FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

MEETING OF THE PSYCHOPHARMACOLOGIC DRUGS

ADVISORY COMMITTEE (PDAC)

Tuesday, December 1, 2015

8:01 a.m. to 4:21 p.m.

FDA White Oak Campus

Building 31, The Great Room

White Oak Conference Center

Silver Spring, Maryland

1	**Meeting Roster**
2	**DESIGNATED FEDERAL OFFICER (Non-Voting)**
3	**Kalyani Bhatt, BS, MS**
4	Division of Advisory Committee and Consultant Management
5	Office of Executive Programs
6	Center for Drug Evaluation and Research
7	
8	**PSYCHOPHARMACOLOGIC DRUGS ADVISORY COMMITTEE MEMBERS**
9	**(Voting)**
10	**David Pickar, MD**
11	Adjunct Professor of Psychiatry
12	Johns Hopkins Medical School and
13	Uniformed Services University of Health Sciences
14	Gabriel Sciences, LLC
15	Chevy Chase, Maryland
16	
17	**Murray B. Stein MD, MPH**
18	Distinguished Professor of Psychiatry
19	Family Medicine & Public Health
20	University of California San Diego
21	La Jolla, California
22	

TEMPORARY MEMBERS (VOTING)

Ralph B. D'Agostino, Sr., PhD

(Acting Chairperson)

Professor of Mathematics/Statistics

Biostatistics and Epidemiology

Mathematics and Statistics Department

Executive Director MA/PhD Program in Biostatistics

Director, Statistics and Consulting Unit

Boston University

Boston, Massachusetts

Dean Follmann, PhD

Assistant Director for Biostatistics

National Institute of Allergy and Infectious Diseases

National Institutes of Health (NIH)

Bethesda, Maryland

1 **Judith D. Goldberg, ScD**

2 Professor of Biostatistics

3 Departments of Population Health and Environmental

4 Medicine

5 New York University School of Medicine

6 New York, New York

7

8 Victor De Gruttola, ScD *(via phone)*

9 Professor of Biostatistics

10 Department of Biostatistics

11 Harvard School of Public Health

12 Boston, Massachusetts

13

14 **Nitin Gogtay, MD**

15 Director, Office of Clinical Research

16 National Institute of Mental Health, NIH

17 Bethesda, Maryland

18

19

20

21

22

Jennifer Higgins, PhD

(Acting Consumer Representative)

Director, Strategic Planning and Business Development

Center for Human Development

Springfield, Massachusetts

Dawn F. Ionescu, MD

Depression Clinical and Research Program

Massachusetts General Hospital

Harvard Medical School

Boston, Massachusetts

J. John Mann, MD

Paul Janssen Professor of Translational Neuroscience

Director, Molecular Imaging and Neuropathology Division

Department of Psychiatry

Columbia University/New York State Psychiatric Institute

New York, New York

1   **Rajesh Narendran, MD**

2   Associate Professor in Radiology and Psychiatry

3   University of Pittsburgh School of Medicine

4   Pittsburgh, Pennsylvania

5

6   **Matthew V. Rudorfer, MD**

7   Associate Director for Treatment Research

8   Division of Services and Intervention Research

9   National Institute of Mental Health, NIH

10   Bethesda, Maryland

11

12   **Natalie Compagni Portis, PsyD**

13   *(Patient Representative)*

14   Oakland, California

15

16   **ACTING INDUSTRY REPRESENTATIVE TO THE COMMITTEE (Non-**

17   **Voting) Robert Russell Conley, MD**

18   Global Development Leader, Pain and Core Therapeutics and

19   Distinguished Scholar

20   Eli Lilly and Company

21   Indianapolis, Indiana

22

FDA PARTICIPANTS (Non-Voting)

John Jenkins, MD

Director

Office of New Drugs (OND)

CDER, FDA

Mitchell Mathis, MD

Director

Division of Psychiatry Products

ODE-I, OND, CDER, FDA

Robert Temple, MD

Deputy Director for Clinical Science

CDER, FDA

Deputy Director (Acting)

Office of Drug Evaluation-I (ODE-I)

OND, CDER, FDA

Lisa LaVange, PhD

Director

Office of Biostatistics (OB)

Office of Translational Sciences (OTS)

CDER, FDA

Peiling Yang, PhD

Biostatistics Team Leader

Division of Biometrics I

OB, OTS, CDER, FDA

C O N T E N T S

```
C O N T E N T S (continued)
```

<u>P R O C E E D I N G S</u>

(8:00 a.m.)

Call to Order

Introduction of Committee

DR. D'AGOSTINO: Good morning. I would first like to remind everybody to please silence your cell phones, smartphones, and any other devices you have if you have not already done so. I would also like to identify the FDA press contact, Kris Baumgartner. If you are present, please stand.

My name is Ralph D'Agostino. I am acting chairperson of the Psychopharmacological Drug Advisory Committee, and I will be chairing the meeting. I will now call the Psychopharmacological Drug Advisory Committee meeting to order.

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

DR. D'AGOSTINO: Dr. Jenkins?

DR. JENKINS: Good morning. I'm John Jenkins. I'm the director of the Office of New

Drugs in the Center for Drugs.

DR. MATHIS: I'm Mitchell Mathis. I'm the director of Psychiatry Products in the Center for Drugs.

DR. LaVANGE: I'm Lisa LaVange, director of the Office of Biostatistics in the Center for Drugs.

DR. YANG: I'm Peiling Yang, Biometrics Team Leader in the Office of Biostatistics.

DR. IONESCU: I'm Dawn --

DR. D'AGOSTINO: Is Victor on the phone?

DR. DE GRUTTOLA: Yes. This is Victor De Gruttola, statistician at Harvard School of Public Health. Thanks for letting me join, but there's terrible echo. I can't really hear.

DR. D'AGOSTINO: Thank you.

DR. IONESCU: I'm Dawn Ionescu. I'm a psychiatrist at Mass General in Boston.

DR. NARENDRAN: I'm Raj Narendran, psychiatrist at the University of Pittsburg.

DR. STEIN: I'm Murray Stein. I'm a psychiatrist at the University of California, San

1   Diego and the VA in San Diego.

2   MS. BHATT: Good morning. My name is

3   Kalyani Bhatt. I'm the designated federal officer.

4   I'm with the Division of Advisory Committee

5   Consultants Management.

6   DR. D'AGOSTINO: I'm Ralph D'Agostino,

7   professor of statistics, mathematics,

8   biostatistics, and epidemiology at Boston

9   University. I was an undergraduate major also in

10   logic, which I think will be helpful.

11   MS. HIGGINS: Jennifer Higgins, consumer

12   representative.

13   DR. COMPAGNI-PORTIS: I'm Natalie

14   Compagni-Portis. I'm a psychologist and the

15   patient rep today.

16   DR. FOLLMANN: I'm Dean Follmann, head of

17   biostatistics at the National Institute of Allergy

18   and Infectious Diseases.

19   DR. GOLDBERG: I'm Judith Goldberg,

20   professor of biostatistics at NYU School of

21   Medicine.

22   DR. GOGTAY: Nitin Gogtay, director of

1   Office of Clinical Research, NIMH.

2   DR. RUDORFER: I'm Matt Rudorfer. I'm a

3   psychiatrist at National Institute of Mental

4   Health.

5   DR. CONLEY: I'm Rob Conley. I'm a

6   psychiatrist. I'm a distinguished scholar of

7   neuroscience at Eli Lilly and industry

8   representative.

9   DR. D'AGOSTINO: Thank you. Could we all be

10   heard appropriately? Thank you.

11   F or topics such as those being discussed at

12   today's meeting, there are often a variety of

13   opinions, some of which are quite strongly held.

14   Our goal is that today's meeting will be a fair and

15   open forum for discussion of these topics and those

16   individuals can express their views without

17   interruption.

18   Thus, as a gentle reminder, individuals will

19   be allowed to speak into the record only if

20   recognized by the chairperson. We look forward to

21   a productive meeting.

22   In the spirit of the Federal Advisory

1   Committee Act and the Government in the Sunshine

2   Act, we ask that the advisory committee members

3   take care that their conversations about the topic

4   at hand take place in the open forum of the

5   meeting.

6   We are aware that members of the media are

7   anxious to speak with the FDA about these

8   proceedings. However, the FDA will refrain from

9   discussing the details of the meeting with the

10   media until its conclusion. Also, the committee is

11   reminded to please refrain from discussing the

12   meeting topics during breaks or lunch. Thank you.

13   I now will pass to Kalyani Bhatt who will

14   read the conflict of interest statement.

15   MS. BHATT: Good morning. Before I read the

16   Conflict of Interest Statement, Dr. Temple, if you

17   could please introduce yourself?

18   DR. TEMPLE: Good morning. Bob Temple,

19   deputy director of ODE1.

20   **Conflict of Interest Statement**

21   MS. BHATT: Thank you. The Food and Drug

22   Administration, FDA, is convening today's meeting

1   of the Psychopharmacologic Drugs Advisory Committee

2   under the authority of the Federal Advisory

3   Committee Act, FACA, of 1972. With the exception

4   of the industry representative, all members and

5   temporary voting members of the committee are

6   special government employees, SGEs, or regular

7   federal employees from other agencies and are

8   subject to federal conflict of interest laws and

9   regulations.

10   The following information on the status of

11   this committee's compliance with the federal ethics

12   and conflict of interest laws, covered by but not

13   limited to those found at 18 U.S.C. Section 208, is

14   being provided to participants in today's meeting

15   and to the public.

16   FDA has determined that members and

17   temporary voting members of this committee are in

18   compliance with federal ethics and conflict of

19   interest laws. Under 18 U.S.C. Section 208,

20   Congress has authorized FDA to grant waivers to

21   special government employees and regular federal

22   employees who have potential financial conflicts

1   when it is determined that the agency's need for a

2   particular individual's service outweighs his or

3   her potential financial conflict of interest.

4   Related to the discussions of today's

5   meeting, members and temporary voting members of

6   this committee have been screened for potential

7   financial conflicts of interest of their own as

8   well as those imputed to them, including those of

9   their spouses or minor children and, for the

10   purposes of 18 U.S.C. Section 208, their employers.

11   These interests may include investments;

12   consulting; expert witness testimony;

13   contracts/grants/CRADAs; teaching/speaking/writing;

14   patents and royalties; and primary employment.

15   Today's agenda involves the discussion of

16   the efficacy and safety data for new drug

17   application, NDA 21164, gepirone hydrochloride

18   extended-release tablets submitted by Fabre-Kramer

19   Pharmaceuticals for the proposed indication of

20   major depressive disorder.

21   This is a particular matters meeting, which

22   during specific matters related to Fabre-Kramer's

1   NDA will be discussed. Based on the agenda for

2   today's meeting and all financial interests

3   reported by the committee members and temporary

4   voting members, no conflict of interest waivers

5   have been issued in connection with this meeting.

6   To ensure transparency, we encourage all

7   standing committee members and temporary voting

8   members to disclose any public statements that they

9   have made concerning the product at issue.

10   With respect to FDA's invited industry

11   representative, we would like to disclose that

12   Dr. Robert Conley is participating in this meeting

13   as a non-voting industry representative acting on

14   behalf of regulated industry. Dr. Conley's role at

15   this meeting is to represent industry in general

16   and not any particular company. Dr. Conley is

17   employed by Eli Lilly.

18   We would like to remind members and

19   temporary voting members that if the discussions

20   involve any other products or firms not already on

21   the agenda for which an FDA participant has

22   personal or imputed financial interest, the

1   participants need to exclude themselves from such

2   involvement and their exclusion will be noted for

3   the record.

4   FDA encourages all participants to advise

5   the committee of any financial relationships that

6   they may have with the firm at issue. Thank you.

7   DR. D'AGOSTINO: We will now proceed with

8   Dr. Jenkins' introductory remarks, followed by

9   Dr. Mitchell Mathis.

10   **FDA Opening Remarks – John Jenkins**

11   DR. JENKINS: Thank you, Dr. D'Agostino.

12   I'd like to welcome you and the members of the

13   committee to today's meeting and to thank you very

14   much for your willingness to serve in this very

15   important capacity as we work on the issues related

16   to this application today.

17   Again, I'm John Jenkins. I'm the director

18   of the Office of New Drugs in the Center for Drugs.

19   I'm going to give some introductory remarks to

20   frame today's discussion about the NDA for gepirone

21   extended-release for the treatment of major

22   depressive disorder or MDD.

1   First, I'd like to give you an overview of

2   the issues we're seeking your input on today at the

3   advisory committee. Gepirone is a new molecular

4   entity, meaning it's not approved in any form

5   currently in the United States, and to the best of

6   my knowledge, it's not approved anywhere in the

7   world.

8   It has been the subject of three complete

9   review cycles by the Division of Psychiatry

10   Products and the Office of Drug Evaluation I.

11   Not-approvable letters were issued on each of those

12   cycles in 2002, 2004, and 2007, so you can get a

13   sense that this development program and the review

14   of this application has been going on for quite

15   some time.

16   In 2012, Fabre-Kramer, the sponsor of the

17   application, submitted an informal appeal to

18   Dr. Temple as the deputy director of ODE1 asking

19   that he reconsider the 2007 not-approvable letter.

20   The psychiatry division in ODE1 has consistently

21   concluded that the sponsor has failed to provide

22   substantial evidence of effectiveness of gepirone

1   in the treatment of major depressive disorder.

2   In 2014, Fabre-Kramer submitted a formal

3   dispute resolution to the Office of New Drugs

4   challenging ODE1's not-approvable decision for

5   gepirone, and that was accepted for review in

6   January of 2015.

7   On June of this year, I, in my capacity as

8   the Office of New Drug's Director and also the

9   deciding official for the dispute resolution,

10   issued an interim response notifying Fabre-Kramer

11   of my intention to refer this application for

12   review by the Psychopharmacologic Drugs Advisory

13   Committee before reaching a final decision on the

14   dispute resolution. So that's why we're here

15   today.

16   The primary issue in dispute is whether

17   substantial evidence of effectiveness has been

18   provided as required for approval under the Food,

19   Drug, and Cosmetic Act.

20   I would note that the 2007 not-approvable

21   letter listed other deficiencies such as chemistry

22   manufacturing and controls, as well as other issues

1 that would have to be resolved if this application

2 were subsequently resubmitted for approval, such as

3 labeling and any postmarketing study commitments or

4 requirements. But those are not the topic of the

5 dispute resolution and are not the topic of today's

6 meeting.

7   Today's discussion will include the typical

8 premarket review of a new drug application seeking

9 approval, meaning it will cover both safety and

10 efficacy of the drug. But the main focus of

11 today's meeting will be on the analysis and

12 interpretation of the clinical trials submitted by

13 the sponsor in support of efficacy.

14   Let's take a minute to discuss substantial

15 evidence of effectiveness. This is always the

16 statutory standard that we apply in deciding

17 whether a drug can be approved, but it's

18 particularly relevant today to go into a bit more

19 detail so that you understand the statutory and

20 regulatory history and specific language of this

21 provision of the Act.

22   This was added to the Food, Drug, and

1   Cosmetic Act in 1962 as part of the Kefauver-Harris

2   Amendments that was the first requirement that

3   drugs be proven to be effective before approved for

4   marketing in the United States.

5   Substantial evidence is defined in

6   Section 505(d) of the Act as -- and I'm going to

7   read this so that you can kind of imprint it on

8   your brain -- "Evidence consisting of adequate and

9   well-controlled investigations, including clinical

10 investigations, by experts qualified by scientific

11 training and experience to evaluate the

12 effectiveness of the drug involved on the basis of

13 which it could fairly and responsibly be concluded

14 by such experts that the drug will have the

15 effective reports, or is represented to have under

16 the conditions of use prescribed, recommended or

17 suggested in the labeling or proposed labeling

18 thereof." That's the congressional language that's

19 in the statute.

20   Now, in interpreting this statutory

21 language, it has long been FDA's position that

22 Congress generally intended that we require at

1   least two adequate and well-controlled studies,

2   each convincing on its own to establish

3   effectiveness.

4   In 1997, as part of the Food and Drug

5   Administration Modernization Act, Section 505(d)

6   was amended to make it clear that the agency may

7   consider "data from one adequate and

8   well-controlled clinical investigation and

9   confirmatory evidence" to constitute substantial

10   evidence if FDA determines that such data and

11   evidence are sufficient to establish effectiveness.

12   That's the one modification to the substantial

13   evidence standard that's been applied since the

14   1962 legislation.

15   Now, the usual requirement for more than one

16   adequate and well-controlled investigation reflects

17   the need for independent substantiation of

18   experimental results. Independent substantiation

19   of a favorable result protects against the

20   possibility that a chance occurrence in a single

21   study will lead to an erroneous conclusion that a

22   treatment is effective.

1   However, the Food, Drug, and Cosmetic Act

2   FDA regulations and our guidance are silent on how

3   the agency should consider what I put in quotes

4   here, "negative evidence on effectiveness of a

5   drug" in reaching a conclusion that substantial

6   evidence of effectiveness has been provided. This

7   is the core of the issue and dispute between the

8   sponsor, Fabre-Kramer, and DPP/ODE1.

9   A brief overview of the gepirone ER NDA. It

10   includes 12 short-term, randomized,

11   placebo-controlled trials, using the extended

12   release formulation in the treatment of major

13   depressive disorder and one maintenance randomized

14   withdrawal trial. Five of the short-term trials

15   include an active control, and you'll see that

16   becomes important later.

17   DPP and ODE1 and Fabre-Kramer agree that

18   two of the short-term placebo-controlled treatment

19   trials are positive, meaning they have a P of less

20   than 0.05 and treatment effects similar in

21   magnitude to other approved antidepressants.

22   DPP and ODE1, however, and Fabre-Kramer

1   reached differing conclusions regarding the

2   interpretation of many of the remaining 11 trials,

3   and the interpretation includes modifiers such as

4   "negative," "failed" and "supportive." And we'll

5   discuss that a bit more in a couple of slides.

6   Now, the analysis and interpretation issues

7   that we're seeking your input on today are listed

8   on this slide. First, you'll hear about the

9   concept of assay sensitivity -- and I'll discuss

10   that in a subsequent slide as well -- and what test

11   should we apply in determining whether a trial has

12   assay sensitivity.

13   For example, should we limit ourselves to

14   the prespecified primary endpoint and analysis for

15   that trial, or can we look at prespecified

16   secondary and/or post hoc analysis even if the

17   primary analysis failed, is a question mark? What

18   if neither drug beat placebo, but the active

19   control drug beats test drug? Does that show assay

20   sensitivity and count as a negative trial?

21   There's also the question of what is the

22   appropriate method to integrate trial results to

1   determine whether substantial effectiveness exists?

2   For example, should we count the number of positive

3   trials out of the total number of positive and

4   negative trials and compute some "overall p-value"

5   for the program in making our determination?

6   If that's the recommended pathway, what

7   threshold for substantial evidence should we apply?

8   Should it be at less than 0.025, which is the

9   single trial standard we normally apply for a

10   statistically significant result?

11   Should it be less than 0.00625, which is

12   0.025 squared, which would be two trials or some

13   other value?

14   Or should we consider an approach that

15   considers the particular results of each study, not

16   just whether they were positive or negative, along

17   with its individual strengths and weaknesses? And

18   if so, what approach, such as a meta-analysis, and

19   what interpretive criteria for substantial evidence

20   should apply?

21   I mentioned assay sensitivity for trials.

22   The psychiatry division has noted for many years

1   that many antidepressants that are thought to be

2   active drugs, when studied in clinical trials, fail

3   to demonstrate efficacy, meaning that they failed

4   to beat placebo.

5   In order to help the division in

6   interpreting these trials, the division has long

7   encouraged sponsors to include a third arm in those

8   placebo-controlled trials, to include a known

9   active drug at known active doses to help assay

10   sensitivity and aid interpretation of the trial.

11   So when you hear us talk about a failed

12   trial, the way psychiatry in ODE1 have

13   traditionally used that term is that the active

14   control and the test drug both failed to beat

15   placebo; in other words, the trial lacked assay

16   sensitivity, could not detect a difference when one

17   should've been there with a known effective drug.

18   So maybe there was a flaw in the trial conduct, or

19   the design, or the endpoints, or the power.

20   A negative trial is used to describe a trial

21   where the active control beat placebo, the test

22   drug did not, so the trial had assay sensitivity,

1   and it suggests that the test drug failed in that

2   study.

3   Now, in DPP and ODE1's efficacy analysis,

4   failed trials are discounted, meaning they don't

5   really count against the drug in establishing

6   substantial evidence of effectiveness, while

7   negative trials raise concern about whether the

8   test drug is, in fact, effective.

9   One last slide to discuss the formal dispute

10   resolution process, this is a process by which

11   sponsors may formally appeal certain decisions made

12   by FDA to higher supervisory levels, and I've

13   listed at the bottom there a footnote that gives a

14   reference to the guidance for this pathway, which

15   was recently updated by FDA just this past

16   September.

17   Now, this process doesn't normally play out

18   in public, so this is a rare opportunity for the

19   advisory committee and the public to not only weigh

20   in on a dispute resolution but also to see how the

21   process actually plays out.

22   As part of our guidance, we say that either

1   the sponsor or the FDA deciding official may

2   request advisory committee input. And in my

3   capacity as the director of the Office of New

4   Drugs, as well as the deciding official for this

5   formal dispute resolution, I requested that we

6   refer this issue to the advisory committee because

7   I think the issues are very complex, but they're

8   also precedent-setting in how we apply substantial

9   evidence.

10   If after today's meeting, I were to grant

11   the dispute resolution, in other words, I determine

12   that substantial evidence of effectiveness has been

13   provided, the sponsor will be directed to resubmit

14   the NDA to the division in ODE1 to complete the

15   other work that's necessary before an application

16   can be approved.

17   In other words, even if I grant the dispute,

18   that is not translated into immediate approval

19   because there's other work that needs to be done to

20   bring the application in line for approval;

21   however, such a decision would effectively result

22   in approval since the other deficiencies are those

that could be corrected, although they may take

some time.

If, however, I deny the dispute resolution,

the sponsor has the opportunity to appeal my

decision to the center director, and then they can

even appeal the decision of the center director to

the commissioner.

With that, I'll turn the microphone over to

my colleague, Mitchell Mathis, to give you a bit

more of an overview of what we normally expect for

programs for the development of drugs to treat

major depressive disorder and some comments about

the application in front of us today. Thank you.

MS. BHATT: Before we start the next

presentation, could we please introduce the people

that have joined the panel?

DR. PICKAR: I'm David Pickar. I was former

chief of experimental therapeutics branch

Intramural, NIMH. I'm an adjunct professor at

Hopkins and Uniformed Services and run Gabriel

Sciences.

