steve Meisel. So I do want to point out I've been on a number of these committees relating to opioids, and I can tell you that almost none of the REMS strategies are effective for opioids because most of them rely upon education, flyers, those kinds of things. And that's all I heard today, is proposals for nice flyers, and education, and websites, and that sort of stuff. And we're deluding ourselves if we think that's going to reduce any risk. So that's one thing we should not do, replicate the failures with the opioids.

I like your idea about your tox screens. I think that might be helpful. Just like clozapine, what we do with that. I think that is helpful.
there. I think it would be helpful if before a drug like this were to be approved, if the manufacturer were forced to have a dosage form that would reduce the likelihood of diversion so that you couldn't separate the products from each other and inject it. I know there's a lot of work going on in that space in the opioid for pain world, some successful partially. A lot of it is unsuccessful, but that's a strategy to reduce the risk of diversion.

In the space of what we talked about earlier with patients who come in with acute pain needs and nobody realizes that they're on a narcotic antagonist that goes along with that, how can we leverage computer systems and other forcing tactics that put into people's face that this patient is on a narcotic antagonist. And here is a strategy and here are the tactics that you should be doing, should a patient need acute pain response.

So to spell those out ahead of time, work with our electronic health record vendors to put that front and center should a patient come in on
one of these products, those things can help mitigate some of the risks.

DR. NARENDRAN: Dr. Besco?

DR. BESCO: I was going to say mainly the same thing, but just not as eloquent as that. But no, I just will concur that I believe that the concurrent proposed REMS strategies are very passive and would like to see something that's more active, that follows the patient throughout the course of their therapy rather than just at the initial, is this an appropriate use case or not.

DR. NARENDRAN: I'd like to reiterate that, too. I think the REMS strategy is inadequate. It should be at least as restrictive as what we do for DATA 2000, I think. And more in line with that, education, keeping track of the providers, knowing how many scripts are being written, and maybe your end talks could be very beneficial, at least if you decide to go forward. Ms. Witczak?

MS. WITCZAK: Kim Witczak. Again, I know this is a totally different department at the FDA, but I think the communication, the outward
communication to the public is really important as part of this. It might be something different because I can hear the commercials right now. "Are you still depressed? Your antidepressant isn't working?"

You can see that being part of it, and I don't know if that's regulatory or if it needs congressional, but I think that piece of it needs to be more than just a pamphlet that's given out at the doctors. I think that's really something to be considered.

DR. NARENDRAN: Thank you. We'll move to question number 4, which is a voting question. Do the available data support a favorable benefit-risk profile of buprenorphine/samidorphan to support approval?

Same directions; press your button on the microphone that corresponds to your vote, 20 seconds to vote, you can always change as far as it's flashing.

(Pause.)

DR. NARENDRAN: We're missing one person.
Please press the button again, everyone. Press it hard so it registers.

(Voting.)

MS. BHATT: The voting results; yes 2; no 21; abstain zero; no voting zero.

DR. NARENDRAN: Do we want to go around? Very quickly, state your name, your vote, and why you voted. We'll start from this end of the table.

DR. CELLA: David Cella, Northwestern University. I voted for the first question yes, as I said, very reluctantly, very small effect, and with the safety data, and the concerns raised through the day, it just didn't put me over the top in terms of risk-benefit.

DR. RILEY: Bill Riley, NIH. I think if the safety were essentially nil, you could maybe take a flyer on the minimal or the edgy aspects of the benefit ratio, but they're more than nil, so no.

DR. CRAWFORD: I'm sorry. I was pressing the yes button instead of the mic. Stephanie Crawford. Having voted no in the first question about substantial evidence of the efficacy, I have
to vote no on the risk-benefit.

DR. MARSHALL: Brandon Marshall. I voted no. I just did not see evidence of benefit, maybe marginal benefit, and that just did not outweigh the safety concerns for me.

DR. FLYNN: Kathryn Flynn. I voted no. I think there's clearly a lot of evidence presented today that looks like, in certain situations, this could work, and there is very clearly unmet need, but overall, especially considering 202 as exploratory or proof of concept, I should say, I voted no on the ratio overall.

DR. JENSEN: Roxanne Jensen. I voted no for efficacy, so I felt that, overall, I couldn't vote yes for this.