DR. MANN: Good morning. My name is

1   John Mann. I'm at Columbia University. I'm the

2   chief of the Division of Molecular Imaging and

3   Neuropathology.

4   **FDA Opening Remarks – Mitchell Mathis**

5   MR. MATHIS: Good morning, and thank you for

6   being here. My name is Mitchell Mathis. I'm the

7   director of the Division of Psychiatry Products.

8   My charge here today is to give the

9   committee an understanding of the usual standard in

10   our division for the approval of drugs to treat

11   MDD. I'll focus on efficacy, safety, and labeling,

12   and then get into the historical experience with

13   active control trials that we have in MDD; discuss

14   the usual postmarketing requirements and

15   commitments; and then I'll go over the discussion

16   and voting questions for the committee.

17   The usual standard of approval, as

18   Dr. Jenkins just went over, is from Section 505,

19   substantial evidence of effectiveness from adequate

20   and well-controlled investigations. That's been

21   interpreted by the division for MDD to mean at

22   least two positive adequate and well-controlled

1   trials.

2   For safety, we need to include all tests

3   that are applicable to show that the drug is safe

4   under the proposed labeling. That's been

5   interpreted by the division in, I think, the usual

6   way. We want to define adverse reactions, common

7   reactions, and any special safety concerns if we

8   can find them in the review.

9   The optimum approach to looking at efficacy,

10   which has been developed over many years, is to

11   have at least one dose response evaluation of a

12   drug. We get a lot of information if we have

13   different doses of the test drug in the same trial.

14   It helps us identify the minimum dose that produces

15   efficacy, the maximum dose above which you get

16   nothing but usually more side effects, and tapering

17   and titration requirements if there are any for the

18   drug. And of course, we can look at subgroups

19   based on those.

20   The use of an active control has been a big

21   part of what we do in depression because it

22   distinguishes, in our minds, the difference between

1   negative and failed trials. Fifty percent of our

2   drugs that we believe to be effective failed to

3   show an effect in active-controlled trials. I'll

4   show you some data on that. Sponsors and FDA

5   benefit, in our view, from the use of an active

6   control.

7   The usual safety evaluation for a drug for

8   MDD includes identifying any safety signals. These

9   include common adverse reactions, which are

10   normally defined as occurring at 5 percent in twice

11   the rate of placebo.

12   We generate for our labels a 2 percent

13   table, which lists everything that happened on drug

14   at 2 percent or more. If there are any drug

15   interactions that we are familiar with and have

16   evidence for, we put those in the label. And then

17   any special safety claims, if there are any, are

18   worked up and also included.

19   The international standards for a safety

20   database, a longer term safety database, we use the

21   ICH guidelines, which defines at exposures at

22   relevant doses to be 1500 overall in the

1   development program as a minimum, 300 to 600 for

2   more than 6 months of exposure, and 100 for more

3   than a year of exposure. That's a hundred people.

4   Our usual approach to labeling, also from

5   505, labeling must not be false or misleading in

6   any particular, which is pretty broad. We

7   interpret that to mean that the indication that's

8   in the label is what the sponsor studied; that

9   dosing will be done as it was done in the clinical

10   trials where we have data. The drug interactions

11   will be labeled, and if there's any deviation from

12   the usual in terms of labeling, extra claims

13   or -- then those would have to be substantiated.

14   The historical use of active control trials

15   in MDD approvals, I think this will give the

16   committee some frame of reference. We have

17   18 drugs that have been approved to treat MDD.

18   Each had at least two positive trials per our usual

19   approach, except for drugs that had had a prior

20   approval in a different formulation, for instance,

21   when there was a good reason to -- we already had

22   data that there was efficacy for the drug.

1   Forty-six studies that make up this group

2   had an active control. Of those 46, in 6 of them,

3   the active control was positive and the test drug

4   was not; and that, by definition, is a negative

5   trial. Now, those drugs are still approved because

6   there were other positive studies.

7   In 15 studies, neither the AC nor the test

8   drug beat placebo, and the studies were considered

9   failed instead of negative. In 8 studies, the

10   active control did not beat placebo but the test

11   drug did. These were considered positive studies

12   to support approval because the test drug beat

13   placebo. So overall, you can add the 15 and the 8

14   and see that the active control beat placebo in

15   50 percent, 23 of these 46 trials.

16   Here is a tabular representation of

17   antidepressant trials when considered at the time

18   of initial approval for new molecular entities, not

19   yet approved drugs. The drugs are listed on the

20   left, the number of positive, negative, and failed

21   studies. I put this here not to draw any

22   conclusions but just to provide it as background as

1   to what we've seen before, and I'm sure there will

2   be more discussion about this as we move forward.

3   The usual postmarketing studies for MDD, we

4   normally receive adult data, and we then ask for

5   maintenance data postmarketing for this chronic

6   disease, and we ask for studies in pediatric

7   depression, children and adolescents.

8   In summary, the usual efficacy requirements

9   for MDD are two positive trials, and we expect

10 about a 50 percent non-positive rate. The concept

11 of active control has been very useful to us. It's

12 been our attempt to help distinguish failed from

13 negative trials.

14   The interpretation of that, of course,

15 requires an understanding of assay sensitivity, and

16 I think we're going to speak about that today. If

17 active control beats placebo, that's been the

18 traditional understanding for assay sensitivity.

19 If active control beats test drug but neither beats

20 placebo -- and there's a question mark there as in

21 Dr. Jenkins' slide -- is that assay sensitivity?

22   Do you have to use a prespecified endpoint

1   or can a common endpoint for measuring depression

2   be used? And what do you do with the data from the

3   failed trials? Are those data to be excluded from

4   the trial or are they interpretable in some way?

5   I think the key question for the PDAC today

6   is going to be, when you have two positive trials

7   that the sponsor and FDA agree are positive, how

8   much should failed or negative data influence the

9   efficacy decision?

10   In summary, for gepirone in adults, there

11   are two positive trials that have been presented

12   and accepted, three negative trials that have been

13   presented and accepted as negative. There are four

14   additional trials in adults that have been the

15   focus of the review team.

16   These trials are failed if you look at the

17   prespecified endpoint and analysis plan from the

18   sponsor, but they could be considered negative, and

19   have been in the past, if examined with a commonly

20   accepted endpoint and a modified analysis using the

21   HAMD-17.

22   I also want the committee to be aware that

1   there are pediatric trials that have not been a

2   part of this application. The sponsor is not

3   seeking a pediatric indication. But there are two

4   trials in pediatric patients that also have

5   negative results. They were not active-controlled.

6   The questions for discussion and voting, the

7   first discussion question will be to discuss the

8   following questions related to substantial

9   evidence.

10   In a situation where two positive and

11   adequate and well-controlled trials have been

12   completed, how much and what type of negative

13   evidence from negative or failed trials would it

14   take to undermine a finding of substantial evidence

15   of effectiveness? What approaches for synthesizing

16   evidence across positive and negative failed trials

17   in a development program are useful for

18   decision-making?

19   Second question we'll ask you to discuss

20   your views on ways to evaluate clinical trials for

21   assay sensitivity and to consider the following

22   questions when you do that: Is the primary

1   endpoint for efficacy prospectively defined in the

2   protocol the only meaningful way to evaluate assay

3   sensitivity? Can post hoc analyses of other

4   endpoints or use of other analysis methods

5   contribute to the determination of assay

6   sensitivity?

7   Questions 3, 4, and 5 are for a vote. The

8   first question, has the sponsor provided

9   substantial evidence of effectiveness for gepirone

10   extended-release in the treatment of major

11   depressive disorder?

12   Question 4. Has the sponsor adequately

13   characterized the safety profile of gepirone ER in

14   the treatment of major depressive disorder?

15   Question 5. Do the available data support a

16   favorable benefit-risk profile of gepirone ER to

17   support approval?

18   After the vote, we'd like to you discuss

19   what, if any, additional studies are needed pre- or

20   post-approval to address outstanding issues, for

21   instance, additional effectiveness study, an

22   additional randomized withdrawal trial, et cetera.

1   That's it. Thank you.

2   DR. D'AGOSTINO: Thank you. And also,

3   Dr. Jenkins, that was extremely important, and I'm

4   sure the committee will benefit greatly when we

5   come to the discussion and vote questions because

6   of it. We're now going to move on.

7   Both the Food and Drug Administration and

8   the public believe in a transparent process for

9   information-gathering and decision-making. To

10   ensure such transparency at the advisory committee

11   meeting, FDA believes it is important to understand

12   the context of an individual's presentation.

13   For this reason, FDA encourages all

14   participants, including the sponsor's non-employee

15   presenters, to advise the committee of any

16   financial relationships that they may have with the

17   firm at issue such as consulting fees, travel

18   expenses, honoraria, and interest in the sponsor,

19   including equity interest and those based upon the

20   outcome of the meeting.

21   Likewise, FDA encourages you, at the

22   beginning of your presentation, to advise the

1   committee if you do not have any such financial

2   relationships. If you choose not to address the

3   issues of financial relationships at the beginning

4   of your presentation, it will not preclude you from

5   speaking. We will now proceed with the sponsor's

6   presentation.

7   **Industry Presentation – Daniel Burch**

8   DR. BURCH: Thank you. Good morning. My

9   name is Daniel Burch, and I'm pleased to be

10 representing Fabre-Kramer in the presentation. I

11 want to go ahead and read my conflict statement.

12 My company has been compensated for my time

13 and travel preparing for this, but I have no

14 financial interest personally, and neither does my

15 company, in the outcome of this meeting.

16 I'd like to tell you a little bit more about

17 who I am. I'm actually head of global neuroscience

18 at a large CRO called Pharmaceutical Product

19 Development. I got to know Fabre-Kramer some years

20 ago when I worked for another organization in

21 collaboration discussions and have been interested

22 in the gepirone story.

1   I've actually been privileged to work on two

2   approved medications for major depressive disorder,

3   as well as six, seven, eight -- I've lost

4   track -- candidate medications for major depressive

5   disorder.

6   But probably more relevant and more

7   important is in my current position and in past

8   positions, now having worked in the

9   neuropsychiatric field for over 12 years, I'm very

10   concerned about this issue of assay sensitivity and

11   the type of impact it has on drug development, the

12   inefficiency that I witness in psychiatry trials,

13   neurology trials, and how it can be improved.

14   In my organization, we're collaborating with

15   academia, with government, with other industry

16   colleagues to try to improve the efficiency and

17   effectiveness of the trials themselves, so this is

18   particularly of interest to me.

19   This has already been said many times;

20   there's a lot of literature on this. About

21   50 percent of all of trials with all approved

22   antidepressants fail. When you come into

psychiatry or pain, for example, and you hear the

statistics, it takes you back a little bit. It

makes you go, "Wow, that's really incredible." But

that's the case, and I think that's been backed up

by numerous studies, including the Division of

Psychiatric Studies.

Some of the things, if you look back over

the last three or four decades -- and there are

publications on this -- you'll see that the placebo

response, for whatever reason, is increasing and

the drug signal is decreasing, from the '80s, to

the '90s to the 2000s. And that's the context that

we're talking about here with the development of

gepirone and other products.

The placebo response gets bashed a lot at

meetings, which is actually the friend of the

clinician, of course. It affects the clinical

trialists because it impairs assay sensitivity.

We'll talk about it, I'm sure, more over today. So

placebo response variability, we're trying to get a

handle on these things with modern methods in doing

clinical trials.

1   Assessing individual subjects independently

2   is another important matter. Independently

3   assessing disease of varying thresholds, endpoint

4   management, adherence and retention, these are all

5   things that we're paying a lot more attention to

6   over the last 5 or 10 years in neuropsychiatric

7   trials.

8   We're going to see a lot of -- Dr. Mathis

9   has already showed some slides. But I wanted to

10   just sort of look at the sort of class of 2000-ish

11   group of medications for major depressive

12   disorder: vilazodone, duloxetine, desvenlafaxine,

13   citalopram. The percentage of positive studies is

14   around 30 percent.

15   What's also important to notice -- I'm sorry

16   I didn't put a column on this -- is that the number

17   of trials -- these are registration-directed

18   trials. The number of trials ranges from 7 to 10.

19   We're talking about 12 short-term trials and

20   1 maintenance trial in this discussion. So

21   already, we're not too far of an outlier.

22   I just wanted to provide perspective. It's

1   not like these major depressive disorder

2   submissions go in usually with two sort of neat and

3   tidy registrational studies. They're usual

4   multiple registrational studies because of the

5   problem of failed trials and negative trials.

6   This is a worthy discussion because this is

7   an important illness. Everybody knows that -- and

8   Dr. Thase is going to talk about this in more

9   detail, but major depressive disorder is the number

10   one cause of disability in developed countries.

11   This is an important subtlety for everyone

12   to realize, that the treatment for major depressive

13   disorder is empiric and it's inadequate. When a

14   patient comes into a psychiatrist's office, they

15   don't have a little sign on them or a biomarker

16   that can be assessed and say I'm going to respond

17   to an SSRI. I'm going to respond to SNRI. I'm

18   going to respond to a tricyclic.

19   There's trial and error involved, and

20   there's art and science and policy issues about why

21   certain patients get certain medications and when.

22   Switching is very common. I cited a recent study

1   where over 11,000 patients received 2.65 average

2   therapies within the prior 18 months, 11,000

3   patients-plus that were undergoing switching. And

4   then the Star*D trial, which many of you are

5   familiar with and Dr. Thase will talk about, showed

6   that only one-third of patients remit on their

7   first course of antidepressant.

8   The safety and tolerability profiles have

9   come a long way since tricyclics and MAO

10   inhibitors, but they still leave something to be

11   wanted. The SNRIs and SSRIs still have issues with

12   weight gain, sexual dysfunction, other CNS

13   symptoms, sedation, and so forth. That is

14   something that frequently drives overall acceptance

15   in both primary and secondary non-adherence.

16   Why are we here? I think this has already

17   been covered by the agency. But we're here to

18   discuss the approvability of gepirone ER for major

19   depressive disorder in adults. We're here to

20   discuss whether substantial evidence of

21   effectiveness has been achieved in major depressive

22   disorder.

1   The sponsor, Fabre-Kramer, is not seeking

2   claims for long-term treatment, children or

3   adolescent claims, or any specific claims in sexual

4   dysfunction. Dr. Stahl will talk about this later,

5   about the rationale, but the proposed dose is 60 to

6   80 milligrams orally per day.

7   Dr. Jenkins already outlined the history

8   here. We're here in the wake of a dispute

9   resolution, and I think there's already been

10   outlined we have areas of agreement. We would

11   submit that the safety of gepirone is in keeping

12   with other generally prescribed antidepressants.

13   We would also submit, and I think we agree, that

14   studies 001 and 007 are adequate, well-controlled

15   and robustly positive independent trials, each

16   convincing on its own.

17   Areas for discussion, the magnitude of

18   treatment effect, I think we'll touch on that and

19   happy to go into more details. I think that was

20   raised in the briefing document by the division.

21   And probably the most core is the interpretation of

22   the remaining 10 short-term trials. There's also

1   one long-term trial that we'll discuss briefly as

2   well.

3   I want to presage Dr. Stahl's talk and just

4   quickly introduce gepirone to everybody. It's an

5   azapirone analogue with a 3-hydroxy metabolite as

6   an active species. The sort of reason to believe

7   on gepirone, it's a selective 5HT 1A receptor

8   partial agonist. It doesn't affect the 5H T2A

9   receptor, which can be associated with adverse

10 events as most everyone in this audience knows.

11 But probably simply put, it's not an SSRI, and it's

12 not an SNRI.

13   When we talk about gepirone, we're going to

14 be talking about the extended-release formulation,

15 and Dr. Stahl will provide a brief rationale why

16 extended release is important.

17   So it's a big hairy beast, the 12 trials,

18 and I'm going to try to simplify it if you will

19 allow me to. The reason it's complicated is

20 because there were three sponsors: There was

21 Bristol-Myers Squibb initially; there was

22 Fabre-Kramer that acquired it from Bristol-Myers

1   Squibb; and then there was Organon that

2   Fabre-Kramer had a relationship with that has been

3   terminated several years ago.

4   With three sponsors, you're going to expect

5   different business agendas, different development

6   strategies, different stakeholders to manage.

7   That's why we see 12 trials here. It just didn't

8   happen by accident because someone was trying to

9   throw something up against the wall again and

10   again.

11   Let's put these 12 trials, the 12 short-term

12   trials, into three different buckets. The first

13   bucket is pretty easy to conceptualize. These were

14   early studies, the earliest studies done, the ones

15   on your left by Bristol-Myers Squibb. And because

16   Bristol-Myers Squibb was developing another

17   antidepressant, nefazodone, and had some positive

18   data on nefazodone, they decided, we want to manage

19   our portfolio, we want to manage our business

20   rationally, and they just discontinued those four

21   trials prematurely, before they met their

22   prespecified sample size requirement.

1   The next bucket is a group of trials that

2   were done by both Fabre-Kramer and Organon. These

3   are replicate trials, 001, 002, 007, 008 and 023.

4   These trials were all done in response to FDA

5   advice, the Division of Psychiatric Product advice,

6   in one way or another, over a period of time.

7   Then finally, there were three studies

8   sponsored by Organon, studies 004, 006, and 017.

9   They were done for market differentiation purposes,

10   trying to show this drug worked in a certain

11   population and so forth.

12   I'm going to show you again the list of all

13   the studies, but I want to draw your attention to

14   the ones that I have in boxes. Again, these are

15   the ones that are replicate trials: 001, 002, 023,

16   007, and 008. Those were all done on advice from

17   the Division of Psychiatric Products.

18   But more importantly, they are trials that

19   did not have an active control, but the internal

20   characteristics -- as a clinical trialist that has

21   done a lot of depression trials, I look at, things

22   like severity threshold, dropout rate, and placebo

1   response -- all those were in keeping with one

2   expecting an outcome of a good clinical assay. In

3   other words, I can trust these studies.

4   The other trials, I want to just review

5   again, the top four there, those are the

6   Bristol-Myers Squibb drugs. I just want to point

7   out that 052 and 053, they had active controls.

8   Technically, all of these four trials on the top

9   were failed trials. We believe the ones in the

10   middle were either positive or negative. There

11   were no failed trials there because of the internal

12   metrics that I just mentioned.

13   The last four trials, 004, 006 and 017 on

14   the bottom, those were all technically, again,

15   failed trials by the definitions that have been

16   discussed because the active didn't separate from

17   placebo.

18   004 and 006 were done in an atypical

19   depression population. That's a special subtype of

20   major depression in contrast to major depression.

21   Dr. Thase will talk about that in his presentation

22   briefly. But it's a separate population.

1   017 was done, again, for market

2   differentiation, specifically interested in the

3   question of sexual dysfunction in gepirone versus a

4   comparator.

5   I wanted to show you another slide, and I

6   was kind of raised in the industry, hearing the

7   FDA, was very much interested in active comparator

8   trials. We always tried to build them in, and we

9   actually felt guilty if we didn't have the

10   resources to do so.

11   In the last five years, frankly, I've

12   migrated to the point where I think if you have

13   good internals, if you have low placebo response,

14   you're paying attention to that, you independently

15   assess subjects, you monitor adherence, you have

16   rater training certifications, surveillance and so

17   forth, that's what you really need. And you look

18   at the results, and you look at placebo response,

19   you look at variability, look at the severity

20   thresholds coming in and so forth, and that's how

21   you judge whether you have a good assay or not.

22   Because the tail of the tape is, again, a

1   contemporaneous group of antidepressants, the

2   percentage of failed studies -- and these again are

3   registrational trials; these come from the summary

4   basis of approval. Sixteen out of 18 failed in

5   this collection: citalopram, escitalopram,

6   duloxetine, desvenlafaxine, and vilazodone.

7   I actually didn't realize this until the

8   last few weeks when I started preparing for this.

9   This is sort of conventional wisdom, but here it is

10   now in black and white.

11   From a clinical point of view, we think the

12   most relevant studies, they're all either positive

13   or negative, are the five that I outlined before:

14   001, 002, 007, 008, and 023, all done in the wake

15   of FDA guidance, all with good internals.

16   001 and 007 both met their primary endpoint.

17   They also met most of their secondary endpoints.

18   We believe that they're robustly positive and each

19   stand convincingly independently on its own.

20   002 and 008 have some directional trends and

21   hit some secondary endpoints, if I ignore the

22   statisticians who say you can't go beyond the

1   primary and so forth. And 023 was just a flat

2   negative study. We see this in major depressive

3   disorder.

4   Now, I want to also presage Dr. Gary Koch's

5   talk. He's done a meta-analysis and he's going to

6   discuss that in detail. The meta-analysis of these

7   five studies, but also the entire range of

8   short-term trials up to 12 trials, was favorable on

9   the primary outcome variable and on the responder

10   variable.

11   Our thesis is that even if you consider the

12   10 short-term trials outside of 001 and 007, that

13   does not undermine our conclusion that these are

14   robustly positive and stand on their own, and

15   substantial evidence of effectiveness has been

16   demonstrated.

17   I wanted to go back to that long-term

18   maintenance trial. I don't really have a slide on

19   it. But we agree that that study -- I believe that

20   study was a failed study. It had a high dropout

21   rate. And the placebo group did not behave as

22   expected, and neither did the drug group in terms

of number of relapses, and it needs to be repeated.

No question about that. No argument with that.

And other long-term maintenance studies have

failed, although we can see that's generally a more

sensitive assay than the short-term trials.

I want to set up the rest of our

presentation. I think there's no disagreement that

major depressive disorder is a major public health

problem. Dr. Koch will talk about the totality of

evidence in his meta-analysis supporting the

substantial evidence of effectiveness assertion

that I just made based on the clinical view.

Dr. Stahl is going to talk about the

different mechanism of action. He's also going to

discuss the two positive trials, as well as all

five trials that I mentioned as being

interpretable. Then finally, Dr. Stahl is going to

discuss the safety and tolerability.

With that, we'll go on with the agenda, and

I'll allow Dr. Thase to start making his way up

here. Dr. Michael Thase from the University of

Pennsylvania will be the next speaker. He's going

1   to be followed by Dr. Gary Koch from the University

2   of North Carolina talking about the totality of

3   evidence for effectiveness. Dr. Stephen Stahl will

4   talk about the mechanism and the clinical data, and

5   then I'll follow up with brief conclusions.

6   I also just wanted to point out that we have

7   some additional consultants involved. We have

8   Dr. Anita Clayton and Dr. Leonard Derogatis from

9   the University of Virginia and Johns Hopkins

10   University, respectively, able to take questions

11   here today; Dr. Mary Johnson, a statistical

12   consultant, also able to take questions today.

13   Dr. L.J. Wei unfortunately had to be out of

14   the country; couldn't be here today. But Dr. L.J.

15   Wei was very instrumental collaborating with

16   Dr. Mary Johnson, collaborating with Dr. Gary Koch

17   on the meta-analysis data that you're going to see

18   presented shortly.

19   Thank you very much. Dr. Thase?

20   **Industry Presentation - Michael Thase**

21   DR. THASE: It's indeed a sobering honor to

22   have the chance to testify before you today or at

1   least to give this presentation. I'm a

2   psychiatrist. My life's work has been about the

3   assessment and treatment of depression. I would

4   say that I've spent more than half of the last

5   33 years doing this, and in the course of that have

6   examined both pharmacotherapies and

7   psychotherapies, ECT, TMS, and other interventions.

8   We can begin with the common need, and that

9   is that major depressive disorder is one of the

10   world's great public health problems. There are

11   just no ands, ifs, or buts about that. In a room

12   with this many folks, at least 20 of us have

13   suffered an episode of depression at some point in

14   our lifetime.

15   It's a condition that is, in industrialized

16   nations, one of the top four public health problems

17   throughout the world, and the United States is

18   projected to be the number 2 public health problem

19   within this decade.

20   In current dollars, nearly $100 billion a

21   year is lost to depression. Most of this is loss

22   from missing work and from early death.

1   Depression, of course, is the leading reason why

2   people kill themselves, and suicide is never lower

3   than the 12th leading cause of death throughout our

4   lifespan.

5   This is a serious problem. It warrants the

6   best possible interventions that we can construct

7   and make. I think that our patients are entitled

8   not only to better treatments, but better tolerated

9   treatments when possible.

10   One of the other honors I've had in my

11   career is to be one of the task force members of

12   the last two practice guidelines for the American

13   Psychiatric Association. In our most recent

14   practice guideline, which was published in 2010,

15   the state of the art was really summarized by the

16   SSRIs and the SNRIs as the treatments most likely

17   to be considered first-choice or first-line

18   options, with mirtazapine and bupropion also

19   considered.

20   There is still some role for older

21   antidepressants, including tricyclics and monoamine

22   oxidase inhibitors, as well as newly introduced

1   antidepressants that have not yet found their

2   proper place in treatment hierarchies.

3   We do not have any treatment that is so

4   superior in efficacy or tolerability that is

5   uniformly regarded as the first choice treatment

6   for all patients, and we make treatment decisions

7   on an individual basis, including likely side

8   effects or the match between the patient symptoms

9   and the drug's clinical characteristics.

10   Often, in a busy market like Philadelphia

11   where I work, even cost and what's available

12   generically might lead to a treatment being ranked

13   first. Ultimately, if you know what's worked for a

14   patient in the past, you pick that medicine first

15   unless there's some new and important reason to go

16   otherwise.

17   I realize in my nervousness, I forgot to

18   read my conflict of interest statement, which I'm

19   obliged to do, so please excuse me.

20   I am here today as a consultant to

21   Fabre-Kramer Pharmaceuticals. I have been involved

22   with gepirone even as a fellow in studies of the

immediate-release form of gepirone and as a

consultant with Organon. Specifically today, I've

been paid for my time and will be compensated for

my travel, but I have no financial interest in the

outcome of this hearing.

Here's my take on the unmet need circa 2015,

and the unmet need number 1 is that we need better

antidepressants. We need stronger antidepressants

if that is humanly possible. If may be that we

need antidepressants that match better for some

patients than for other patients.

But what we have are antidepressants that

have modest effects, on average, in absolute terms,

only about 10 to 20 percent better than placebos

and double-blind studies.

For given patients, we need better tolerated

antidepressants. In our practices, about 1 in 10

patients stop a medication within the first month

of treatment because of an intolerable side effect.

We need antidepressants that match particular

symptom clusters better than our available agents;

insomnia, anxiety, two examples of symptoms that

1   often predict poor outcomes.

2   All of our reuptake inhibitors are slow to

3   onset. This is probably the nature of this

4   mechanism of action, that these antidepressants set

5   in motion a series of events that lead to changes

6   in genes that are probably stress response genes,

7   perhaps resilience genes. But this process takes

8   weeks, if not several months, to go from ill to

9   well.

10   Finally, we just need antidepressants that

11   work differently because the vast majority of our

12   medications work through serotonin reuptake

13   inhibition or serotonin and norepinephrine reuptake

14   inhibition.

15   These are very useful, very safe

16   antidepressants. They have made a world of

17   difference in the likelihood of dying overdose, but

18   they don't treat everyone. And as I'll show you in

19   a moment in our largest scale study ever done in

20   the U.S., about a third of the patients failed to

21   remit even after a year of treatment by dedicated

22   physicians.

1   These are the data that I wanted to show you

2   from a study called STAR*D, which was conducted

3   across the country in a relatively representative

4   population of outpatients seeking treatment for

5   depression between about 2000 and 2005. We're

6   looking at remission rates here.

7   If we wanted to talk about response rates,

8   just start the first number up at 50 and then

9   proportionally go down. But here's the probability

10   of remitting with the first treatment trial, if

11   that didn't work with the second, if that didn't

12   work with the third, and if that didn't work with

13   the fourth.

14   You will see that the easy-to-treat are

15   easily treated, but the folks who don't benefit

16   from the easy treatments have a very low likelihood

17   of responding with our best available strategies,

18   best available strategies circa 2005.

19   The probability that a third-in-line

20   treatment will work is close to 20 percent for

21   response rate; 15 percent for remission rate. Then

22   likewise, if you look at the probability of still

1   remaining ill at the end of up to a year of

2   treatment, you've got about a third of the

3   population still not remitted at the end of the

4   treatment. So there is a great unmet need for

5   different treatments and, of course, for better

6   treatments.

7   Since STAR*D, the public health demand for

8   better treatments has attempted to be addressed by

9   a series of new antidepressants, none of which have

10   been paradigm changing. We've seen a new monoamine

11   oxidase inhibitor. This one delivered by a skin

12   patch. We've seen two more

13   serotonin-norepinephrine reuptake inhibitors.

14   We've seen two modified serotonin reuptake

15   inhibitors, modified to have other secondary

16   effects.

17   But all of these have similar limitations as

18   the existing agents. So there has not been a

19   paradigm-changing solution to the problem, the

20   unmet needs that I've described.

21   I'm just going to say the tiniest bit about

22   signal detection, but the attempt to find new

1   treatments has really, really been compromised by

2   the inability to show that good treatments actually

3   work in contemporary clinical trials.

4   You hear the number 50 percent bandied

5   about, but it's probably closer to 45 percent in

6   the newer generation studies, and the placebo

7   response rate grew each decade from the '80s

8   through the time from 2000 to 2010.

9   Now, the major single factor for this

10   problem is the rising placebo response rate.

11   Placebo has not become better. The patients who

12   enroll in these studies are apparently less

13   representative of the whole than the patients who

14   were 20 or 30 years ago; or maybe we're providing

15   generally better care, and the better care part

16   inflates the nonspecific benefit of the treatment.

17   Now, the average placebo response rate is

18   over 40 percent in contemporary studies, and we

19   know looking backwards that when the placebo

20   response rate goes over 40 percent, the chance of

21   study failure skyrockets. So instead of having

22   50 percent chance of failure, maybe it's more like

1   80 percent chance of failure with a high placebo

2   response rate.

3   Importantly, it's hard to even know if a

4   good antidepressant actually works, so imagine the

5   challenges that makes for a manufacturer with an

6   unknown potential antidepressant.

7   This is one of the meta-analyses, and I just

8   pulled this. There are dozens I could have pulled.

9   This one I've pulled for two reasons: First of

10   all, each of these data points represents a study.

11   I think it's a pool of five investigational, the

12   new drug application submissions. What we've got

13   here, each dot is a study. The size of the dot is

14   the size of the study.

15   But the point I simply want to make is

16   you'll notice that the average pre-treatment

17   severity score here is around about a 25. The

18   average difference in all the studies is around

19   about 2 and a half points on Hamilton Scale, and

20   you just see a scattering of study results, half

21   above this mean, half below this mean as a normal

22   distribution should be.

1   There's a single study that enrolled very

2   mildly depressed people. That study failed

3   miserably, no drug placebo difference whatsoever.

4   Even for a drug that works, occasionally, you see a

5   nominal mean difference in which placebo is

6   slightly more effective than drug.

7   You see no example where placebo was

8   significantly more effective than a proven

9   antidepressant, but you see plenty of examples in

10   which the active medicine is not much better than

11   the placebo.

12   I think importantly, you'll notice that the

13   cutoff for the average severity is well to the

14   right of the Hamilton score of 20 and that the

15   regression line shows that the drug placebo benefit

16   grows as severity increases.

17   This has been replicated dozens of times.

18   I'm pretty sure this finding was in the first

19   review papers of the efficacy of tricyclics in the

20   mid-1990s so that when patient illness severity is

21   greater, the absolute benefit or the relative

22   benefit of drug over placebo is larger.

1   Now, these data do not come from industry.

2   They actually come federally-funded or

3   foundation-funded investigators working with even

4   more mildly depressed people. But now we're

5   looking at -- each circle represents a bunch of

6   patients at a given level of severity. The

7   regression lines, again, show the exact same

8   finding in the earlier meta-analysis, namely that

9   the more severity increases, the relatively greater

10   advantage for the active drug over placebo.

11   You've got drug placebo difference starting

12   to separate in, say, 20 to 25 range, and in the

13   less numerous most severely depressed patients, you

14   get a large drug placebo difference. But really,

15   importantly for patients below this Hamilton score

16   of 20, which constitute about one-half of depressed

17   outpatients seeking treatment in non-research

18   settings, many of these patients don't even meet

19   criteria for a major depressive episode. What you

20   see is there's just no hint of a reliable

21   separation between drug and placebo.

22   Placebo effects are good things. As

1   Dr. Burch said, we capitalize on them every day in

2   our treatment process. They represent hope and

3   expectation and being cared for in the process of

4   caring for, but they do compromise detecting drug

5   effects in mildly depressed populations.

6   Among the other unmet needs, we can talk

7   also about tolerability because one of the greatest

8   reasons people stop an effective antidepressant is

9   because of an intolerable sexual side effect.

10   Sadly, our most popular and widely prescribed

11   antidepressants reliably suppress desire and

12   reliably make it more difficult to have an orgasm.

13   Depression reduces sexual desire.

14   Depression reduces interest in sexual activities,

15   may be associated with some increased rate of

16   specific sexual dysfunctions. But then,

17   ironically, the very treatments we've often

18   prescribed to treat depression can amplify these

19   problems.

20   To just give a kind of meta-analytic

21   perspective of the current landscape, here, we're

22   looking at percents of patients that were likely to

1   report sexual dysfunction, a wide range of

2   contemporary antidepressants. Here's placebo, just

3   to lock your eyes, so about 15 percent of patients

4   will report some sexual difficulty that they've

5   attributed to the placebo. This is the base rate.

6   You will notice that at the level of

7   placebo, there is only a small handful of

8   antidepressants, but none of them are called

9   serotonin, selective serotonin reuptake inhibitors.

10   The novel mechanism antidepressants are the ones

11   that are not associated with sexual dysfunction.

12   Having such a small group of antidepressants

13   really does represent a final unmet need, and that

14   is for treatments that are effective but have a low

15   risk of adverse sexual events.

16   I've had the easiest task of the people

17   speaking today, and that's to make the case that

18   depression is a serious problem and that we need

19   better treatments, both better and different

20   treatments.

21   We need treatments that work for those who

22   aren't helped by our -- those who fail to benefit

1   from our most widely used antidepressants and we

2   need treatments that can be tolerated and helpful

3   for people who can't tolerate otherwise effective

4   antidepressants. So thank you so much for your

5   kind attention.

6   **Industry Presentation – Gary Koch**

7   DR. KOCH: I'm Gary Koch from the University

8   of North Carolina, Chapel Hill. My only financial

9   relationship with the sponsor is through a

10   cooperative agreement with the University of North

11   Carolina. That cooperative agreement provides

12   partial funding, for my regular university salary

13   has a structure for reimbursement of travel.

14   I have no other financial relationship to

15   report with respect to the sponsor. Also,

16   cooperative agreements with the University of North

17   Carolina is the only relationship that I have with

18   any sponsor from industry.

19   I'm going to go over my interpretation of

20   the totality of evidence for effectiveness. I have

21   reviewed the sponsor's briefing book, the FDA

22   briefing book. I've looked fairly closely at what

1   the questions are and what the evidence is, and I

2   will essentially describe how I have identified a

3   process to work through this information, as well

4   as what my interpretations are.

5   Initially, I will review studies 134001 and

6   the 007 study. As you've already heard, these are

7   two adequate and well-controlled studies that

8   demonstrated efficacy through a statistically

9   significant result on the primary endpoint by the

10   prespecified method of analysis for that. And

11   also, they had robustness through the finding for a

12   number of secondary endpoints.

13   I will then describe how meta-analysis can

14   be a tool to identify the extent to which the

15   efficacy from the two positive studies is still

16   demonstrated from a sensitivity analysis point of

17   view after dilution via their integration with

18   three additional studies, which were the three

19   additional studies that were most comparable to

20   them in terms of study design, regulatory

21   objectives, and a number of other considerations.

22   I will then further describe how

1   meta-analysis can again be a tool for sensitivity

2   analysis for efficacy results with more extensive

3   dilution through inclusion of additional and less

4   comparable studies.

5   Here, I will initially consider the two

6   studies, which have been identified as positive,

7   the 01 study and the 07 study. On the left, you

8   see a summary of the results for HAMD-17, change

9   from baseline to endpoint with an LOCF way of

10   dealing with missing data. The p-value as re 0.018

11   for the HAMD-17 in 01; 0.032 in 007.

12   You can see that there are also low

13   p-values, below 0.05, for a number of the other

14   continuous endpoints for both 01 and 07. Then

15   there were also responder variables in these

16   studies, and this summarizes some of their results.

17   The primary endpoint met its objective with

18   a statistically significant p-value below 0.05 by

19   the prespecified primary method of analysis, and

20   then the secondary endpoints, many of them, in

21   fact, most of them, also had p-values below 0.05 as

22   well, supporting the robustness of the findings for

1   the primary endpoint.

2   Now, here, we will consider the role of

3   meta-analysis for sensitivity assessments for the

4   efficacy results. We'll evaluate homogeneity and

5   strength of evidence as we broaden the criteria for

6   the inclusion of studies in this meta-analysis.

7   This is a different type of use of

8   meta-analysis. It's essentially using

9   meta-analysis as a tool to do sensitivity analysis

10   because in this particular paradigm, there are

11   already two self-standing positive studies. One

12   does not need a meta-analysis to make a number of

13   studies that had borderline findings have a

14   positive overall result.

15   Here, the purpose is to actually identify

16   the extent to which when you apply meta-analysis,

17   the results weaken by essentially integrating in

18   less favorable results with your two positive

19   studies at the beginning. And a major role of

20   these meta-analyses is to try to get a further

21   assessment as to whether the findings from 01 and

22   07 are chance findings when you consider them in

1   the background of the totality of all of the other

2   studies.

3   Now, we begin with strict criteria to

4   identify the interpretable studies for integration

5   through the fact that they have the most in common

6   with the two positive studies in terms of the same

7   primary endpoint, the same inclusion criteria in

8   terms of baseline severity, and a number of other

9   considerations.

10   Now, we will expect the treatment effect to

11   be diluted as less comparable studies are

12   additionally included and the integration becomes

13   more heterogeneous. Reasonable preservation of

14   overall strength of evidence well supports the

15   efficacy results from 01 and 07, not being due to

16   chance, regardless of their dilution via their

17   heterogeneous integration with less comparable

18   studies.

19   We have some criteria for classifying an

20   individual study as interpretable for

21   efficacy: adequate and well-controlled, completed

22   in accordance with the protocol; appropriate study

1　population; major depressive disorder; sufficient

2　severity of illness at baseline; minimum HAMD score

3　with required at entry or substantial majority of

4　patients with HAMD-17 greater than or equal to 20.

5　　　　The reason why I focus on HAMD-17 greater

6　than or equal to 20 was that was an entry

7　requirement for 01 and 07. I'm interested in the

8　extent to which when I consider the other studies,

9　do they contradict what was in 01 and 07, and that

10　is essentially an important focus. To do that, I

11　must work with the same population that 01 and 07

12　included, which is the population that has HAMD-17

13　greater than or equal to 20.

14　　　　Sufficient dosing is another issue, and two

15　of the studies that were very early had an

16　assessment that involved the average of a high dose

17　and a low dose versus placebo. Then the studies

18　that had early termination had the issue that the

19　patients didn't complete the full follow-up period

20　because of this termination for business reasons,

21　so they didn't get the full benefit of the dosing

22　throughout the study.

1   Then we have the issue of assay sensitivity.

2   Now, in my opinion, when assay sensitivity applies,

3   it shouts at you that it's something that's clearly

4   demonstrated, that the active control will separate

5   from placebo just as strongly as when we go back to

6   the two positive studies, the test treatment

7   separated from placebo; that here, we see in these

8   two positive studies how the test treatments

9   separates from placebo.

10   If we're looking at assay sensitivity in an

11   active control study, we would want to see a range

12   of p-values like this, particularly including the

13   primary endpoint. In my opinion, when you have to

14   go looking for assay sensitivity, you really don't

15   have that.

16   Now, that still will not be a reason for

17   totally dismissing these other studies, and the

18   meta-analyses that I will consider will indeed

19   include them. But in some sense, their role does

20   need to have some recognition of discounting, and

21   that's why I include them at the latter part of the

22   meta-analyses rather than at the beginning.

1   Now, there also are criteria for classifying

2   an individual study as uninterpretable for

3   efficacy. This can include premature termination.

4   This has the issue of both low power and incomplete

5   data.

6   Insufficient dosing, we've already noted

7   that. Two of the studies had a low-dose group.

8   Insufficient baseline severity, the previous

9   presentation touched on the inability of patients

10   within insufficient baseline severity to be

11   sensitive to the detection of treatment

12   differences.

13   Then we just noted previously the issue of

14   lack of assay sensitivity, a then a further

15   criterion from the FDA briefing document is simply,

16   there are some studies that the FDA and the sponsor

17   mutually agree are very difficult to interpret for

18   efficacy.

19   Here, we have a summary, again, of

20   interpretable versus uninterpretable.

21   Interpretable studies are more suitable for a more

22   interpretable sensitivity-oriented meta-analysis,

1   and that's the type of meta-analysis that we're

2   doing here. It's a sensitivity analysis, adequate,

3   well-controlled, homogeneous with respect to key

4   design features regardless of outcome.

5   This will include positive studies, those

6   that are statistically significant for the

7   treatment effect for the primary endpoint; studies

8   that demonstrated assay sensitivity when they

9   included an active control arm. We'll also include

10   negative studies as previously noted even though

11   they do not identify a significant separation.

12   Uninterpretable or failed studies can have

13   inclusion in what I call a pessimistic sensitivity

14   meta-analysis for which heterogeneity can adversely

15   influence interpretation. These kinds of studies

16   may have deficiencies in design or conduct, as well

17   as having lack of assay sensitivity.

18   Here's an overview of the criterion for

19   exclusion of studies. Prematurely terminated as

20   you've heard before is one criterion. Insufficient

21   dosing is another criterion. Lack of assay

22   sensitivity is the third criterion. Insufficient

1   baseline severity is a fourth criterion.

2   You can see all of the studies that have X's

3   are studies that might be excluded. A couple of

4   the question marks simply mean that insufficient

5   dosing or lack of assay sensitivity may not be

6   clear.

7   I actually agree with the FDA that in the

8   053 study, there was a reasonable demonstration of

9   assay sensitivity for a more correct analysis. I

10   also found the 053 study interesting because it had

11   two centers. One center actually completed the

12   study, and in that center, you actually have assay

13   sensitivity demonstrated for the active control,

14   but you also have separation of gepirone from

15   placebo in that particular center.

16   Now, as a whole, that study is difficult to

17   interpret. But among the active control studies,

18   that particular center of that particular study

19   actually had the clearest demonstration of assay

20   sensitivity post hoc or exploratory as it might

21   otherwise be.

22   This summarizes some characteristics of the

1   respective studies. You can see that the four

2   studies that were early terminated have low power.

3   You can see that the 04 and the 06 studies have a

4   mean baseline HAMD-17 that's actually less than 20

5   because about 50 percent of their patients had a

6   baseline HAMD that was less than 20.

7   That was a major deficiency of these

8   particular studies in terms of identifying the

9   extent to which they replicate the findings of 01

10   and 07. Again, if the key question is, are the

11   findings from 01 and 07 trustworthy or are they due

12   to chance, when we consider other studies like 04

13   and 06, we need to do that for the same population

14   that 01 and 07 had in them. And there's some other

15   summary information here as well.

16   This gives, again, a summary description of

17   the seven 2-arm studies, as well as some comments

18   about what they showed. We've already noted these

19   previously. This essentially gives a summary of

20   the five 3-arm studies, and in particular, noting

21   how their primary analysis, as specified in the

22   protocol, did indeed have p-values greater than

1   0.05 for the assay sensitivity assessment.

2   Again, for the 053 study, if you use a more

3   correct method of analysis, the p-value would

4   become 0.05. That more correct analysis

5   corresponds to omitting the treatment by center

6   interaction from the primary model as recommended

7   in ICH-E9 guidance.

8   053 study I think was actually analyzed, and

9   planned for analysis when it was traditional to

10   include treatment by center in a model regardless

11   of its particular significance. 053 is the one

12   that comes the closest to demonstrating assay

13   sensitivity.

14   These are the five studies that are

15   considered interpretable together with a summary of

16   their estimates of effect size, which have been

17   noted to be in the clinically relevant range, as

18   well as their p-values both for the baseline

19   adjusted analysis as well as the protocol-defined

20   analysis.

21   Now, we move to the meta-analyses. This is

22   the first meta-analysis. It works with the five

1   interpretable studies. In the lower right-hand

2   corner, there's the p-value for judging

3   homogeneity. It's 0.26. This meta-analysis

4   supports the five studies being comparable or

5   compatible with homogeneity.

6   The overall effect size is -1.3 because

7   we've averaged the effect sizes from the two

8   positive studies in with the supportive studies and

9   the study that has been agreed upon as negative.

10   You can see that the weights for the five studies

11   are reasonably similar here because their sample

12   sizes were reasonably similar, although the

13   greatest weight is actually given to the 02 study.

14   The overall p-value is 0.002.

15   The next thing that was done in order to get

16   a better sense of clinical interpretability was to

17   look at a responder variable. This responder

18   variable is 50 percent reduction in the HAMD-17,

19   and it's looked at as a dichotomy in an odds ratio

20   metric. The overall odds ratio was 1.4, and that

21   corresponds to roughly a risk difference in the

22   range of 8 to 9 percent or an NNT in the vicinity

1   of 11 or 12.

2   Again, the assessment of heterogeneity has a

3   p-value of 0.30, so there indeed is reasonable

4   compatibility with homogeneity in this analysis

5   even if one recognizes limitations in statistical

6   power for these assessments.

7   More importantly, as an additional

8   sensitivity analysis, the responder analysis was

9   repeated in which all of the patients who

10   discontinued or dropped out were classified as non-

11   responders.