DR. DE WIT: I also voted no because of the efficacy issues.

DR. ACRI: I voted no even though I voted yes for efficacy. The effect was small and, obviously, there were a lot of considerations about the magnitude of the benefit. And I felt that the risks were equal to the benefit at least, that the
benefits did not outweigh the risks.

  DR. MEISEL: Steve Meisel. I voted no for the reasons I've already described.

  DR. GRIFFIN: Marie Griffin. I voted no because of the risk-benefit.

  DR. BESCO: Kelly Besco. I also voted no for reasons that I've already shared.

  MS. JONIAK-GRANT: Elizabeth Joniak-Grant. I voted no. I'm not convinced of efficacy and I'm not wholly convinced of the safety.

  MS. WITCZAK: Kim Witczak. I voted no. I don't think we can approve something if it doesn't work. And I also think, because it has the opioid, we almost have to create an even higher standard for approvals.

  DR. FIEDOROWICZ: This is Jess Fiedorowicz. I voted no for reasons I've already described.

  DR. NARENDRAN: Raj Narendran. I voted no for reasons described already.

  DR. JAIN: Felipe Jain. I voted no for previously described reasons.

  DR. DUNN: Walter Dunn. I voted yes because
I voted yes on the first two questions, but also
the caveat that I would use this medication very
carefully and in probably a much more of a
restrictive population than what was looked at by
the sponsor, so perhaps for patients who have
failed for five different types of treatments and
maybe several other augmentation strategies. Even
then, I perhaps would only use it for a limited
amount of time.

DR. IYENGAR: Satish Iyengar. I also voted
no because of the efficacy issues. I had voted no
on the first question and also after hearing all
the concerns about the risks.

DR. HERNANDEZ-DIAZ: Sonia Hernandez-Diaz.
I voted no because the potential benefit doesn't
seem to outweigh the potential risks.

DR. HABEL: I'm Laurie Habel. I voted no.
I voted no on the first two questions, so that
seemed appropriate.

DR. KULLDORFF: I'm Martin Kulldorff, no.

DR. WARHOLAK: I'm Terri Warholak, and I
voted no for reasons already stated.
DR. RUHA: Michelle Ruha. I voted yes even though I voted no on the substantial-evidence question, and that was because I did think there was evidence. It was more the FDA's definition of substantial evidence, I felt I needed to vote no the first time.

But I did feel like there was evidence. I was actually surprised and moved a bit by the graph that compared the results to the aripiprazole and other atypical antipsychotics like quetiapine. It looked pretty similar in efficacy there, and I was moved by some of the efficacy results. And I voted yes on safety.

I would support much stricter REMS. I agree with that, possibly even the DATA ex-waiver like we talked about, although I know that's for opioid treatment. And I would restrict the patient population to really resistant depression.

DR. NARENDRAN: Okay. Just to summarize, I heard a lot of people voted no, mostly based on the efficacy. There was a few people who also had concerns about safety. The people who voted yes
also felt it was reasonable, but with a stronger REMS and more restrictive prescription status.

Question number 5 is a discussion question. What, if any, additional data are needed pre- or post-approval to address outstanding issues with buprenorphine samidorphan? Please be clear whether you believe these data should be required prior to approval. We'll start whoever's ready to go to provide the discussion. Dr. Crawford?

DR. CRAWFORD: Thank you. I don't have any additional suggestions about data, but I think I would like to underscore that several of us would feel better if it were a more delineated REMS versus what was given within the briefing documents and the applicant's description.

DR. NARENDRAN: Dr. Cella?

DR. CELLA: I would just like to suggest much more clarity about responders and percent of remitting patients to get to the bottom of is there really this subset of patients that has a tremendous benefit. I hate to see that opportunity lost in those patients, but it just didn't come
through in the presentation today, and it might be there based upon some of the public comment.

DR. NARENDRAN: Dr. Joniak-Grant?

MS. JONIAK-GRANT: Elizabeth Joniak-Grant.

In addition to some of the comments I've already made about data that I'd like to see, I think it would be good to see data about the super responders. There's been a lot of reference to them. One of the comments was about them. But there's no data to say, well, are there two of them? Are there 50 of them? And what do we mean by a super response? So I think that data would be really useful.