12   In this particular analysis, a responder was

13   the patient who had a 50 percent reduction in

14   HAMD-17 and they had to complete follow-up. Any

15   patient who dropped out is managed as a

16   non-responder in this analysis.

17   In terms of issues about missing data, in

18   this particular responder analysis, we're using an

19   estimand that is accounting for the fact that

20   dropouts actually are not eligible to be a

21   responder. So this analysis is consistent with the

22   way such data might be dealt with in a more modern

way today.

Again, the odds ratio actually is 1.4. The
p-value for the homogeneity assessment is a little
bit more border line. But still the odds ratio of
1.4 fits, again, with an NNT in the vicinity of 11
or 12. The overall p-value is 0.009, which again
is essentially supporting overall efficacy when all
five studies are included.

Now, the next study that was added was 053
because, as I've mentioned previously, in my
opinion, 053 actually demonstrated assay
sensitivity. Its only limitation was it was
prematurely terminated. But as was indicated in
the first presentation, that premature termination
was apparently because that other sponsor had
positive studies for another product and chose to
move on with the development of the other product.

Essentially, there were external reasons why
these studies were prematurely terminated. They
weren't necessarily prematurely terminated because
of the data. I think it, therefore, is reasonable
to include them in these integrated analyses. Now,

1   the overall p-value is at 0.001, and the p-value

2   for homogeneity is 0.36.

3   If we move further and add in 078 and 083,

4   which are two other studies that were early

5   terminated, again, the p-value stays at 0.001. The

6   homogeneity assessment is 0.555. Again, there is

7   compatibility with homogeneity in this analysis.

8   So again, when we consider these eight studies,

9   what we see overall is a positive result. We see

10   nothing that seems to be contradictory to what was

11   shown as the efficacy in 01 and 07 from a

12   sensitivity analysis point of view.

13   Then if we move to nine studies and

14   additionally include 017, the p-value becomes

15   0.009, homogeneity 0.201. Again, this analysis,

16   when we are additionally adding the 017 study, is

17   comparable.

18   If we include all studies except 04 and

19   06 -- and in this meta-analysis, we're essentially

20   considering all studies that essentially have a

21   population, which is like the population that was

22   in 01 and 07, the p-value is 0.009.

1   Finally, if we move to all 12 studies but we

2   invoke the restriction that HAMD-17 must be greater

3   than or equal to 20 because that was the enrollment

4   criterion for 01 and 07 -- so again, we're using

5   this as a sensitivity tool to address whether the

6   findings in 01 and 07 are due to chance.

7   We want to look at those studies that are as

8   much like 01 and 07 as possible, also taking the

9   view that we don't want to prematurely dismiss a

10   particular study because it had early termination

11   or perhaps other issues. Here, the p-value does

12   increase to 0.02 and the heterogeneity again is

13   0.225.

14   Now, I also requested our statistical

15   colleagues to do another type of analysis. You've

16   heard something about looking at p-values and

17   classifying them as to whether the one-sided

18   p-value was below 0.025 or not as a way of then

19   identifying the role of chance and classifying how

20   many studies had p less than 0.025 one-sided.

21   But this ignores the actual magnitude of the

22   p-value. If you're concerned about integrating

1   p-values as opposed to effect sizes as I've

2   mentioned earlier, RA Fisher, a prominent

3   statistician in the history of statistics, actually

4   proposed a method for combining independent

5   p-values.

6   The way it works is you take minus 2 times

7   the log of the p-value. That has a chi-square

8   distribution with 2 degrees of freedom. If you add

9   that up across a set of studies, that will have

10   2 degrees of freedom times the number of studies.

11   It's a well-established method for combining

12   p-values across studies. So the bottom row of this

13   table indicates what those Fisher combined p-values

14   are for the set of 5 studies, the set of 6 and

15   others, including all 12 studies regardless of the

16   restriction of baseline HAMD-17 greater than or

17   equal to 20.

18   This analysis is useful because it

19   recognizes that a p-value of 0.01 is more

20   interesting than of 0.024. It also recognizes that

21   a p-value of 0.075 is more useful than a p-value of

22   0.95. It takes the actual magnitude of the

1   p-values into account in the integration.

2   It also weights all the studies

3   equally -- it does that -- whereas the previous

4   meta-analyses weighted them on a precision basis by

5   inverse-variance. But regardless of which of these

6   you use, the efficacy is reinforced.

7   In summary, meta-analysis of the primary

8   efficacy parameter and responder rates for the five

9   interpretable studies were clearly statistically

10   significant in favor of gepirone ER. The effect

11   for responder rates was maintained when all

12   dropouts were considered as non-responders for the

13   five interpretable studies, a more pessimistic

14   assumption than that used in the original analyses.

15   Meta-analyses for the five interpretable

16   studies also show a clinically meaningful effect

17   size for responder rate for gepirone, basically NNT

18   in the vicinity of 11 or 12. The scope of more

19   inclusive sensitivity meta-analyses provide

20   complimentary support for the findings from the

21   individual interpretable studies that gepirone ER

22   is effective in the treatment of MDD.

1   Thank you for your attention. The next

2   speaker.

3   **Industry Presentation – Stephen Stahl**

4   DR. STAHL: Good morning. I am Steve Stahl,

5   and I've been compensated for my time and travel.

6   However, I have no financial interest in the

7   outcome of this meeting.

8   I am a professor of psychiatry at UCSD for

9   almost 30 years and have been interested in

10   serotonin for almost 40, dating back to my PhD

11   dissertation on a serotonin transporter. I've

12   actually been interested in 5-HT 1A receptors for

13   over 20 years, including one of the books that I've

14   written just on that topic.

15   I'm perhaps better known for my textbooks,

16   which I'm going to share a cartoon or two with you

17   in a minute, the number 1 and number 2 best-selling

18   textbooks in psychiatry for the past 20 years.

19   I've been asked to do two things: one, a

20   couple of slides on mechanism of action; and

21   number 2, look at this not from a statistical or

22   regulatory point of view but from a clinical point

1   of view. I'm a clinical psychiatrist. I see

2   patients, so I am bringing that perspective. But

3   I'm also a psychopharmacologist. We'll start with

4   that.

5   What's an overview of gepirone? It's a

6   so-called azapirone. There's an additional

7   azapirone on the U.S. market already for a

8   different indication and another one actually on

9   the Japanese market for another indication.

10   This one is metabolized to a 3-hydroxy

11   metabolite, which actually is twice as much in the

12   plasma as the parent compound. When you give this

13   drug, the biological activity is due to the

14   combination of its 3-hydroxy and its parent

15   compound together.

16   The far right-hand part of the formula there

17   is clipped off. It's called 1PP, and it's about

18   half the plasma level of either the active or the

19   parent compound. And it has some activity, but

20   more of an alpha 2 antagonist, but not very potent

21   on serotonin. And actually, its potency on alpha 2

22   is about two orders of magnitude less.

1   This 3-hydroxy metabolite and its parent

2   acts specifically and uniquely at the 5-HT 1A

3   receptor, which I'll show you in a table later, is

4   unique for this drug and no other drug has that

5   property.

6   It specifically doesn't affect 5-HT 2A

7   receptors. 5-HT, as you know, is 5-hydroxy

8   tryptamine. It's a fancy way of talking about

9   serotonin, so I'll use those interchangeably.

10   There are actually 14 known subtypes of serotonin

11   receptors and gepirone only acts as one of them.

12   5-HT 2A receptor notably is linked to various

13   adverse effects such as sexual dysfunction,

14   insomnia, agitation, et cetera, and this does not

15   interact with that receptor.

16   Not only did the pharmacology become more

17   specific than any known agent on the market to just

18   the 1A receptor, it actually was reformulated.

19   This is not trivial. Sometimes these are seen as

20   convenience things for once-a-day or patent

21   extension gimmicks, but this is neither in the

22   gepirone case.

1   What it is, is that if you have a rapid

2   onset of Cmax, which is your concentration that's

3   the highest and it's fast, you don't have

4   tolerability. And that's what's the problem of the

5   immediate-release formulation of this drug and the

6   two other azapirones actually that are on the

7   market. It reduces the tolerability, so therefore

8   reduces the dose. And then, of course, that can

9   reduce the efficacy.

10   Here it is. Pharmacokinetics are shown with

11   the intermediate release in the blue compared to

12   the controlled-release ER, extended-release, in the

13   orange. If you slam the receptor by a rapid rise

14   of the blue and then you rapidly come off, and then

15   you're quiet for a long time and then slam it again

16   12 hours later, you will maximize the potential for

17   side effects, and this decreases tolerability quite

18   a bit.

19   What the ER has done is it more gently rises

20   in the orange, and it doesn't go as high, and it

21   stays there longer. This also is a property to

22   actually desensitize receptors more robustly and

also to bathe them, if you will, in the drug for

more parts of the day, which seems to enhance

efficacy at least in animal models.

Here's a cartoon to compare and contrast

what might be different about gepirone. We'll

start with the world famous SSRI serotonin and

selective reuptake inhibitors. And here you've got

a cartoon of a serotonin neuron. On the left, you

have only 5-HT 1A receptors, and this is one of the

targets of gepirone by the way. On the right, you

have 14 receptors on the post-synaptic phase.

Now, the model of depression could be that

sometimes in some patients, for whom we do not yet

have biomarkers, the 5-HT 1A receptors could be

dysregulated, and the serotonin could be also

dysregulated perhaps low.

This is a common so-called monoamine

hypothesis. It's been around for 50 or 60 years.

It's oversimplified. There's many other forms of

depression perhaps related to glutamate and

inflammation and all sorts of other things. But it

is seemingly true for at least some patients, so

1   it's come to this formulation.

2   Now, if you look at the rest of the diagram,

3   you see some yellow pieces of serotonin floating

4   around on the left and little on the right. You

5   see blue circles, which are the reuptake pump, and

6   some lightning or, in other words, firing of the

7   neuron inside.

8   The first thing you do is if you put an SSRI

9   into the reuptake pump, you will actually

10   interestingly, almost immediately, block the 5-HT

11   reuptake both on the left and on the right, and

12   you'll raise serotonin, though, over here

13   immediately.

14   You can show this with dialysis probes in an

15   animal. But even with some neuro imaging, which I

16   know some members of the panel do in this area, you

17   can show that you can get 5-HT going up, but only

18   in the left part, and you do not have an

19   antidepressant effect yet.

20   So how do you get an antidepressant effect?

21   We believe that what happens is that these

22   receptors on the left desensitize, down-regulate.

1   So here they are. What are those receptors

2   normally there for? They're there as breaks, or as

3   inhibitory, or, if you will, auto-receptors. When

4   serotonin goes in to those auto-receptors, it

5   normally turns off further release. It's one of

6   the ways the neuron controls its own serotonin

7   release.

8   What happens is they desensitize and go

9   away. Guess what that does? If you don't have any

10   breaks any more, you've cut the break cable, look

11   at the lightning in the neuron, it will go up, and

12   then you get further release with the delay on the

13   far right. It's thought that that release of

14   serotonin in the synapse on the right occurs in the

15   same time course as recovery from depression.

16   Does this happen in all patients? We don't

17   know, but probably not. When it doesn't occur,

18   could this account for treatment resistance in some

19   patients? Possibly, but we don't know that.

20   That's under investigation, but it's certainly a

21   possibility. But when it does occur, it goes into

22   all the receptors.

1   Now, what's interesting, for example, when

2   you stimulate a 5-HT 2A receptor and a 5-HT 1A

3   receptor, you cancel out many of the downstream

4   effects. There's actually a whole bunch of other

5   receptors there. For example, 5-HT 2A receptor

6   stimulation causes sexual dysfunction and other

7   things. The 5-HT 3 receptor can cause nausea,

8   et cetera. Obviously, the 5-HT 1A receptor, as

9   I'll show you in a minute, can cause dizziness and

10   nausea in its own right.

11   You give this portfolio of outcome with what

12   I call mixed downstream effects because what

13   happens is that you have whatever net-net occurs

14   after all these receptors are stimulated.

15   Sometimes when I lecture, I say it's like

16   dunking the brain into a bucket of serotonin

17   because you're actually stimulating every receptor

18   and every field all of the brain. Some of that is

19   good because you could have an antidepressant

20   effect if you're so fortunate to have that kind of

21   depression, but you could also have side effects,

22   cost of doing business.

1   What about gepirone? Well, it's an agonist.

2   On the left and on the right of that neuron is a

3   partial agonist. If you care at the presynaptic

4   5-HT 1A auto-receptors, those are the guys on the

5   left -- but since it's a partial agonist, it

6   doesn't down regulate them at least in animal

7   models as robustly and not therefore

8   surprising -- it results in less 5-HT in that

9   synapse. But it, meanwhile, is acting as a full

10   agonist at the 5-HT receptor on the right.

11   Net-net, this results in far more stimulation of

12   the 5-HT 1A receptor and far less stimulation of

13   the other guys.

14   Here, you've got the cartoon of gepirone,

15   same neuron, same serotonin, same receptors. But

16   now you have gepirone going into the left and into

17   the right. And what happens is that eventually on

18   the left, these receptors desensitize, and you'll

19   get some increase of serotonin, but you will

20   actually have a predominance of 5-HT 1A receptors

21   being stimulated on the right because there's less

22   serotonin. If you dose the drug right, you'll get

1   one heck of a lot of 5-HT 1A-driven downstream

2   events.

3   Well, this is really kind of critical

4   because if you do not dose this drug adequately, it

5   won't work. I'll actually show you a dose response

6   curve. When you go into post hoc analyses of some

7   of these other studies and you don't give this drug

8   long enough or high enough, you might as well not

9   give it because it won't work. It's not fair to

10   have those sorts of things compared unless there's

11   adequate dosing.

12   The other thing is that the downstream

13   events, when you just stimulate the 5-HT 1A

14   receptor is that you get dopamine being increased

15   downstream. This does not happen with the SSRIs.

16   That's thought to be a favorable thing for certain

17   symptoms but also for mitigating certain side

18   effects such as sexual dysfunction.

19   Also, finally, one of the things that

20   gepirone has, at least, a theory of being able to

21   work at is it doesn't require any serotonin to

22   work. It can go right into the 5-HT 1A receptor as

a sort of artificial serotonin itself. In patients

who don't have enough serotonin, if you give them a

reuptake blocker, nothing happens because it's

blocking the reuptake of released serotonin. If

you ain't got no serotonin, you're not going to

have any enhancement of that.

These are other possible patients that don't

respond to SSRIs. Some of them may have that. And

also, some people who we call "poop out" in which

people take a medicine and they stop having a

response to it, possibly due to depletion of 5-HT

in those people. You need a mechanism that can

bypass the serotonin need.

Now, whether this is actually true or not,

it's promise. Whether this is something that's

more an article of faith at this point or hope as

opposed to being demonstrated, this relates to a

belief that we all have as clinicians in this field

that different people respond differently.

If you ask a room full of clinicians, as I

have many times, have any of you ever seen a

patient respond to one antidepressant and not in

1   another, you get all the hands going up. If you

2   ask them if they've seen a patient who tolerates

3   one antidepressant and not another, all the hands

4   go up.

5   Who that will happen in will require further

6   studies if and when this drug is reproved. But it

7   certainly doesn't work like the other drugs, so it

8   has at least the possibility that it will work in

9   different patients.

10   This is the final pharmacology slide, and it

11   shows you that gepirone is a 5-HT 1A partial and

12   full agonist. Buspirone, its friend from the past,

13   not well known by a lot of people, is actually a D2

14   antagonist. In fact, here in town at the NIDA,

15   Nora Volkow is using buspirone as a D3 PET

16   ligand as far as I know, and also for substance

17   abuse things.

18   As you may know, buspirone was actually

19   developed -- it didn't go very far -- as an

20   anti-schizophrenia drug. So buspirone actually has

21   some neuroleptic-like properties, which create

22   problems.

1   The problem with buspirone is it can't be

2   dosed very high because it's not tolerable. It's

3   got off-site actions at D2 receptors. It's

4   immediate release, and so it tends to be

5   underdosed. Despite that, you may recall in the

6   same study that Mike Thase told you about, the

7   STAR*D study, buspirone was one of the best second

8   treatments in augmentation, which could give you a

9   hint that gepirone could go in the direction of

10   either switching from SSRI non-responders or

11   augmenting partial responders to SSRIs going

12   forward.

13   Now, the SSRIs are the only one in that

14   whole list there that don't have 5-HT 1A. I could

15   have put a "/SNRI" there, but then I would have had

16   to put not just SRI serotonin reuptake inhibition

17   but also NRI norepinephrine reuptake inhibition.

18   The antidepressants, vilazodone and

19   vortioxetine actually share both properties.

20   They're both 5-HT 1A NSRIs. The interesting thing

21   about those two drugs is they're dosed so that only

22   about 50 percent of 5-HT 1A and 50 percent of

1   transporters are occupied.

2   You'll see that we don't actually

3   have -- other than calculations, but gepirone, when

4   dosed at around 60-80 milligrams is occupying about

5   70-80 percent of 5-HT 1A receptors. Buspirone, in

6   its normal doses, occupies 30 percent. Everything

7   else in that list is probably around 50 percent.

8   When an SSRI is dosed, it blocks almost

9   100 percent of SSRI pumps, but these two new drugs

10   block about 50 percent. It's sort of like the two

11   new drugs are trading in their 5-HT reuptake

12   blockade for a little bit of 5-HT 1A, and they have

13   net-net perhaps the same efficacy but maybe better

14   tolerability.

15   Finally, the last four drugs are drugs that

16   are approved for one or more forms of depression.

17   But as you know, these are also drugs that are

18   D2 antagonists; they have off-site and off-target

19   actions of 5-HT receptors and other neuro

20   transmitters.

21   Some of that could be good to contribute to

22   efficacy, because as I said, they do not have as

1   robust of 5-HT receptor occupancy as gepirone. But

2   these drugs also cause weight gain. They cause

3   problems with metabolics and akathisia and other

4   things that gepirone does not.

5   That kind of, in a nutshell, tells you that

6   this is unique pharmacology. And the way you start

7   with a drug with unique pharmacology is show

8   ordinary pharmacology, which is you have to have

9   studies that work. Then once you've done that, you

10   can try to differentiate it. So the first step is

11   what we're here to talk about, clinical experience

12   of efficacy and safety.

13   Again, you've seen this table before. It's

14   the five interpretable studies. Again, I'm just a

15   country doc. I'm told which ones are interpretable

16   and which ones aren't. To me, this seems to make a

17   lot of sense.

18   Remember, these are 5 of the 7 studies that

19   are placebo-controlled, and I'm going to go through

20   those a little bit from a clinician's point of

21   view. Here, you've got -- all of them are

22   double-blind, placebo-controlled, flexible dose,

multicenter.

Vanilla ice cream; no Häagen-Dazs here. This is just a very interesting matter, to severe major depression with a HAMD of -- we talked about in the earlier presentation of 20 or more, diagnosed by a clinical interview. Also did what's called the Montgomery Asberg Depression Rating Scale for one of the studies and also a secondary feature of them.

As you know, some studies used MADRS as primary and the Hamilton as secondary and flop it around. This one used the Hamilton 17 as the primary, as you know.

Treatment duration was 8 weeks except for one study was 9, except this is also a bit of a lie. This could come up when you try to do post hoc retrospective analyses of assay sensitivity.

If you take a drug that takes two weeks to work and two weeks to be titrated before you actually can work it, and compare it to a drug that takes eight weeks and it's always at therapeutic levels, this is not a fair comparison.

1   I'm not sure any of the assay sensitivities

2   have ever been done with that in mind, and I remind

3   you that iloperidone had a similar flaw and also

4   compared studies where a drug was given as a

5   comparator for eight weeks that worked against it

6   and it took a few weeks to titrate.

7   I just bring that up. This drug won't work

8   if you don't have it at the right dose. If you

9   compare it to drugs that work longer than it, it's

10   not fair.

11   The primary efficacy endpoints are mean

12   change from baseline and total HAMD as you heard

13   many times. Secondaries are here. One of the

14   important ones, of course, is depressed mood. As

15   you know, major depression is not just depressed

16   mood. It's a syndrome that has other things like

17   sleep abnormalities, energy, guilt, et cetera.

18   But the depressed mood is an important item

19   of course. It has phase validity as well. The

20   MADRS is another way. It will score depression.

21   The clinical global impression, you see the

22   severity on the first set of changes in baseline on

1   the bottom. The CGI improvement, you can actually

2   use that. It's sort of a Likert scale I'll talk

3   about in another slide.

4   Another secondary way of looking at this is

5   with the responder remitter analysis, which

6   Dr. Koch told you a little bit about. You can do

7   it with the HAMD-17 or the HAMD-25. You can look

8   at the CGI or the remissions, which is different.

9   What's a responder? You got to be

10   50 percent better or more. If you're a remitter,

11   it's supposed to be essentially not mentally ill

12   but what is defined in practice as a HAMD of 8 or

13   less.

14   Let's take a look at the severity scales.

15   Max Hamilton invented about 50, 60 years ago now

16   the 17-item questionnaire used to rate the severity

17   of depression in the tricyclic era. It's long

18   considered the gold standard since then, although

19   Stuart Montgomery and Marie Asberg may disagree

20   though because the MADRS has come in as a new one.

21   A score on the Hamilton 17 of 20 is a

22   prerequisite of these clinical trials, and in

1   retrospect, perhaps one of the markers of

2   responsivity to serotonergic drug is a HAMD of 24

3   or 25, not 20. But we didn't do that, so it

4   probably is good enough, but it could've been

5   strengthened perhaps if the patients had even

6   higher scores.

7   Two-point reduction on the HAMD is

8   considered clinically relevant, and the most

9   important item is item 1. MADRS is only 10 items.

10   It's also very similar to the HAMD, but it's

11   designed to be more sensitive to the effects of

12   antidepressants, although I'm not sure Max Hamilton

13   will agree with that. But the score of 25-30,

14   because it's scored differently, is a prerequisite

15   for entry of trials, and a 2 and a half point

16   reduction is considered clinically relevant.

17   Finally, the CGI, clinical global

18   impression, is a 7-point scale. It's sort of a

19   clinical, just like it says, impression; you know,

20   somebody that adds everything all up and tries to

21   say exactly, not with a rating scale but with all

22   information available, in their judgment is the

1   patient's severity or improvement getting better.

2   Now, how are these five interpretable

3   studies all designed? Here it is. Drug versus

4   placebo, there was not an internal comparator in

5   these five. Everybody got 20 milligrams for 3

6   days; and on the fourth day, there were forced

7   titration to 40 milligrams. And at day 7, they had

8   the option to go to 60. In retrospect, it probably

9   was an error not to force them to 60.

10   Forty doesn't work very well. I can show

11   you that, and it may not even work at all. There's

12   a real question as to whether 40 is the minimally

13   effectively dose or the null effect dose, but it's

14   right on the corner.

15   But even though you could keep people at 40,

16   in practice, only about 10 or 15 percent were

17   at 40. Almost everybody on day 7 went to 60, and

18   actually half of the people that were at 60 then

19   went to 80. You'll see the doses at the end.

20   It's a good thing they went to 60 or 80

21   because that's what works. Then you can see that

22   an effective dose is not attained until probably

1   two weeks into the studies, at least a week.

2   That's the way this was designed, eight weeks in

3   most of the studies.

4   There was not a lot of post-treatment

5   assessment to the extent it was anecdotally done.

6   There was actually no observation of a lot of

7   withdrawal effects, but it hasn't been formally

8   looked at.

9   What is the dose response to the effect that

10   we have it? If you look at all the five

11   interpretable studies, plus the two others that

12   were just placebo-controlled, you'll see a dose

13   response curve. These were most robust. If you

14   took 80 -- there was actually two of the studies

15   that went above 80, the so-called 073 and 083,

16   which some people don't think are interpretable.

17   I'm just trying to tell you that this is

18   called pharmacology, is a dose response curve.

19   Mostly, we have dose response curves for side

20   effects in this field and not so much for efficacy.

21   You could say that if you don't have -- the lower

22   dose is, probably the longer you have to treat to

1 give it a good chance of working.

2   If you look at old studies, we had a lot of

3 dropouts before two weeks of treatment at 60.

4 That's just not fair because there's no chance that

5 those people would have gotten better. We just

6 know that from pharmacology.

7   Here, they are. Now, ladies and gentlemen,

8 I think that's a drug. It's called two positive

9 studies with the predesignated endpoints, positive

10 in both cases, 001 and 007. They separated at

11 endpoint when they were getting 74.7 milligrams on

12 the left and 71.7 on right at the endpoint.

13   Although it says it separates at week 3 on

14 the one on the left and week 4 on the one on right,

15 that's really separating at week 1 on the left

16 after adequate dosing and week 2 on the right.

17 Both of these were separating at endpoint.

18   What about secondary variables? You've got

19 the top part of the table, which talks about the

20 HAMD item 1, the MADRS, CGI, all positive for the

21 one study, 001 and two out of the three positive

22 for 007.

1   Look at the responder remitter analysis on

2   the bottom. Two of the four were positive for the

3   first one on the left, and all were positive on the

4   right. These supported what that -- so that's a

5   very positive study.

6   How about these two? They missed. But they

7   also are somewhat supportive if you're willing to

8   look at that. The one on the left had no trend

9   except maybe the orange line is numerically better

10 than the blue line. But the blue line is very

11 much -- too much of the placebo response, which

12 probably killed the efficacy in that study.

13   There are actually two points on this study

14 on the right, on week 2 and week 6 that separated,

15 but unfortunately for the drug, it lost its

16 separation at week 8. There is some supportive

17 evidence and directionality and earlier onset in

18 these studies.

19   If you look at the secondary variables for

20 that, you see that the study on the left, which

21 didn't separate one bit, actually did on the HAMD

22 item 1, and almost hit it on the MADRS and did hit

1   it on one of the responders.

2   The one on the right, which separated early

3   but not late, still, even at endpoint, had some

4   separation on the HAMD-25 responder. So we would

5   consider these directional and supportive.

6   This is the fifth of the five studies,

7   complete dud, no separation on the left, no

8   separation on secondaries on the right.

9   As you've heard many times -- we keep

10   hammering this -- gepirone ER meets the standard of

11   effectiveness in MDD by having two well-controlled

12   studies, both which hit its primary. Of course,

13   the ones that hit the primary hit almost all their

14   secondaries.

15   Of the other three, there's a directional

16   trend on the primary and definitely certain

17   secondaries that were positive. And one was a

18   complete a dud with Dr. Koch's meta-analysis

19   showing that both of these favorable.

20   Remember that all five of these studies are

21   placebo-controlled and did not have an active

22   comparator arm. Just like some of the newest drugs

1   being approved by the FDA, the comparator studies

2   were not included in the package of positive

3   studies.

4   How big is this? It's normal. In other

5   words, the purple shows you the range of HAMD-17's

6   drug placebo differences at endpoint that are

7   existing for the last zillion antidepressants out

8   there, between 1 and a half and 3.8 points. And

9   the two positive studies are not trivial. They

10   have effect sizes that are even greater than

11   2 points right in the middle of the curve.

12   If you look at the MADRS, it's the same

13   thing. If you say 2 and a half points is what

14   you'd like to see or greater, both of these were

15   greater, 3.3 of the first one and 4.2 on the right.

16   So these are meaningful, and I think everybody

17   agrees on that.

18   In summary, I think two studies show

19   substantial evidence of effectiveness in treatment

20   of major depression, and these achieve statistical

21   significance on the primary and also on supportive.

22   Effect size was clinically meaningful in which

1   you'd expect, and there's additional supportive

2   evidence in two of the other studies, so four of

3   the five had some evidence of efficacy.

4   Safety. This will be quick because there's

5   really no safety issues with this drug. But just

6   to tell you so, you can weigh safety against

7   efficacy. There are 5800 patients included in the

8   safety database, almost 5000 in phases 2 and 3.

9   The safety population was from 19 various

10   controlled phase 2 -- and I'll show you those -- in

11   3 ER, extended-release studies, with almost 2000

12   receiving this drug, 1275 placebo and several other

13   comparators receiving a certain amount depending on

14   the study size.

15   Overall, the safety profile of gepirone is

16   consistent with that of other antidepressants.

17   It's even more consistent out of its mechanism, and

18   it wouldn't be surprising, which is that it's a

19   5-HT 1A agent and therefore has 5-HT 1A side

20   effects.

21   If you want details of this, we can show

22   you, but the withdrawals through the serious AEs

1   were 2 and a half percent in gepirone and

2   1.2 percent in placebo but were not thought be drug

3   related. Rate of withdrawal to any AEs is about

4   16 percent in gepirone and 7 percent in placebo.

5   Not shown here are the withdrawals due to

6   lack of efficacy in which that would be flipped.

7   There were five deaths, none drug related in either

8   placebo or the two -- with gepirone or two with

9   placebo.

10   Here they are. What are the problems with

11   gepirone? They had to be dizziness and nausea,

12   some headache, insomnia, and fatigue. This is a

13   well-known mechanism-related side effect. But the

14   interesting thing about dizziness is that's really

15   not presyncope dizziness. It's a form of

16   wooziness, or light-headedness, or

17   fuzzy-headedness. It's a kind of unique way of

18   looking at it.

19   Other kinds of nausea for the SSRIs would be

20   the 5-HT 3 kind of nausea, which is well-known to

21   be linked to cancer chemotherapy and so forth.

22   There's at least a couple of mechanisms of nausea

1   and dizziness. This drug has it, of course,

2   related to its mechanism.

3   They were typically mild to moderate and

4   usually lasted a few hours after dosing and for

5   only a couple of days. People got tolerant to it.

6   It would come back if you'd raised the dose, but

7   new dizziness and nausea eventually declined

8   towards the placebo rate after a couple of weeks.

9   Interestingly enough, you get tolerant to

10   this drug about the same time course as it works.

11   So the desensitization of some receptors leading to

12   tolerance is occurring at the same time; not

13   surprisingly that the desensitization of other

14   receptors is linked to efficacy.

15   Dizziness was dose-related but nausea was

16   not particularly. None were coded as serious, and

17   there was no syncope or fainting because that's not

18   the mechanism of this dizziness. Two and a half

19   percent of gepirone patients and half a percent of

20   the placebo patients withdrew due to dizziness.

21   I have two slides of AEs reported by greater

22   than 5 percent in any treatment group. These are

1   not statistical comparisons. These are safety

2   comparisons. These are numerical.

3   You can see a little bit of the orange

4   bar -- the blue bar on headache, more so for

5   dizziness, a smidge for somnolence, but really more

6   for nausea. You can see the other kinds of side

7   effects, some of which are shared by the SSRIs, and

8   of course some of the anticholinergic things, which

9   are prominent for imipramine.

10   Continuing this on the second slide, you'll

11   see some insomnia. But what you also interestingly

12   see is a little bit of a reduction in

13   sexual-related adverse experiences in the whole

14   portfolio of studies; whereas, as Dr. Thase showed

15   you, it's well-known that SSRIs have treatment-

16   emergent sexual dysfunction, that wasn't the

17   purpose of this study, but it was observed here not

18   unsurprisingly.

19   Actually, this gives you a hint of where

20   this drug could go, again, with the properly

21   designed study to look at whether it either has

22   promotable advantages and less sexual dysfunction

1   from treatment-emergence. There's even the

2   possibility that this could mitigate the sexual

3   dysfunction of other drugs when added on, as has

4   been seen anecdotally. That's not a claim. You

5   can't read that out of these data, but this is

6   where the mechanism could be exploited later.

7   I think gepirone ER is safe and

8   well-tolerated. The human safety data for ER is

9   extensive, long-term exposure more than 1500

10   subjects has not turned up any new safety concerns,

11   almost 700 for a half a year and over 100 for a

12   year as required. It's well-tolerated with

13   dizziness the most common, which is typical of 1A

14   mechanism. And it's got a low risk of sexual

15   adverse experiences comparable to placebo.

16   Overall conclusions. Two studies achieved

17   statistical significance on the predesignated

18   primary efficacy variable and nearly all of its

19   secondary efficacy variables for those studies.

20   Effect sizes in those studies is clinically

21   relevant and comparable to every other approved

22   antidepressant. Long-term studies, exposing more

1   than 1500 subjects, have demonstrated that gepirone

2   ER is well-tolerated.

3   **Industry Presentation – Daniel Burch**

4   DR. BURCH: Thank you. Daniel Burch with

5   Pharmaceutical Product Development again. I wanted

6   to go ahead and quickly conclude. I think you've

7   heard from the speakers, Dr. Thase in particular,

8   that the unmet need for antidepressants is high and

9   new classes are needed.

10   I'll just go ahead and state the obvious.

11   Gepirone is going to work in some patients and is

12   not going to work in others just like all other

13   antidepressants.

14   As you heard from Dr. Stahl, the safety and

15   tolerability is in keeping with the generally

16   prescribed antidepressant. You heard about the

17   mechanism. You also heard Dr. Stahl argue from a

18   clinical point of view that studies 001 and 007 are

19   robustly positive and alone meet substantial

20   evidence of efficacy. They met their primary

21   endpoints.

22   They showed a reasonable effect size, very

1   much in keeping with what's been approved by the

2   Division of Psychiatric Products, and they were

3   internally consistent on the secondary endpoints.

4   As I asserted earlier, they had good internal

5   metrics in terms of reliability and assay

6   sensitivity.

7   Separately, Dr. Koch talked about the

8   meta-analyses and demonstrated to us that the

9   findings in 001 and 007 just didn't occur by

10   chance. That would be exceedingly unlikely.

11   We submit that this has met a fairly high

12   evidentiary standard and we think this argues for

13   the approval of gepirone. Obviously, we think the

14   risk-benefit profile is favorable. Thank you.

15   **Clarifying Questions to Industry**

16   DR. D'AGOSTINO: I want to thank the sponsor

17   for the detailed and enlightening presentation. We

18   are now going to look at the questions for the

19   panel, the clarifying questions from the panel.

20   I want to ask are there any clarifying

21   questions that we want to address to the sponsor.

22   When you do have the questions, please state your

1   name for the record before you speak. If you can,

2   please direct your questions to a specific

3   presenter.

4   I'm going to start by asking if Victor has

5   any questions. Being not present here, I won't see

6   his hand go up. But for the rest, I'd like you to

7   raise your hand, and we will get your name down and

8   go around a circle and hopefully address everybody

9   before we have our break.

10   Victor, do you have any questions?

11   Dr. De Grutolla? Is he still on the line?

12   DR. DE GRUTTOLA: Yes. Can you hear me?

13   DR. D'AGOSTINO: Yes.

14   DR. DE GRUTTOLA: I had a question for

15   Dr. Koch. I understand that he did a sensitivity

16   analysis where non-completers were considered

17   non-responders. I just wanted to ask what the

18   rates were of the non-completers in the study and

19   were they differential between the active treatment

20   and placebo?

21   DR. BURCH: I think that question was to

22   Dr. Koch, correct?

1   DR. DE GRUTTOLA: Yes.

2   DR. BURCH: Okay.

3   DR. KOCH: Could we have the core slide 27?

4   Maybe 37. Sorry my eyes aren't -- yes. Sorry.

5   I think that this slide in the columns

6   marked percent dropout answer the question. For 01

7   and 02 for the gepirone, it's 27 and a half and

8   31.8 percent, and for placebo 23.6 and 28.7

9   percent. For 07 and 08, they're smaller, 21.8,

10   24.0 versus 17.8 and 21.5; 023, 26 and 21. So

11   they're in the 25 to 30 percent range, and the

12   gepirone rate of dropout I think is somewhat

13   higher.

14   DR. DE GRUTTOLA: I have a follow-up

15   question. I understand you did an analysis that

16   treats the non-completers as non-responders. Did

17   you do any other kind of modeling for what could be

18   informative non-completion?

19   DR. KOCH: No. The extent to which analyses

20   could be done was very limited. Involvement with

21   respect to this project didn't really begin until

22   October. It required a lot of extra effort to get

1 access to the data. The early studies were done in

2 the early 1990s. The other studies were done, I

3 think, prior to 2007. There really weren't

4 resources to go back and do extensive model-based

5 analyses.

6   The responder variable that we considered

7 was essentially considered to address an estimand

8 where a good outcome was completion of the

9 follow-up period. And in doing that, you achieved

10 a 50 percent reduction in the HAMD-17, which was

11 your primary endpoint.

12   This was not a prespecified analysis. Of

13 course, in protocols written today, it would be.

14 But at the time these studies were written, that

15 kind of responder variable was not really

16 considered. It was to help interpret the HAMD-17

17 in a more useful way in the modern environment, is

18 the reason why it was done.

19   But it was not really possible to do more

20 extensive model-based analyses in view of the fact

21 that all of these analyses were done since October,

22 and there were limitations on how much could be

1   done. But I think it is a meaningful estimand. It

2   is a good outcome to complete the follow-up period

3   and achieve a 50 percent reduction in the HAMD, but

4   we're limited with that.

5   DR. D'AGOSTINO: Dr. Compagni?

6   DR. COMPAGNI-PORTIS: Yes. Just a follow-up

7   on that, I have a couple of questions. Do you know

8   a reason for the great number of dropouts in the

9   various studies?

10   DR. KOCH: I think others can speak to that.

11   I think it was given in the briefing book, so we

12   would need to be pulling up tables from the

13   briefing book for each of the respective studies.

14   But I believe the briefing books do identify which

15   percent of dropouts were for lack of efficacy and

16   which percent were for adverse events.

17   The previous presentation summarized it for

18   adverse events, and I've given you what the total

19   dropouts were. I apologize for not being able to

20   do a subtraction, but probably the difference

21   between the adverse event dropouts and the total

22   dropouts would be either for lack of efficacy or

1   things that are allegedly lack of efficacy, like

2   just dropping out without stating a reason, which

3   most of the time is interpreted as lack of

4   efficacy.

5   DR. COMPAGNI-PORTIS: Thank you. My other

6   question is I know that it was said that the

7   gepirone hasn't been approved in any other country.

8   Has it been presented and not approved in any other

9   country?

10   DR. BURCH: I'll take that question. No, it

11   has not.

12   DR. FOLLMANN: Dean --

13   DR. D'AGOSTINO: Please say your name before

14   you speak.

15   DR. BURCH: I'm sorry. This is Daniel

16   Burch.

17   DR. D'AGOSTINO: I'm sorry. Do you still

18   have another question?

19   DR. COMPAGNI-PORTIS: Sorry. One last

20   question. I know that some of the early studies

21   were looking at anxiety, and I know one of the

22   adverse events was insomnia. Do you know what the

1   impact on anxiety was for patients? Also, did you

2   include data on whether patients were in

3   psychotherapy? Then I really will stop with my

4   question.

5   DR. BURCH: I'll ask Dr. Stahl to take that

6   question.

7   DR. STAHL: As far as I know, it wasn't

8   properly studied or nor was it even a secondary

9   efficacy variable like the HAM-A often is done with

10   the HAMD. Of course, this class of drugs is

11   well-known to do that. It's approved in Japan for

12   anxiety with tandospirone. It's approved in the

13   United States with buspirone for that. Good

14   question, but I don't think it has been properly

15   studied.

16   DR. COMPAGNI-PORTIS: [Inaudible.]

17   DR. STAHL: I didn't catch that one. Could

18   you repeat it?

19   DR. COMPAGNI-PORTIS: Do you know how many

20   people were also in psychotherapy during the

21   studies, in any of the studies?

22   DR. STAHL: Correct me if I'm wrong, but I

1   believe that they were not allowed to be in

2   psychotherapy during the study.

3   DR. D'AGOSTINO: Dean?

4   DR. FOLLMANN: Yes. This is Dean Follmann.

5   I have a couple of questions. The first is sort of

6   prompted by what I just heard. The dropouts for

7   lack of efficacy, I guess, they're not showing

8   benefit on the drug or placebo, so they drop out.

9   And then you quit measuring them, you quit

10   following them, but often you're obliged to

11   use -- I guess last observation carried forward as

12   the endpoint determination?

13   DR. BURCH: Dr. Koch is going to respond to

14   that.

15   DR. KOCH: Yes, that's my understanding,

16   although I'll probably look at Mary Johnson for

17   correcting it. But my understanding, again, these

18   studies were done a number of years ago. A number

19   of them were done in the early 1990s. Others were

20   done between 2000 and 2006, or thereabouts.

21   I think the corresponding review division

22   found last observation carried forward a tolerable

1   method at that particular time. They may have

2   found it tolerable because continued follow-up on

3   active treatment hypothetically might have let

4   those patients do better had they been followed

5   longer with those on placebo doing worse, so you

6   possibly get an under-estimate of effect size.

7   It's definitely recognized you get an

8   under-estimate of variance when you do LOCF, but

9   that was the prespecified method.

10   In order to adjust for that for the

11 presentation in 2015, we defined the responder

12 variable to be 50 percent improvement, and you had

13 to complete the study with all dropouts classified

14 as failure. That was our effort to approach 2015

15 in spite of the architecture of these studies and

16 their analysis plan 10, to 15, to 20 years ago.

17   DR. FOLLMANN: Okay. My second question has

18 to do with a meta-analysis. You showed a

19 meta-analysis using all 12 of the studies, which I

20 like because it's using sort of all of the data.

21 You have a p-value on that of 0.021.

22   In the sponsor's material, Mary Johnson also

1   did a meta-analysis of the 12 studies and had a

2   p-value of about 0.09. I was wondering if you or

3   Mary could comment on the difference in those two

4   p-values. I assume you used different methods to

5   come up with different p-values.

6   DR. KOCH: My understanding is the p-value

7   in the vicinity of 0.09 includes the 04 and 06

8   studies with respect to all of their patients. The

9   meta-analysis that I showed in slide --

10   DR. FOLLMANN: 48.

11   DR. KOCH: -- 48 only focuses on the

12   patients who have the HAMD-17 greater than or equal

13   to 20. In doing that, it only includes about half

14   of the patients from 04 and 06 because 04 and 06

15   were the least compatible studies with the

16   population enrolled in 01 and 07.

17   Now, the meta-analysis I showed in the next

18   slide, which is 49 -- if we go to 49 -- which uses

19   the Fisher method. So we go to 49 -- am I able to

20   do that? Yes, okay. Let's see, 49.

21   Yes, so when I used the Fisher combined

22   method, then the p-values are 0.012 when it's all

1 12 studies and 0.009 when it's restricted to the

2 greater than or equal to 20. But you're correct.

3 If you do the meta-analysis based on effect sizes,

4 then it's all 12 studies regardless of this

5 restriction on baseline, the overall p-value is

6 around 0.10.

7   I believe the heterogeneity p-value also is

8 around 0.10 suggesting more heterogeneity as one

9 might expect when one includes those additional

10 studies that are less of a good match to the other

11 studies.

12   DR. FOLLMANN: All right. Then my final

13 question is to Daniel Burch concerning slide 11.

14 This might be a slide for the FDA actually. But

15 you show a lot of approved drugs that have an

16 enormous number of failed studies.

17   I guess I'm wondering how they could

18 approved. And I guess they were approved because

19 you -- they also included in their packets studies

20 with just two arms, the new drug and the placebo,

21 and those studies tended to show good success.

22 This is a sort of noteworthy to me, and I was

1   wondering if the FDA or you could comment on how

2   these drugs were approved.

3   DR. D'AGOSTINO: Maybe we should hold that

4   off and --

5   DR. FOLLMANN: Fine.

6   DR. D'AGOSTINO: -- let the -- that's a good

7   question, but maybe to bring -- I want to make sure

8   the sponsor gets their time to respond to

9   questions.

10   DR. FOLLMANN: Sure.

11   DR. D'AGOSTINO: Dr. Pickar?

12   DR. PICKAR: Thank you. Just to follow up

13   the last observation, LOCF carried forward,

14   Michael, you discussed how placebo response is

15   really hindering our ability to demonstrate

16   efficacy, sensitivity, or whatever you want to call

17   it.

18   I wonder how much that LOCF is contributing

19   to that. When someone drops out and they're not

20   feeling well, that's usually not gotten in the

21   assay at the time they dropped out because it's a

22   numeric thing that's carried off as their last

observation. It may actually reflect what their

clinical state is.

In psychosis, it's a huge problem, and I

thought we were moving away from LOCF but perhaps

not. Is there a way look at it, Mike? And

Dr. Koch, it's simply completers, and does that

help you or does not invoke -- but Mike, I can't

help but think that's a big problem.

DR. THASE: It's Michael Thase still from

the University of Pennsylvania. In newer studies,

we stay away from the LOCF because of that bias.

DR. PICKAR: Okay.

DR. THASE: Generally, studies that use

modern imputation methods or MMRM are somewhat more

positive to medications than LOCF is. My favorite

with access to all the data would actually be

time-to-benefit or time-to-response where people

are censored as they dropout.

But really, they all should converge if a

drug really works. There are somewhat different

levels of precision and somewhat different errors,

but it shouldn't just hinge on whether you have an

1   LOCF or MMRM. But yes, these studies were done in

2   the era in which LOCF was king or queen, and they

3   carry that penalty.

4   Particularly, in those four early studies,

5   which were stopped for nonclinical reasons, the

6   dropout rates are higher because some patients were

7   actually in week 4 of their treatment trial when

8   the sponsor pulled the plug.

9   DR. PICKAR: Okay. One other quick question

10   to Steve. Did I misunderstand you? Gepirone at

11   dose ranges has alpha 2 antagonist properties? Did

12   I hear that? It's interesting because you

13   emphasize how important it is to get it at a full

14   dose, and I'm wondering if you're getting

15   overlapped outside the serotonin induced and

16   adrenergic processes.