DR. NARENDRAN: Dr. Jain?

DR. JAIN: Yes. So I echo a call for more data on response and remission. I also think that it would be quite beneficial to do a trial in depressed patients with comorbid pain since it's the population in which prescribers are going to be really quite drawn to use this. And I don't think we can ignore that premarketing and not have any data on that premarketing.
It would also be helpful to have greater
data on an individual level withdrawal symptoms and
greater characterization of what that is, not only
from a statistical perspective, but from a
narrative perspective in patients who are
experiencing symptoms such as these, particularly
given the FDA's analyses on the percentages that
showed higher side effects within the BUP/SAM group
than the placebo.

DR. NARENDRAN: Dr. Dunn?

DR. DUNN: Given the observation that
perhaps some of the mechanism is potentially purely
at the mu opioid receptor, I would like to see what
the FDA had mentioned before, long-term efficacy
data on patients who are on this for perhaps about
a year, with a concern that perhaps they're going
to develop tolerance to it and then lose that
antidepressant effect.

Then also along that line, if you were able
to show that this was truly acting at the kappa
opioid receptor, so perhaps a combination of
buprenorphine plus naltrexone still retained the
antidepressant effect, I think that would contribute a lot to the safety concerns.

DR. NARENDRA: Dr. Hernandez-Diaz?

DR. HERNANDEZ-DIAZ: Sonia Hernandez-Diaz from the Harvard Chan School of Public Health. I was going to say the longer-term effectiveness, I think we have heard that comment many times. I was going to also say that now that the outcomes have been discussed and the dose is more clear, having another efficacy trial focusing on more patient-centered outcomes that are relevant for patients and clinically meaningful. And potentially, there might be studies to identify a group a priori and have a study that identifies that group, and have this stage 2 randomization rather than for non-placebo responders, for null responders being kind of dropped, and focus on a population. Like, you could replicate that study in clinical practice.

Also, in other comments, we have been trying to compare informally with other adjuvant therapies that in clinical practice, you will not give
placebo, or you will not stay in the same dose in these patients. You will increase the dose. You will add other psychotropics.

So I wonder if the comparison group is more appropriate, that has the equipoise, not placebo, but maybe considering other studies that you will follow in clinical practice with these patients, just a comment.

DR. NARENDRAN: Dr. Kulldorff?

DR. KULLDORFF: Martin Kulldorff, Harvard Medical School. If 1.5 reduction in the scale between drug and placebo is clinically important, then I would like to see a randomized trial that's powered to detect a difference of 1.5, and I think that should be done prior to approval.

DR. NARENDRAN: Thank you. I would like to add my comments in terms of I think it would be good to see a study with in vivo pharmacologies clearly defined at 2 and 2. I mean, there are really good mu opioid receptor radiotracers. They're really good kappa opioid receptor radiotracers that could provide that kind of in
vivo pharmacology data. I think one more trial of everything we have learned, perhaps SPCD, longer duration, averaging their outcome measure, and to prove substantial effectiveness would probably be very welcome, I think. That's my last comment.

Anybody else have any suggestions?

DR. ACRI: I'm sorry. You passed me by.

I'm Jane Acri from NIDA. And I guess the one thing that I would like to see is an actual study of whether or not you produce withdrawal with discontinuation because I feel like withdrawal signs are one of the reasons that people keep taking opioids.

So if I ran out of my BUP/SAM, would I just take another? Would I take an oxy? Would I take something to alleviate withdrawal and then get hooked on another opioid?

So the ability to create a physical withdrawal is really important in terms of maintaining use and possibly transition to other opioids. But I would also like to see more basic pharmacology, and as you were saying, some work to
determine the mechanism of action, to look at kappa
opioid receptors as well as mu opioid receptors and
try to get to the root of what's happening in the
long term.

DR. NARENDRA: I think, if there's no
further comments, I think we can close the session.
Thank you, everybody.

DR. MATHIS: Thank you for your service. We
appreciate it.

Adjournment

DR. NARENDRA: Thank you. We'll meet again
tomorrow.

(Whereupon, at 5:01 p.m., the meeting was
adjourned.)