17   DR. STAHL: I don't know if you can quickly

18   get it. We have a binding table in the back --

19   DR. PICKAR: I'd like to see that. Sorry

20   it's so tedious, I know.

21   DR. STAHL: -- in which it shows the 1PP

22   binding. Gepirone itself carries no alpha 2.

1   DR. PICKAR: It's a metabolite, is that

2   right?

3   DR. STAHL: It's the metabolite. But as I

4   recall, the KIs or the PKIs are like 8 for 5-HT 1A

5   and like 6.9 -- no, it's not that one, but it's in

6   that area -- alpha 2 for MPP. But the plasma

7   levels are also one-third as high, so you end up

8   having 1 and a half-fold more 3-hydroxy than

9   parent. So basically, you've got -- add those

10 together, right?

11   DR. PICKAR: Yes.

12   DR. STAHL: Because that's active 5-HT 1A.

13 And then you've got 10 fold to 20 fold less 1PP at

14 10 to 20 fold less potency. You will have a

15 shoulder -- there it is.

16   So you see the gepirone has a PKI of 7.42 on

17 5-HT 1A and so does its 3-hydroxy metabolite. And

18 as I said, I can show you elsewhere, but the PK

19 shows that those two are much higher than the 1PP

20 on the right. So you've got an order of magnitude

21 or so on 5-HT 1A. There should be another one on

22 alpha 2.

1   DR. PICKAR: I didn't see that.

2   DR. STAHL: There should be another slide

3   for alpha 2 some place there.

4   DR. PICKAR: Do you know if there's any data

5   on plasma norepinephrine levels? Does this drug

6   increase plasma norepi?

7   DR. STAHL: Plasma norepi, I don't know.

8   Its microdialysis mostly has been shown on

9   dopamine.

10   DR. PICKAR: In all the clinical studies, is

11   there any biological measures that went on during

12   the studies at all to help us, any academic

13   studies, norepi increases? I'm just curious

14   because it's interesting because it's overlapping

15   to an alpha 2, which is a whole other game.

16   DR. STAHL: Here we go. Here we go. This

17   is the slide. You see the PKs for 1PP is almost 7

18   for alpha 2A and 2B and a little less for 2C. This

19   is about tenfold less than the PK for 5-HT 1A for

20   gepirone. So therefore, you're going to have a

21   little bit of alpha 2 in net-net.

22   DR. THASE: Michael Thase. The best

1 biologic assays in people during clinical trials

2 are pulse and diastolic blood pressure, and then

3 the incidence of side effects such as constipation

4 or sweating. And there was just no signal -- no

5 effect on diastolic blood pressure, no effect

6 on --

7 DR. PICKAR: There have been no biological

8 measures during --

9 DR. THASE: To my knowledge, no.

10 DR. PICKAR: Okay.

11 DR. D'AGOSTINO: Dr. Narendran?

12 DR. NARENDRAN: This is probably a question

13 from Dr. Stahl. I know there's no human PET

14 studies with 5-HT 1A receptor binding. Was there

15 any non-human primate data or anything that could

16 estimate what the occupancy is at 60 and 80 versus

17 lower doses?

18 DR. STAHL: The question I've asked myself

19 without an answer. All you can do is on the back

20 of an envelope where you can -- in fact, they don't

21 really even have good estimates of what 40, 60, and

22 80 Cmaxes are. There have to be imputed from the

1 normal volunteers studies that were done at lower

2 doses.

3   If you play that game, you can get yourself

4 to believe that 80 milligrams is around 70 percent

5 occupancy at Cmax and maybe about 50 percent

6 occupancy at trough. That was for 60, and 80

7 milligrams is a little bit more. But that is a

8 calculation, so there's no good studies to directly

9 answer your -- even in animals.

10   DR. D'AGOSTINO: Dr. Mann?

11   DR. MANN: Thank you. I have a couple of

12 questions for Dr. Koch. I infer from what you were

13 saying that you had a limited amount of time to

14 work with the data set. But as a statistician,

15 given that variables like baseline severity, dose,

16 duration of treatment, time-to-response,

17 alternative response criteria and so on are all of

18 interest, what do you feel about the idea of having

19 done a patient level analysis instead of trying to

20 lump together all of these studies?

21   DR. KOCH: Well --

22   DR. MANN: I have a second question

1   unrelated to this.

2   DR. KOCH: Okay. My understanding is that

3   when the meta-analyses were done initially,

4   particularly those that used the sponsor's

5   prespecified method of analysis, we used estimates

6   of effect size and standard errors from various

7   tables that were available. Is that correct, Mary?

8   MS. JOHNSON: Yes.

9   DR. KOCH: Ultimately, we did get access to

10   the data. The data files did become available to

11   us. And when we did the ANCOVA analyses, which we

12   tended to emphasize in what I reported today, we

13   had access to the data, and those ANCOVA analyses

14   were done.

15   Now, they were done on a last observation

16   carried forward file to produce those results. We

17   didn't do any kind of other data management. To

18   some extent, we moved in that direction as we

19   learned more about what was in the FDA briefing

20   book to sort of harmonize what we were doing in

21   reporting as much as possible with what the FDA

22   briefing book had been doing and emphasizing.

1   Am I correct there, Mary? Okay. She's

2   nodding. Thank you.

3   DR. D'AGOSTINO: Dr. Gogtay?

4   DR. MANN: I actually have a second

5   question. I understand your answer. You didn't

6   actually -- in those slides that you just showed us

7   in your talk, you didn't actually present person

8   level analysis data. They were study level

9   meta-analyses; is that right?

10   DR. KOCH: Well, for each study analysis, a

11 person level data produced an estimate of effect

12 size --

13   DR. MANN: I understand that. But the whole

14 point of the question is, wouldn't it be more

15 interesting to have combined person level data

16 across all of the studies?

17   DR. KOCH: Well, we did indeed do that but,

18 again, we emphasized the five studies first, and

19 then we added other studies to it. Ultimately, we

20 had the 12 studies for the patients with the

21 baseline severity of greater than or equal to 20.

22   Now, those analyses used person level data

1   to produce the estimate of effect size and standard

2   error. Instead of doing a computing process that

3   had a data structure with all 12 studies and

4   fitting a statistical model that had study

5   baseline, baseline by study treatment, treatment by

6   study, and then essentially producing an estimate

7   statement that averaged across, we didn't do it

8   that way.

9   We simply did a by study analysis at person

10 level data. So they are person level data analyses

11 to get the effect sizes, but then they are

12 aggregated in a usual meta-analytic way by

13 weighting the studies by the inverse variance,

14 which would typically be like what a person level

15 analysis would seek to do as well.

16   DR. MANN: Okay. I understand that you

17 didn't present the person level data pooling all of

18 the studies. Now, I have a study level question.

19   In table 1, in the FDA presentation, I'd

20 like you to comment on table 1, 12 short-term

21 studies with HAMD-17 results. I wondered whether

22 you looked at the four studies that are in the

1   middle of this table in which the division finds

2   negative finding and the sponsor finds the studies

3   to be failed.

4   Did you attempt to replicate the FDA

5   findings and try and resolve, to some degree

6   statistically, why your conclusions are different?

7   DR. KOCH: With respect to study --

8   DR. MANN: I can give you the numbers of the

9   studies if that's helpful.

10   DR. D'AGOSTINO: Are you able --

11   DR. KOCH: Are you putting the slide up --

12   DR. D'AGOSTINO: Gary, are you ready to

13 address that question now or --

14   DR. KOCH: Yes. With respect to the 04 and

15 the 06 studies, we regard them as failures on assay

16 sensitivity because they did not demonstrate assay

17 sensitivity on their prespecified primary endpoint,

18 and that is the same with 17. But the

19 meta-analyses I reported towards the end.

20 Particularly the one for all 12 studies in the

21 patients with baseline greater than or equal to 20,

22 all 12 studies were included.

1   I did point out in my presentation that I

2   actually consider study 53 to be a study in which

3   assay sensitivity was demonstrated when, like the

4   FDA, you use a statistical model that does not

5   include treatment by center.

6   I also have noted that study 53 is actually

7   partially a positive study because one of its

8   centers actually completed enrollment prior to when

9   the premature termination occurred. If you look

10   only at the data for that center, both the active

11   control and gepirone separate from placebo.

12   But our analyses, when we work with all 12

13   studies, do indeed include all 12 studies with

14   baseline HAMD greater than or equal to 20 when

15   we're looking at effect size. And when I use the

16   Fisher method in the last slide, it looked at all

17   12 studies regardless of whether the baseline HAMD

18   was greater than or equal.

19   Our analyses are equivalent to analyses on

20   patient level data, but the style was statistical

21   computing, was patient level data within studies in

22   an inverse variance weighting across the studies of

1 the estimated effect sizes.

2 DR. MANN: So where the FDA has focused on

3 the comparator drug in all of these four studies, I

4 have a showing that the comparator drug beat the

5 placebo or the comparator drug beat gepirone. Did

6 you actually repeat any of those analyses to see

7 whether you agreed or disagreed?

8 DR. KOCH: No, I did not have either the

9 resources or the time to repeat that. But as I

10 said, I agree that the active comparator beat

11 placebo in the 053 study on the prespecified

12 primary endpoint of HAMD-17. The active comparator

13 did not beat placebo in any way whatsoever on the

14 prespecified primary endpoint of HAMD-25 in 04 and

15 06 or in MADRS on 017.

16 It was exploratory analyses, albeit logical

17 ones to do as part of due diligence, that

18 identified a possible suggestion of assay

19 sensitivity for 017 and perhaps 04 and 06 as well.

20 But I didn't replicate those analyses. I took

21 those analyses at the face values that the FDA

22 reported them as.

1   Mary may have a comment to add.

2   MS. JOHNSON: Hi. Mary Johnson, statistical

3   consultant to Fabre-Kramer. Regarding your

4   question about patient level data, we did have

5   access to patient level data. Basically, I was

6   preparing analysis of covariance models for

7   Dr. Koch to implement the meta-analysis.

8   In the process of doing that, I tried to

9   replicate the FDA's p-values comparing active

10 control to placebo, test drug to placebo, and I was

11 able to match those p-values. However, that was

12 not the pre-planned analysis. Most of those

13 studies were basically comparing the active control

14 to placebo only if the test drug was significant.

15 They had these sequential rules in place to avoid

16 multiplicity, so they were not really a full 3-arm

17 analysis, which the FDA had done post hoc.

18   The other point you made was about the

19 meta-analysis, did we ever do a patient level

20 meta-analysis. Yes, we did an analysis of

21 covariance looking at study as a main effect,

22 treatment as a main effect, and looking at

1   treatment by study interaction. We combined five,

2   eight -- you know, all the various combinations of

3   studies.

4   The results were very significant, favoring

5   gepirone versus placebo in those meta-analyses.

6   But I think Dr. Koch was more comfortable using a

7   more classic meta-analysis, where he's looking at

8   the weighted combination of treatment effects

9   across the studies, which is you're not

10   interpreting as a patient level analysis.

11   But if you did a patient level analysis,

12   ignoring center and just blocking by study, the

13   test of treatment effect of gepirone versus placebo

14   was highly significant in the pool of 12 studies.

15   DR. MANN: Thank you.

16   DR. D'AGOSTINO: Do you have further

17   questions?

18   (No audible response.)

19   DR. D'AGOSTINO: I'd like to break now.

20   We've gone over in terms of the amount of time we

21   were going to direct questions to the sponsors. We

22   will have the opportunity in the afternoon to pick

1   up those questions. Those who are being asked to

2   hold their questions, we will come back to you.

3   I do want to say this very last discussion,

4   I think it was one of my questions, I think it's

5   very important to have a sense of what would you

6   have gotten if you looked at that sub-sample,

7   Dr. Koch's analysis is on. If the sponsor can come

8   up with the answer for that, I think that would be

9   one of the things we'd want to look at here about

10   later on.

11   Right now, we're going to take a 10-minute

12   break because of the extensive questions, very good

13   questions. For the panel members, please remember

14   that there should be no discussion of the meeting

15   topics during the break amongst yourselves or any

16   member of the audience. We'll resume in 10

17   minutes, which is 10:44.

18   (Whereupon, at 10:33 a.m., a recess was

19   taken.)

20   DR. D'AGOSTINO: We were thinking during the

21   break that the questions seem to have a lot of

22   relation to each other, and it might be useful to

1   take a couple of hands that were still not answered

2   or addressed the question.

3   Judith, would you like to ask a question?

4   DR. GOLDBERG: It's for Gary.

5   DR. D'AGOSTINO: Gary is behind you.

6   DR. GOLDBERG: It was more about the

7   meta-analyses. I mean, did you at any point -- you

8   used the inverse variance weighting, so basically

9   what you were -- that's the fixed effect model.

10   Did you do any work with the random effects models,

11   which generally increased the variability so that

12   you might've gotten a different result? Have you

13   done that or thought about it?

14   Have you done any random effects models for

15   the meta-analyses, and if so, what were your

16   results?

17   DR. KOCH: Both fixed and random effects

18   analyses were done in parallel. I think for the

19   five studies, the meta-analyses for both the fixed

20   and random effects analyses were reported in the

21   briefing book of the sponsor, or at least that's my

22   recollection.

1   DR. GOLDBERG: I don't recall.

2   DR. KOCH: Now, what happens is I tried to

3   show the assessments of homogeneity tended to have

4   fairly large p-values greater than 0.20 for the

5   majority of the analyses, which then essentially

6   supports the use of the fixed effects analysis.

7   Now, the random effects paradigm tends to

8   assume the studies you have are like a random

9   sample of potential studies, but really these

10   studies were very different. They were the early

11   Bristol-Myers studies. They were the five studies

12   that focused on responding to FDA advice, and then

13   there were the three other studies.

14   Also, the random effects analysis needs a

15   sufficiently large sample size in terms of number

16   of studies for the confidence intervals to make

17   sense because you're assuming a particular

18   distributional structure when you do that.

19   By and large, the findings from the random

20   effects models were comparable to those from the

21   fixed effects analyses. The random effects

22   analyses had slightly wider confidence intervals

1   and slightly bigger p-values, but the conclusions

2   were generally similar.

3   We emphasize the fixed effects analyses

4   because these studies were done as a fixed set of

5   studies. And further, the assessments of

6   homogeneity supported the use of the fixed effects

7   analyses.

8   DR. GOLDBERG: Okay. Thank you. That's it

9   for now.

10   DR. D'AGOSTINO: Dr. Gogtay?

11   DR. GOGTAY: I have a few questions. The

12   first one, knowing very well that I'm not a

13   statistician, but I'm curious that if you were to

14   remove those first two studies from the

15   meta-analysis, would any of the results survive?

16   DR. KOCH: I don't know because the purpose

17   of the meta-analysis was not to focus on an

18   integrated analysis substituting for two pivotal

19   studies. My understanding of the regulatory review

20   process is that two positive studies are needed,

21   and in this particular paradigm, it's what do the

22   other studies take away from the two positive

1   studies.

2   Now, I suppose that if you were to do some

3   sort of meta-analysis of the other 10 studies, I

4   don't know what you would find. You probably would

5   find something either border line or slightly

6   trending. I think the majority of studies did have

7   point estimates in the predicted direction.

8   You might've gotten a border line trend of

9   0.09, 0.15, I have no idea. But that wasn't the

10   question we were doing. We're not using

11   meta-analysis as a proof of efficacy. The proof of

12   efficacy is the two studies, the 01 and the 07.

13   We're using meta-analysis as a sensitivity

14   analysis tool to identify how much do the other

15   studies take away from the usual conclusion that

16   one would draw on the two studies, and that was

17   what we did. And we did it by starting with the

18   studies that were most comparable to the two

19   studies first, and then adding in all of the other

20   bodies of information until we got to all 12

21   studies, albeit with the restriction that baseline

22   HAMD-17 had to be greater than or equal to 20

1   because that's the feature of the two positive

2   studies.

3   DR. GOGTAY: There was a mention of two

4   negative pediatric trials. I somehow must've

5   missed that while reading the brochure. I'm

6   curious as to what the results of those were, why

7   they were not included in this meta-analysis or

8   these 12 studies, and do we have any data on those?

9   DR. BURCH: Okay. I will ask Dr. Thase to

10 respond to that.

11   DR. THASE: I understand that they were not

12 included in the specific package and were not

13 included in our presentation because this was

14 submission for adults. The studies did fail or

15 they were negative since there wasn't an active

16 comparator.

17   The internals looked good. The studies were

18 considered failed or negative studies. We can make

19 the results available to you. I don't have a slide

20 to put up to show you.

21   DR. GOGTAY: Do you know the doses used for

22 the --

1   DR. THASE: I'm sorry?

2   DR. GOGTAY: The medication dose used, do we

3   have any idea about -- are you able to --

4   DR. THASE: I assume the dose was up

5   titrated to 60 to 80 milligrams, although there may

6   have been a downward adjustment.

7   I am not a pediatric psychiatrist and make

8   no claims. I have some sense of the data. You

9   know, 80 plus percent of pediatric depression

10   studies fail to show assay sensitivity or failed to

11   show significant drug placebo differences.

12   Gentlemen, are there two or three approved

13   antidepressants for pediatrics? Two?

14   DR. MATHIS: We have two.

15   DR. THASE: Two, whereas we have 18, 19 for

16   adults, only two have passed muster in peds.

17   DR. GOGTAY: A couple more follow-up

18   questions. On the dose of the other studies, I

19   think Dr. Stahl showed the graph in slide, I think,

20   69 where the dose separation at 60, or 60 onwards.

21   I was curious to see if there was doses in

22   the following or subsequent graphs, which I did not

1   see them on the graphs. So I was curious to see

2   what the dose points were on those if they were

3   available.

4   DR. STAHL: I don't know off the top of my

5   head to answer to your question. Steve Stahl. I

6   think one of Dr. Koch's slides actually shows the

7   mean dose. Do you have that one?

8   DR. KOCH: Slide 37 that I showed earlier on

9   dropout rates.

10   DR. STAHL: Does that answer your question?

11   DR. GOGTAY: No, because I was curious to

12 see -- the trending separation that you showed on

13 the subsequent graphs, I was curious to see what

14 dose level did that happen in those?

15   DR. STAHL: It's very interesting that

16 study 078 is a partial answer to your question.

17 That was one of those prematurely terminated

18 studies. It turns out, for reasons that escape me,

19 they actually did two doses against placebo, but

20 the primary predesignated outcome was the

21 combination of those two arms together.

22   If they actually just took 078 high-dose

1   arm, it worked. There's actually evidence

2   that -- and that would be, in that case, 60 to 80

3   as opposed to 40 to 60. So there is evidence in

4   that specific study of a dose response curve.

5   DR. GOGTAY: If I could ask one more thing

6   along those lines, I think at some point I'd read

7   in the brochure that there was a decision made to

8   skip the regular dosing, I mean regular formulation

9   and go to the ER? And I was unclear about the

10   rationale for that. What was the reason for

11   dropping the regular formulation?

12   DR. STAHL: I wasn't there at the time, but

13   it's not tolerable is the short answer to your

14   question. It's twice a day; it's not tolerable.

15   They had many more dropouts --

16   DR. GOGTAY: Not tolerable because of side

17   effects?

18   DR. STAHL: Yes.

19   DR. GOGTAY: Or because of dosing

20   inconvenience?

21   DR. STAHL: Well, in fact, you can't get the

22   dose. In fact, as I recall it, of all the studies

1   that were done with the IR, only one of them was

2   over 40. All the ones below 40 didn't work. There

3   are supportive data if you want to go on to the IR.

4   That gets messy. But there is a positive IR study.

5   It was the one study that had a dose over 40.

6   All the other studies were less than 40 and

7   they didn't work. That's part of the rationale for

8   them when they went to the ER. Also, the dizziness

9   and the nausea, this is typical of this group of

10   drugs. It's also why another drug called

11   ipsapirone never went forward, although it's got

12   the positive evidence in phase 2 for depression

13   because it just had too many peak dose side

14   effects.

15   DR. GOGTAY: One final question. Are there

16   any ongoing trials of this?

17   DR. STAHL: As far as I know, no.

18   DR. GOGTAY: Okay.

19   DR. D'AGOSTINO: Dr. Rudorfer?

20   DR. RUDORFER: Yes. Thank you.

21   Matthew Rudorfer. A question perhaps for Dr. Burch

22   just about the recruitment methods. Could you say

1   something about where the subjects came from and

2   were any treatment-seeking, were they answering

3   advertising, or was there any differences across

4   the studies?

5   DR. BURCH: Yes, these studies, most of the

6   studies that were -- this is Daniel Burch of

7   course. Most of the studies were done earlier this

8   millennium, I guess, in the early 2000s.

9   My understanding was that they were done

10   with investigators that had their own databases, so

11   patients that had their own clinical practices. As

12   Dr. Stahl said, they were sort of plain vanilla

13   patients with a HAMD severity threshold that the

14   investigators themselves actually adjudicated.

15   Does that answer your question well enough

16   at this point?

17   DR. RUDORFER: Yes.

18   DR. D'AGOSTINO: Thank you. We're going to

19   now continue and move on to the FDA presentations.

20   It'll be Dr. Yang, Mathis, Temple, and LaVange.

21   **FDA Presentation – Peling Yang**

22   DR. YANG: Thank you. I will provide some

1   highlights on the issues regarding the evaluation

2   of effectiveness.

3   First, a brief history of this NDA. The

4   application was first submitted in 1999, but FDA

5   refused to file. This submission included 16

6   trials, of which 6 were with ER formulation and 10

7   with IR formulation.

8   The sponsor focused on 4 trials only. One

9   of them was a short-term ER trial. This trial was

10   conducted at two sites, and the sponsor intended to

11   use the results of a single site to support

12   efficacy. The other three trials were with IR

13   formulation. Two were short-term and one long-

14   term.

15   In 2001, the NDA was resubmitted. It

16   included two more ER trials. But in the initial

17   submission, the focus was on four trials, one ER

18   trial but different from the initial submission,

19   and three IR trials, same as in the initial

20   submission.

21   ODE1 agreed that the ER trial was positive,

22   but concluded that only one of the three IR trials

reached statistical significance. This IR trial

was a small atypical depression trial, and the

primary endpoint was HAMD-17.

ODE1 initially was concerned about the

characterization of the population studied because

the mean baseline HAMD score was only 13 to

14 units, roughly 10 units lower than the usual

scores, but subsequently concluded that this study

could be considered supportive to provide

additional support of efficacy.

Despite that, the preponderance of negative

trials in the overall development program remained

as a concern although many of the trials employed

relatively low doses. ODE1 requested an additional

positive ER trial to demonstrate efficacy, and that

the trial should be a fixed-dose trial.

In 2003, the second resubmission was

received. It included a randomized withdrawal

study with ER formulation instead of a short-term

trial required in the not-approvable letter.

ODE1 concluded that this study was negative.

By then, there were a total of 19 trials with only

1   two positive short-term ER trials, one with two

2   positive short-term trials, one was ER, and the

3   other was IR. At the same time, another ER trial

4   was unblinded, and the results were not positive.

5   ODE1 asked for at least a robustly positive short-

6   term trial that was a fixed-dose study and a

7   positive randomized withdrawal study.

8   In 2007, the third resubmission was

9   received. It included a second positive ER trial

10   along with various meta-analyses of 12 short-term

11   ER trials. It also included post hoc analysis

12   results of the same maintenance study as in the

13   second resubmission.

14   Although there were two positive ER trials

15   by then, ODE1 concluded lack of substantial

16   evidence for efficacy and noted that the magnitudes

17   of observed treatment effects among those ER trials

18   were very small. But this was not the reason for

19   the non-approvable action.

20   In 2011, sponsor requested ODE1 for

21   reconsideration. A type C meeting was held. In

22   response to sponsor's argument, ODE1 clarified a

1   few misunderstandings from the sponsor and

2   suggested that sponsor provide additional

3   justification to support their argument. In 2012,

4   sponsor submitted an amendment in support of

5   informal appeal. After thorough review, ODE1 still

6   concluded lack of substantial evidence for

7   efficacy.

8   Early this year, Office of New Drugs headed

9   by Dr. Jenkins, accepted sponsor's formal dispute

10 resolution request. The main issue is on the

11 substantial evidence for efficacy. Dr. LaVange,

12 director of Office of Biostatistics, was consulted.

13 This AC meeting was called for to seek input from

14 the committee.

15   This slide summarizes the 13 ER trials, 12

16 short-term and 1 long-term. By the time of the

17 last review cycle, there was agreement on the two

18 positive studies as in the top green block,

19 studies 07 and 01, and three uninterpretable

20 studies in the pink block.

21   In the white block are the three negative

22 studies, but the sponsor considered two of them

1   supportive. In the yellow block are the four

2   short-term trials in dispute. They were 04, 06,

3   17, and 53.

4   In the orange block at the bottom is the

5   long-term maintenance study 09. For this trial,

6   sponsor acknowledged that it was negative but noted

7   that the results would have been positive if the

8   trial were designed and conducted appropriately.

9   Here is a brief summary of the five

10   interpretable studies agreed upon. The top two

11   trials were positive. The daily dose ranges for

12   these trials were 20 to 80 milligrams, and the

13   prespecified primary endpoint was HAMD-17.

14   The third column is the HAMD-17 results, LS

15   mean difference, that is, the treatment effect

16   relative to placebo. A negative difference

17   indicates gepirone is better than placebo.

18   For the two positive trials, the magnitudes

19   of observed treatment effects were around 2.5

20   units. For the other three studies, the magnitudes

21   were relatively small. In study 23 at the bottom,

22   the observed effect was numerically in favor of

1   placebo.

2   This is a brief summary of the three

3   uninterpretable studies in consensus. These trials

4   investigated different dose ranges. Study 52

5   included fluoxetine as an active comparator. The

6   magnitudes of observed treatment effects were very

7   small in these studies.

8   This slide is the four trials in dispute.

9   Each of these trials includes an active comparator,

10   fluoxetine for studies 04 and 17, paroxetine for

11   study 06, imipramine for study 53. Studies 04 and

12   06 enrolled patients with atypical depression. The

13   protocol specified primary endpoint was HAMD-25;

14   it's HAMD-17 plus eight additional items assessing

15   reversed vegetative symptoms.

16   In study 17, the protocol specified primary

17   endpoint was MADRS. In study 53, there were two

18   co-primary endpoints, HAMD-17 and CGI, in the sense

19   that gepirone needed to beat placebo on both

20   endpoints to demonstrate efficacy.

21   This slide summarizes the review team's

22   efficacy results of the four trials in dispute

1   based on HAMD-17 regardless of the protocol

2   specified primary endpoint. The results are

3   divided into two panels. The left panel summarizes

4   the mean changes from baseline for each arm, that

5   is, the drug and placebo responses. A more

6   negative change indicates more improvement.

7   For the two atypical depression trials, the

8   top two, the observed drug and placebo responses

9   were generally smaller compared with the bottom two

10   trials.

11   The right panel summarizes the pairwise

12   comparison results. A negative difference

13   indicates the former is more effective than the

14   latter. The first column suggests that gepirone

15   was numerically worse than placebo in three of the

16   trials. The second column suggests that active

17   comparator was numerically better than placebo in

18   all four trials, and two of them, studies 06 and

19   53, reached nominal statistical significance level

20   of 0.05.

21   The third column suggests that gepirone was

22   numerically worse than active comparator in all

1   four trials, and three of them, except study 53,

2   reached nominal statistical significance. However,

3   based on sponsor's prespecified primary endpoint

4   and analysis, none of the comparisons reached

5   statistical significance.

6   What caused the disparity in conclusions

7   between ODE1 and sponsor? There were two major

8   factors. Number 1, definition of assay

9   sensitivity. In ODE1's view, the trial has assay

10   sensitivity if active comparator beats placebo or

11   test drug. But in sponsor's view, assay

12   sensitivity should not be judged based on

13   comparisons between any two drugs.

14   Number 2, efficacy endpoints used with

15   determining outcomes. HAMD-25 was the primary

16   endpoint in the two atypical depression trials,

17   MADRS in study 17. Nonetheless, ODE1 used HAMD-17

18   for each trial to assess the totality of evidence

19   for efficacy.

20   What was ODE1's rationale for using HAMD-17?

21   HAMD-17 was the protocol specified primary endpoint

22   in 9 of the 12 trials. It was a reasonable

1   endpoint to use as a common metric for each trial

2   in the overall assessment.

3   For the two atypical depression trials, the

4   primary endpoint was HAMD-25. It adds eight

5   additional items that measure reversed vegetative

6   symptoms. The review team noted that HAMD-25 was a

7   secondary endpoint in the two positive studies, 01

8   and 07. This endpoint was also statistically

9   significant in both positive studies.

10   To compare the baseline HAMD scores between

11   the two positive studies and the two atypical

12   depression studies, the review team explored the

13   distributions of baseline HAMD-25 total score,

14   HAMD-17 total, and the sum of the eight item

15   scores. Of the eight items, five measure atypical

16   features, so the review team also explored the

17   total of the five item scores at baseline. These

18   distributions were compared among the four studies.

19   The results suggest that the patient

20   populations were generally comparable among the

21   four trials. The review team felt that any

22   depression rating scales commonly used in clinical

1   trials should be sensitive to show an

2   antidepressant effect.

3   For each trial, the sponsor deemed

4   uninterpretable, sponsor listed possible reasons

5   for trial failure. While the reasons pertain to

6   the two atypical depression trials, the sponsor

7   believes that fluoxetine and paroxetine should not

8   have been chosen as the active comparators because

9   of their unknown effects in atypical depression.

10   If so, we wondered why they were included as active

11   comparators in these trials.

12   In addition, if these active comparators

13   were ignored, the two trials will still be

14   considered negative because gepirone did not beat

15   placebo on either endpoint.

16   Sponsor considers high placebo response as

17   another reason for trial failure in studies 04, 06

18   and 17. In ODE1's view, the observe placebo

19   responses in these studies were generally

20   comparable with those observed in the two positive

21   studies.

22   High placebo response was commonly observed

1   in depression trials. Based on our depression

2   trials database, the average placebo response on

3   HAMD scale was around 8.2 with a standard deviation

4   of around 1.8. The observed placebo responses in

5   these three trials do not appear to be extreme

6   compared with those observed in our database.

7   Study 53 was one of the four studies in

8   dispute. This study consisted of two sites, site A

9   enrolling 123 patients and site B, 47 patients

10   only. This trial was included in the sponsor's

11   initial NDA submission to support efficacy based on

12   the results from site A despite overall negative

13   findings.

14   The sponsor attributed trial failure to

15   several reasons such as early termination leading

16   to a very small sample size at site B and

17   conflicting efficacy results between the two sites.

18   At site B, the HAMD-17 baseline scores were higher.

19   The average modal dose for gepirone was lower, but

20   the placebo response was higher at this site.

21   We were aware that the mean modal dose for

22   gepirone was lower at Site B, the smaller site.

1   But the mean modal dose for imipramine was also

2   lower at this site. When we compared the observed

3   drug responses between the two sites, there was

4   really not much difference whether for gepirone or

5   imipramine.

6   We noted that the observed placebo response

7   was higher at site B, but it was difficult to

8   ignore the efficacy outcome from this site although

9   differential placebo responses might have increased

10   the variability of efficacy outcome.

11   Based on the sponsor's analysis, study 53

12   was not interpretable, but ODE1 concluded that this

13   study had assay sensitivity because the active

14   comparator beat placebo. The difference attributed

15   to the statistical model for analysis.

16   The sponsor included a treatment-by-center

17   interaction term in the model, but we did not. By

18   including the interaction term in the model in

19   sponsor's analysis, each site is assigned equal

20   weight when integrating treatment effects from both

21   sites, whether the site is large or small.

22   Our approach by excluding the interaction

1   term weighted two sites approximately by their

2   sample sizes, roughly, 72 percent weight to site A

3   and 28 percent to site B when integrating treatment

4   effects from both sites.

5   Before moving on to the discussion of the

6   maintenance study, I would like to give you

7   graphical displays of the 12 short-term trials.

8   This plot displays ODE1 results using the common

9   endpoint HAMD-17 for each trial. The vertical axis

10   represents the treatment effect relative to

11   placebo. More negative indicates larger treatment

12   effect.

13   The horizontal axis represents study ID.

14   The plot is divided into four panels. From your

15   left to right are the two positive studies, the

16   three negative studies, the four trials in dispute,

17   and the three uninterpretable trials.

18   Each closed circle represents a treatment

19   effect: gepirone in blue and active comparator in

20   yellow. The size of a circle is proportional to

21   the sample size.

22   In the first panel, the two blue circles are

surrounded by a red ring, denoting that gepirone

beat placebo in the corresponding studies 07 and

01. Above the red reference line, there are four

blue circles, denoting that gepirone was

numerically worse than placebo in the corresponding

four trials. One of them was negative and three

were in dispute.

In the dispute panel, there are two yellow

circles surrounded by a red ring denoting that

active comparator beat placebo. In the same panel,

on the top, there are three blue circles with

yellow stars inserted denoting that active

comparator beat gepirone.

This is a similar plot but based on the

prespecified primary endpoint and the analysis

method. In the dispute panel, no red rings, no

yellow stars, indicating none of the pairwise

comparisons reached statistical significance.

The previous two plots are displayed side by

side here for quick comparison. A noticeable

difference in the two plots is observed treatment

effects in study 53, the last one in the dispute

1   panel.

2   This was a two-center study with imipramine

3   as the active comparator. The difference resulted

4   from the statistical model used in analysis. The

5   sponsor included a treatment-by-center interaction

6   term but we did not.

7   Now, moving on to the maintenance study,

8   this was a randomized withdrawal study. In the up

9   to 12 weeks open-label phase, patients received

10   flexible dose of gepirone. Responders at the end

11   of open-label phase were randomized to continue

12   gepirone or switch to placebo. In the double-blind

13   phase, the ITT analysis set consisted of 250

14   patients.

15   The primary endpoint was the proportion of

16   patients who had relapsed. The primary analysis

17   was CMH test adjusting for center. Time-to-relapse

18   was a secondary endpoint analyzed by the log rank

19   test.

20   The maintenance study was first included in

21   sponsor's second resubmission. At that time,

22   sponsor considered the primary efficacy result

1 positive, but we considered negative.

2 Nevertheless, both sides agreed on the negative

3 finding of time-to-relapse.

4   Why was there a disagreement on the primary

5 efficacy conclusion? There were 5 patients on

6 gepirone who actually relapsed per the case report

7 form records but were not coded as relapses in

8 sponsor's analysis. In addition, there were 32

9 patients excluded from sponsor's analysis because

10 they came from small centers where all patients

11 were randomized with the same treatment arm or

12 small centers where no patients relapsed.

13   In order to bring these patients back to

14 analysis, we used a few common approaches to pool

15 these centers to form pseudo-centers and found that

16 results were negative regardless of pooling

17 algorithm.

18   The maintenance study was subsequently

19 included in the third resubmission. A major

20 difference from the previous submission was that in

21 this submission, the sponsor modified the original

22 ITT analysis set by excluding 40 patients who were

1   considered protocol violators.

2   The study was also included in sponsor's

3   informal appeal in 2012. A major difference from

4   the previous submissions was that the sponsor

5   redefined true responders for randomization into

6   the double-blind phase.

7   From both re-evaluations, the sponsor

8   concluded all the results positive because all the

9   p-values were less than 0.05. But sponsor believes

10 that the trial would have been positive if it were

11 adequately designed and conducted.

12   We were concerned about the redefining

13 patient population after data unblinding. Also, we

14 did not find the sponsor's post hoc evaluations

15 reached statistical significance because sponsor

16 did not correct the status of the five relapsed

17 patients on gepirone in their analysis; they did

18 not add in the 32 deleted patients who came from

19 the small centers and did not remove all patients

20 who should've been removed per their various

21 post hoc redefinitions of responders.

22   After necessary corrections according to

1   sponsor's various redefinitions, we found all the

2   results inconclusive with p-values ranging from

3   0.11 to 0.25.

4   Despite the controversies on whether certain

5   studies should be considered failed or negative,

6   there was an agreement on two positive studies out

7   of 12 short-term studies. The maintenance study

8   was negative but with the trend in favor of

9   gepirone.

10   The big challenge is how to evaluate the

11 overall evidence for effectiveness. Dr. Temple and

12 Dr. LaVange will present perspectives to address

13 this issue later on. Thank you.

14   **FDA Presentation – Mitchell Mathis**

15   DR. MATHIS: Mitch Mathis from Psychiatry

16 Products again to present the safety review. The

17 outline for our review of safety will look at

18 exposure, deaths, serious adverse events or SAEs,

19 dropouts, common adverse events, AEs, common and

20 drug-related AEs, and laboratory findings.

21   In terms of exposure, we used a cutoff date

22 of 6-23-03 because we had those data. There were

1   almost 5000 subjects exposed to at least one dose

2   of this drug, 732 for at least 6 months; and it

3   looks like nearly 500 for at least 6 months at

4   doses at or equal to 40 milligrams, 260 exposed for

5   a year and that includes a 170 exposed for

6   12 months. The dose is greater than 40 or greater.

7   There were 9 deaths reported in the phase

8   2/3 studies, 5 in the gepirone group, one in the

9   immediate-release, 4 in the extended-release.

10   There's one in paroxetine, one in fluoxetine, one

11   on placebo, and one during washout period prior to

12   randomization. The review team assessed these

13   deaths as not being study drug-related.

14   Here they are so that they can be in the

15   record. I should note that there were some

16   suicides as unfortunately there are in these

17   trials. This was at time prior to our active

18   inquiring about suicidality in clinical trials

19   where they were reported as adverse reactions.

20   Other serious adverse events, there were

21   none in the phase 1 studies. In the phase 2/3

22   studies, you can see that 2 and a half percent of

1   patients had SAE on gepirone, which is more than

2   placebo and imipramine, but less than fluoxetine

3   and paroxetine. The most commonly reported SAEs

4   with gepirone were depression, suicidal ideation,

5   suicide attempt, and pneumonia. The review team at

6   the time agreed that these serious adverse events

7   did not define a safety concern for the drug.

8   In terms of dropouts, percentage with an AE

9   leading to premature withdrawal from the trial,

10   highest was in the imipramine group

11   percentage-wise, 24 percent. Gepirone IR and ER

12   groups are there; they're broken out by ER and the

13   IR. You can see the percentage lower on ER.

14   Twelve percent, paroxetine; 9, fluoxetine;

15   7 percent on placebo.

16   The most frequency AEs leading to dropout

17   from the combined IR and ER groups were dizziness,

18   nausea, headache, insomnia, and anxiety. For the

19   ER, it's the same adverse events but smaller

20   percentages. There were no safety concerns

21   identified from the dropout data from the review

22   team.

1   Common and drug-related adverse events,

2   defined as occurring at 2 percent of gepirone

3   ER-treated patients and at a rate that's at greater

4   than or equal to 1 and a half times the rates seen

5   in placebo patients, included dizziness,

6   paresthesia, tremor, nausea, dyspepsia, insomnia,

7   abnormal dreams, middle insomnia, agitation,

8   palpitations, increased weight, tinnitus and

9   blurred vision. The safety review team concluded

10   that there were no adverse events that would

11   preclude approval among common and drug-related

12   adverse events. There were no relevant changes in

13   vital signs or laboratory values.

14   In conclusion, the team of evaluated the

15   safety data multiple times. There are multiple

16   submissions of this application and have not

17   identified any relevant safety issues that would

18   preclude an approval recommendation.

19   DR. D'AGOSTINO: Dr. Temple? Are you

20   presenting?

21   **FDA Presentation - Robert Temple**

22   DR. TEMPLE: Hello, everyone. I will try

1   not to repeat too much of what Dr. Yang has already

2   said, so there may be pauses in there where I do

3   that.

4   I should say at the outset -- it's not on

5   this slide -- that we're going to describe

6   reservations about whether the drug has fulfilled

7   the substantial evidence requirement. That does

8   not represent a conclusion by us that we know the

9   drug is ineffective or anything like that. It's a

10   question of whether the evidence of effectiveness

11   is sufficient, meets the usual legal standard or

12   what we have interpreted to be the standard.

13   We have always said that what Fabre-Kramer

14   needs is one more adequate and well-controlled

15   study. So we don't dismiss the available evidence.

16   We just have concluded that, for various reasons,

17   it's not sufficient.

18   The issues here are whether there really is

19   substantial evidence of effectiveness that gepirone

20   is effective, and an important component of that is

21   what are the implications of studies that are

22   failed or negative? And there is some debate about

1   which those are.

2   A critical question is whether four

3   placebo-controlled studies with an active control,

4   that clearly showed no effect of gepirone, should

5   be considered failed or negative. Failed studies,

6   in our view, do not undermine an effectiveness

7   finding. They're simply evidence that the study

8   couldn't show anything. But negative studies,

9   depending on their number, could have a negative

10   effect on your conclusions.

11   FDA's conclusion that those four studies

12   were negative for gepirone but did have assay

13   sensitivity is based on findings that the active

14   control was significantly superior to placebo in

15   two studies on the HAMD-17. And in three cases,

16   three of those studies were actually superior to

17   gepirone, which I have to say is a very unusual

18   finding in 3-arm studies.

19   We've looked a little bit -- and Dr. Mathis

20   may know -- and I think we're aware of one such

21   study, but it's very uncommon. And to have all

22   three of them go that way is very unusual. In any

1   event, we believe all four of those studies had

2   assay sensitivity so that they were in fact

3   negative gepirone studies.

4   In two of the four active control studies,

5   the primary endpoint was a HAMD-25, chosen because

6   we believe -- this could be explained further -- it

7   might have a possibly greater ability to assess

8   atypical symptoms of depression. Perfectly

9   reasonable. We didn't object to using it had the

10   study succeeded.

11   In those studies, the active control clearly

12   failed on the HAMD-25. It didn't show an effect.

13   A third study using MADRS as the primary endpoint

14   also failed. But all three of those studies showed

15   a significant effect on the HAMD-17 compared to

16   gepirone. They didn't all beat placebo.

17   In the fourth study, the HAMD-17 was the

18   primary endpoint, and the trial showed an

19   effectiveness of the active control versus placebo

20   but no significant effect of gepirone, although

21   gepirone leaned in that study.

22   Now, our interpretation of the active

1   control studies and what they mean thus rests in

2   part on the demonstrated effect of the active

3   control on an endpoint that was not the primary

4   endpoint. We know that, and that's been challenged

5   and this is going to be something that needs to be

6   discussed.

7   These are not easy questions, and you'll

8   hear various -- or you have already heard various

9   views on how to interpret the available evidence,

10   especially the active control studies.

11   The Food, Drug, and Cosmetic Act in 1962, as

12   amended, created the effectiveness standard.

13   Approval requires substantial evidence that the

14   drug will have the effect claimed in labeling,

15   et cetera, et cetera, as Dr. Jenkins has already

16   said. And the only source of such evidence is

17   adequate and well-controlled investigations

18   conducted by appropriate experts.

19   The legislative history in the center report

20   makes it very clear that the term investigations

21   was deliberately plural, and we have held, with

22   some exceptions -- and there's FDAMA legislation in

1   1997 that said there was circumstances in which one

2   study would do, that we generally expect two

3   adequate and well-controlled studies to show

4   effectiveness.

5   You might wonder how that came to be. Well,

6   we think the legislative history is compatible with

7   that but the legislative history didn't go through

8   what the statistical reasons for that are.

9   Nonetheless, we've always thought that you

10   need two studies because there's a possibility that

11   a single favorable study could be a false positive.

12   It's perfectly obvious that if you do a lot of

13   studies, 5, 10 or whatever, the chance that one of

14   them will show nominal significance starts to

15   increase as you do that.

16   Using a sort of crude calculation, again,

17   related to one study, considering trials as either

18   yes, less than 0.05 or no, greater than 0.05,

19   calculating the chance of having one trial show

20   significant effect, P less than 0.05, when there

21   was, in fact, no effect, you can calculate that for

22   each trial, there's a 5 percent chance of a false

1 positive and a 95 percent probability that there

2 will not be a false positive.

3 So with two trials, the chance of a false

4 positive is 1 minus 0.95 times 0.95 or about 10

5 percent. But if you now get to, say, five trials,

6 the chance of a false positive, that is, concluding

7 that the drug works when there was no effect at

8 all, is 1 minus 0.95 to the fifth or about

9 23 percent.

10 For one study, we know that everybody is

11 reasonably comfortable with those kinds of numbers.

12 I should note that these trials are treated as yes

13 or no. They're not meta-analyzed to see whether

14 the leans make it or something like that. And the

15 reason, I think -- and it's sort of implied in the

16 legislative history -- is that they thought there

17 were various reasons why a trial might look

18 positive, either chance or something funny going

19 on, and they really wanted two independent

20 findings. I think that we have persisted in having

21 that attitude although there is some debate about

22 it.

1   The reasons for wanting to be sure that a

2   study has shown an effect are sort of obvious. We

3   always want approved drugs to be effective. That's

4   particularly true when the effect of the drug is

5   very important, as is the case for antidepressants.

6   You really don't want to use an

7   antidepressant that doesn't work. There are many

8   effective antidepressant alternatives, and you

9   don't want to -- even though we all recognize the

10   need for more effective drugs -- you don't want an

11   ineffective drug to go out there.

12   When there are two positive trials, as

13   actually Mary Johnson calculated a very long time

14   ago for Fabre-Kramer in an early submission, the

15   chance of both being false positives is quite

16   small. Even for two positives out of 12, the true

17   nominal overall p-value false positive probability

18   is still less than 5 percent, maybe it's 0.02 or

19   something like that.

20   I would argue, however, that is not near the

21   strength of evidence that the legislators had in

22   mind as a basis for approval. It's a relatively

low probability, but it's still a possible error

and maybe that's too much, which again is why we've

asked for a third study.

Now, turning to a different subject, why do

active control trials? We know from very extensive

experience with antidepressants that about

50 percent of trials of effective antidepressants

drugs we concluded work do not show a statistically

significant effect. They're negative studies.

Whether that number is growing or not, I'm not

sure, but it's been true for many, many, many

years. Perfectly good trials of effective drugs

fail.

These negative studies, since these are all

effective drugs, must mostly be studies without the

ability to detect the drug effect, that is, they

are failed studies, studies without assay

sensitivity. And many, many, many years ago,

because of the high rate of failures -- of negative

studies, of what appeared to be negative studies of

antidepressants -- we began to advise sponsors to

include a third arm so they could distinguish a

failed trial, one that lacked assay sensitivity and

that would have no implications for the

effectiveness of the test drug from a truly

negative trial, where the test drug did not work

even though the trial did have assay sensitivity.

Assay sensitivity was considered present if

the active control was statistically significantly

superior to placebo. We did not, to my best

knowledge, address the question: suppose it didn't

beat placebo but beat the test drug?

Again, not to repeat myself too much, but a

failed study really doesn't show ineffectiveness of

the test drug. It is totally uninformative, and we

all agree with that. A study with assay

sensitivity, however, that fails to show an effect

of the test drug is a truly negative study.

Our purpose in all in this was to try to

keep drugs from looking like they didn't work when

they didn't deserve to. It was our attempt to be

helpful. Always good to remember.

Whether a trial was negative or failed

matters. If we have two positive studies out of

1 five, which could be the case if we dismiss those

2 four active-controlled studies and considered

3 failed, then the chance of the positive findings

4 being a chance occurrence is pretty low. I don't

5 think we'd be too worried about that.

6   But if there are two positive studies out of

7 nine, which would be your conclusion if the four

8 active controls did have assay sensitivity but

9 failed to show an effect of gepirone, a spurious

10 result becomes more possible, and that's what we're

11 sort of worried about.

12   Now, I realize that there could be an

13 argument that treating trials as yes or no, with

14 respect to success or failure, loses information,

15 as Dr. LaVange has written. And to some extent,

16 Gary Koch's analysis puts all the studies together

17 and tries to use all of their information.

18 Although I have to note, he did not show you an

19 analysis that left out the two positive studies.

20 We've done those analyses; they show nothing.

21   If you take all the negative studies, they

22 don't show anything. Anyway, we've long held that

1   the law, with respect to effectiveness, wanted

2   independent confirmation of the effectiveness

3   finding in contrast to an overall p-value from

4   upping the data.

5   I mean, as everybody would recognize, a

6   single positive study with a P less than 0.01 or

7   something like that, pooled with one or two

8   negative studies with a weak trend toward

9   effectiveness, might show a nominal p-value of

10 0.05. That doesn't do it for us. We wouldn't

11 consider that substantial evidence. We want a

12 second confirmatory study.

13   That's easy when there's only one study

14 that's positive. How to translate that into two

15 positive studies out of nine is a more difficult

16 question. That's what we're faced with.

17   As you've heard from Dr. Yang -- I won't try

18 to repeat it -- there are two clearly successful

19 studies. Everybody agrees. There are three

20 largely uninterpretable studies with gepirone, and

21 everybody largely agrees on that. There are three

22 negative studies, although the sponsor considers

1   just two of them supportive and leaning a little

2   bit. And then there are four negative or failed,

3   and there's a debate about which they are.

4   These are them. You've seen this before.

5   Probably the main point here is that if you look at

6   the effect of the active control versus gepirone,

7   in the first one, fluoxetine was more effective and

8   significantly so. In the second one, paroxetine

9   was more effective and significantly so. It also

10   happened to beat placebo in that one. And in the

11   third one, fluoxetine, again, beat gepirone, and

12   not by just a little bit, but by effect sizes that

13   are not so different from the effect sizes we're

14   used to. That is very unusual.

15   We all know -- you've heard this -- in three

16   of the four trials, the active control failed on

17   its primary endpoint, HAMD-25 or MADRS. Dr. Yang

18   has explained why we thought the HAMD-17 was

19   reasonable. Obviously, it's an endpoint used in

20   depression studies. It's well-established. It

21   worked in two positive studies as did also the

22   HAMD-25.

1   We thought it was reasonable to use the

2   HAMD-25 as the primary endpoint, but we also think

3   it's reasonable to use the HAMD-17 to assess assay

4   sensitivity. That's plainly one of the matters

5   under debate.

6   Well, I won't say it again. The active

7   control beating the test drug, apart from being

8   very surprising and unusual, we believe supports

9   assay sensitivity. And as I said, it's almost

10 never seen.

11   It's always important to worry about

12 changing endpoints. Heaven knows, if someone has a

13 primary endpoint and wants to declare effectiveness

14 by looking at a secondary endpoint, we would have

15 major reservations about that. We're worried about

16 that.

17   But in this case, our conclusion has been

18 that the HAMD-17 in this case is not something

19 picked out of the air and is a reasonable basis for

20 assessing assay sensitivity. Remember, we don't

21 have to be absolutely sure.

22   What in this case I would say is the truth

1   is that these four studies may or may not be

2   definitive on whether gepirone doesn't work. We're

3   not concluding that it proves that it doesn't work;

4   it raises doubts about the available data and makes

5   us believe there needs to be more data.

6   All in all, it's very impressive that in

7   three trials, with what we believe had assay

8   sensitivity, gepirone was actually worse than

9   placebo and statistically significantly inferior to

10   the active control. That has to make you worry

11   about what the effect of the drug is and create a

12   need for further data. Thanks.

13   DR. D'AGOSTINO: Dr. LaVange?

14   **FDA Presentation – Lisa LaVange**

15   DR. LaVANGE: Good morning. I'm

16   Lisa LaVange. I'm the director of the Office

17   Biostatistics in CDER, and I will give the

18   perspective from the office on the evidence

19   provided for gepirone extended-release as a

20   treatment for major depressive disorder.

21   You've already heard from Dr. Peiling Yang,

22   one of our many talented team leaders in the Office

1   of Biostatistics. Dr. Yang directed the

2   statistical reviews of gepirone. This involves

3   several different reviewers over many years.

4   While she accurately represented the outcome

5   of that long review process, I'd like to bring us

6   up to date on the office's current thinking about a

7   couple of issues that are encountered in these

8   reviews. In particular, I'd like to talk about

9   trial sensitivity and about how to synthesize

10   information across trials in a program.

11   Now, I'll just make a comment following the

12   sponsor's presentation. I won't particularly

13   discuss issues with handling missing data, but I

14   wanted to comment on them since they'd been raised

15   and were the source of some questions; in

16   particular, Professor Koch mentioned of the use of

17   an estimand that considers dropouts as non-

18   responders and that that is consistent with our

19   current thinking about ways to handling missing

20   data, and that is correct.

21   Simple, single valued imputation methods,

22   like last observation carried forward, LOCF, is no

1   longer recommended due primarily to issues with

2   underestimation of the variance as was mentioned.

3   And even some newer methods that are based on mixed

4   models or repeated measures that require missing at

5   random assumptions are also problematic because

6   they often treat dropouts as if they had continued

7   taking the drug.

8   Now, the fact that the results of the two

9   meta-analyses Professor Koch presented with and

10 without the treatment of dropouts as

11 non-responders, give similar results, I think,

12 points to a conclusion that we had arrived at, that

13 the differential response rate here is not a major

14 factor. I won't say anything further about it, but

15 others might.

16   All right. I'll go back to the topic I

17 mention, trial sensitivity. Trial sensitivity is a

18 concept used to characterize trials as being

19 informative, particularly when it's not ethical to

20 include placebo in the trial. It's also called

21 assay sensitivity. It's discussed in the CDER's

22 2010 draft guidance on non-inferiority trials. The

1   definition from that guidance is quoted here.

2   It's a critical concept for non-inferiority

3   trials because the effectiveness of the test drug

4   in those trials relative to placebo can only be

5   inferred based on its comparison to an active

6   control in conjunction with the assumption that the

7   active control had the effect that it would be

8   assumed to have based on past history.

9   In other words, the active control is

10   assumed to have a particular level of efficacy

11   based on past studies of that active control versus

12   placebo. There's no placebo in these trials, so

13   you can't prove that.

14   The concept is thought to be useful in

15   developing antidepressant trials but for different

16   reasons. Here, it's considered ethical to include

17   a placebo arm in the trial, so the effectiveness of

18   the test drug is based on a direct comparison and

19   not inferred indirectly as in an non-inferiority

20   setting.

21   It's also fairly common, as you heard today,

22   for approved antidepressants to fail to

1   differentiate from placebo in many trials. The

2   agency, therefore, advises sponsors to include an

3   active control as a third arm in the trial of an

4   antidepressant to aid in its interpretation. If

5   the trial fails to show an effect to the test drug,

6   the comparison of the active control to placebo can

7   help explain why.

8   There's currently no agency guidance on when

9   or how to determine trial sensitivity. The

10   non-inferiority guidance that I cited defines the

11   concept, but then states that it has to be assumed.

12   It can't really be calculated.

13   In the absence of such a guidance, the

14   approach that's evolved for antidepressants and the

15   approach that was taken in our original reviews of

16   gepirone was to base sensitivity on a test of the

17   active control versus placebo conducted when the

18   test of the new drug versus placebo was not

19   significant.

20   The trial sensitivity is a concept, not a

21   dichotomy, in my view. And information from a

22   trial is lost when reduced to a single yes/no

1   quantity. I think defining it as such poses

2   particular difficulties.

3   First, we have no idea what the operating

4   characteristics are with basing a decision about a

5   drug's effectiveness in a trial on this type of

6   sequential determination. Second, there are a

7   variety of analyses that can be used to produce the

8   trial sensitivity p-value. Interpreting the

9   significance level of any one of these can be

10 difficult.

11   Particular problems with determining trial

12 sensitivity based on the results of those

13 statistical tests are illustrated in this

14 application and include the problem with early

15 terminations, which has already been discussed.

16   There were four trials terminated early for

17 different reasons, two with no active control, so

18 you can't determine sensitivity, and two with

19 active control that showed no sensitivity by the

20 primary analysis method. One of these, the 053

21 study, the primary analysis method was questionable

22 due to the way the interaction term was handled. A

1   reanalysis gives sensitivity and it also gives a

2   favorable result for gepirone.

3   Early terminations cause a problem because

4   they usually are thought to have insufficient power

5   to test for significance. Another presenter has

6   mentioned the problem of early stopping of the dose

7   titration, which could also be at play here.

8   Acknowledging that the estimated treatment

9   effects may be underestimated due to the early

10   stopping, the effect estimates may still contain

11   information. And it's been pointed out, and I'll

12   point out again, that in the 053 study, the effect

13   of gepirone at the early stopping point was

14   actually quite close to the effect in the two

15   positive studies, which I think carry some

16   information.

17   The particular problems with determining

18   trial sensitivity based on the results of the

19   statistical test also include multiplicity. This

20   has been discussed. I listed the three studies

21   here that used different endpoints. At least two

22   of them had different target populations; there

1   were other differences with these trials. And the

2   primary analysis of these trials and other trials

3   yield unambiguous results with respect to trial

4   sensitivity.

5   Now, to repeat myself, there's no agency

6   guidance on what endpoint or what analysis method

7   to use to test for trial sensitivity, but I feel

8   the best way to avoid multiplicity due to multiple

9   analysis choices is to stick with the primary

10   endpoint and the primary analysis method, provided

11   it's not incorrect.

12   This approach is unambiguous. There's no

13   choice made about which p-value to apply. There

14   are always alternative analyses that the sponsor

15   could apply, that FDA reviewers could apply. Those

16   discussed here and by Dr. Yang and also Dr. Temple

17   include changes in endpoint, changes in the

18   residual variance computation, changes in the

19   analysis model, and the handling of interaction

20   terms.

21   Many of the post hoc analyses, we, at FDA,

22   have applied to this application were seen as

1   preferable to the sponsor's analysis by the primary

2   reviewers at the time, and I might agree absent

3   prespecification. Their application may produce

4   estimates of effect that have some advantage for

5   interpretation such as the 053 study mentioned on

6   my previous slide.

7   The problem, to me, arises when the p-value

8   is the sole basis for making a determination about

9   sensitivity. When there are multiple p-values to

10 choose from, then what's the appropriate decision

11 rule to use?

12   I will just mention -- because this is

13 something we think about a lot at the FDA -- the

14 Office of Biostatistics follows a key principle of

15 good statistical review practice, namely that

16 sensitivity analyses we conduct should be designed

17 to assess the impact on the study's results of an

18 assumption required for the analysis that was

19 prespecified to be correct and if that assumption

20 is difficult to verify. Post hoc analyses that are

21 conducted for the sole purpose of changing a good

22 result to bad are not particularly useful for

1   decision-making.

2   Now, I'll mention one other particular

3   problem, and that's basing trial sensitivity on a

4   test of the active control versus the test drug

5   when neither the active control nor the test drug

6   differentiates from placebo. This is not a

7   multiplicity problem. This rather relates to the

8   meaning of sensitivity itself.

9   In these cases, the active control failed to

10   demonstrate its assumed effect. I think that's

11   really not under argument. This indicates, to me,

12   a lack of sensitivity in the sense that it is

13   defined in the one guidance that tries to do that.

14   Interpretation of the results of these trials, I

15   think, needs to take that into account.

16   I'll turn to the second issue I wanted to

17   talk about, and that is the synthesis of

18   information and in particular some comments I have

19   about the counting trial method that Dr. Temple had

20   just talked about.

21   Assessing trial sensitivity can help explain

22   why a trial is negative, and I think it's useful in

1   the sense both for people developing drugs as well

2   as for us reviewing drugs. I don't think it's as

3   useful as a criterion for determining which trials

4   to consider in synthesizing information. I like to

5   look at the information I can get from all of the

6   studies if it makes sense to do so.

7   I've mentioned already I view trial

8   sensitivity as a concept, not a dichotomy. Its use

9   in determining which trials we can or can't

10   consider in making a decision, I think, is

11   problematic.

12   Going a step further and computing the

13   probability of a false result based on a Bernoulli

14   outcome of each trial considered positive or

15   negative just exacerbates this problem. With this

16   approach, all positive results are treated the same

17   regardless of how positive. All negative results

18   are treated the same regardless of how negative.

19   While this is a synthesis, it is a

20   meta-analysis. It's a meta-analysis of Bernoulli

21   outcomes of a series of trials. It's a method that

22   I think had a long tradition in the medical

1   literature in particular. Several decades ago, it

2   was realized that the conclusions from these

3   procedures can be misleading, and that's why the

4   development of other methods of meta-analyses then

5   proceeded.

6   If you can indulge me, a simple example

7   here. If you have two studies of exactly the same

8   size and an observed T statistic of 1.96 is

9   observed in those studies, the studies are barely

10   significant at the 0.025 one-sided level that we

11   require. A meta-analysis approach, given that

12   they're the same size with the same result, is just

13   this average the two statistics. And if you do

14   this, you get a probability or a p-value of 0.003.

15   The Bernoulli trial meta-analysis, that I

16   call the method that's been talked about here

17   today, gives you a very small p-value, as

18   Dr. Temple mentioned, and it's less than 0.001.

19   Now, suppose a third study is run and the

20   third study is negative, how do you factor that

21   information into what you already know? Well, if

22   you apply this Bernoulli trial meta-analysis at

three trials, you will always get a weaker result

provided that the third study is negative of any

size. Even if the p-value is 0.026 or 0.027,

you'll get a weaker result. If you do a proper

meta-analysis of the three trials, you'll get a

stronger result as long as the third study has a

trend. I find this to be more informative.

I think we can conclude that based on a

Bernoulli trial meta-analysis of the gepirone

trials, two positive trials out of two gives a

smaller overall p-value than 2 out of 12. On this,

Dr. Temple and I agree, I'm pretty sure.

I would also conclude that the global null

hypothesis that gepirone has a true effect of zero

in every trial is most likely false. As I

mentioned earlier, the Bernoulli trial analysis is

valid, but it's an inefficient test of the global

null hypothesis. And any type of meta-analysis

applied to the actual outcomes and not just the

zero-1 outcome of positive or negative will yield a

valid, more efficient test of the same hypothesis

and also with a more reasonable alternative.

1   I will just comment also the reason

2   Dr. Temple and I disagree on that global null

3   hypothesis, whether it's been rejected or not, is

4   the significance or test level that we're applying

5   to the analysis.

6   I actually don't think it's appropriate to

7   use the single-study criteria of 0.025 squared or

8   0.00065 because we're not talking about replacing

9   the evidence of two studies. The replication has

10   already happened here. This isn't a single study

11   submission, so we don't need to replace the

12   replication. But also, replicating has value other

13   than just a small p-value.

14   Believe it or not, I manage to have in my

15   slide a meta-analysis that you haven't seen yet, so

16   apologies for that. It's hard to believe there's

17   still another one.

18   Actually, in, I think, 2007 in the sponsor's

19   submission at that time, they presented a

20   meta-analysis of seven trials, which were the two

21   positive, the three others that we have been

22   considering to be interpretable, and then two of

1   the studies that were stopped early and didn't have

2   a third arm, so you couldn't tell whether they had

3   assay sensitivity.

4   I just present that here and comment that

5   the meta-analysis of the five interpretable studies

6   the sponsor presented gives an effect size in the

7   same range of about minus 1.32. Then FDA conducted

8   another meta-analysis of nine trials, but these

9   were the nine trials where HAMD-17 was prespecified

10   as primary endpoint.

11   Now, we use that as sort of a proxy of these

12   studies being homogeneous enough to meta-analyze.

13   They are not the same nine trials that the sponsor

14   included in their meta-analyses. We did not

15   include the -- I think it was the study 017 because

16   it had a different primary endpoint. But we did

17   include the studies that stopped early. So it's

18   not quite a worst-case scenario. All three of the

19   effects in the meta-analysis are about the same,

20   and they're all about half of what was observed in

21   the two positive trials.

22   My conclusion based on this is that

1   gepirone's effectiveness varies under some

2   conditions of use or with some patients. And I

3   will mention the analyses we did also -- I'm

4   reporting the inverse weighted mean method here,

5   but we did look at the random effects

6   meta-analysis, and the standard error estimates

7   were slightly smaller as were the point estimates.

8   All right. Just a couple of comments about

9   the maintenance trial. I think you've seen already

10   there were clearly some quality issues that

11   negatively impacted the study, conduct and its

12   ability to succeed. The sponsor's analysis was

13   incorrect due to exclusion of patients from sites

14   with only one treatment arm. FDA's reanalysis

15   shows a favorable trend, but the results are not

16   significant. However, I don't believe the results

17   are conclusive either in showing that gepirone does

18   not work and that another maintenance trial is

19   needed. I think everyone agrees on that too.

20   In terms of just my conclusions, I think

21   gepirone has been shown to be effective in two

22   positive phase 3 trials. I think everybody thinks

1   that.   They were of reasonable size and quality and

2   they had similar results.

3   There's inconsistency of effectiveness.

4   It's shown across the trial regardless of whether

5   you synthesize this with a Bernoulli outcome

6   meta-analysis or a meta-analysis of the actual

7   effects of the studies. The synthesis I mentioned

8   shows that the average effect can be about half

9   that of positive trial results.

10   The maintenance trials are inconclusive, so

11   I think different people have different

12   conclusions. Mine would be that there's evidence

13   that gepirone works but perhaps not reliably.

14   Thank you.

15   **Clarifying Questions to FDA**

16   DR. D'AGOSTINO: Thank you. We're now going

17   to move to asking the FDA for clarifying questions.

18   Before we do that, I'd like to just -- on behalf of

19   the committee and taking prerogative as chair,

20   hearing the FDA, it seems to me -- the FDA

21   presentations, reading the materials, it seems to

22   me that we have a number of disputes with

1 the -- the FDA has a number of disputes with the

2 company.

3   One of the major causes is the shifting from

4 the HAMD-17 to HAMD-25, where the sponsor claimed

5 that they had justification for it, looking for

6 atypical parts, features, and so forth in their

7 studies.

8   I'm concerned -- and I'd like to hear the

9 FDA speak to it. I'm concerned about the

10 multiplicity and the post hoc aspects that come

11 with doing that. How do you pick an endpoint that

12 wasn't the sponsor's endpoint, chase it down, it's

13 post hoc? You've got all these multiplicities

14 then.

15   One of my concerns also is when you start

16 finding things strange, you keep looking and you

17 find more and more strange things. We may be

18 just -- in some of the things that are being picked

19 up, it may just be the fact that there's some

20 strange studies that maybe should be discarded or

21 maybe should be focused very much. Dr. LaVange's

22 approach is quite useful in what she's saying here.

1   Then, I guess I'm interested in -- maybe

2   it's premature to ask -- how does the

3   sponsor -- excuse me -- how does the FDA respond to

4   the meta-analysis that was presented by the sponsor

5   where they were looking for HAMD-17 greater than or

6   equal to 20 as an entry criteria for their

7   analysis.

8   Could I get a quick response to some of

9   those questions? And then, I'm sure the panel will

10   ask much more penetrating questions, much more

11   clearer questions.

12   DR. TEMPLE: You probably have to hear from

13   several people. Remember, as I said at the

14   beginning, we have not reached the conclusion that

15   gepirone couldn't work. The question is whether

16   the evidence that it does work is sufficient.

17   We recognize that going from HAMD -- on the

18   so-called failed studies, on the studies that are

19   debatable about whether they failed or were

20   negative, we recognize that going from the HAMD-25

21   to the HAMD-17 raises a multiplicity question. We

22   know that. We worry about that all the time.

1   What's of interest here is a couple of

2   things. One, the HAMD-17 was widely used by the

3   company in all of its trials. This is not out of

4   the blue. It's not one of 50. It's a major one.

5   It's commonly used, and it is the endpoint on which

6   the control agents are known to work in depression.

7   So it isn't crazy to do it, but you're going to

8   have to worry about whether the fact that there

9   were at least two analyses done undermines the

10   meaning of it. That's a perfectly good question.

11   We know that.

12   Second, remember, what we're looking here is

13   whether there are things that make you doubt the

14   results of the two positive studies. There's no

15   question the two studies are positive. That's the

16   usual standard for approval. My contention would

17   be there's enough in the four studies with active

18   controls to make you worry about whether that

19   finding is true. And the findings are, one, the

20   drug didn't win, and two, very surprisingly, on a

21   well standardized endpoint, well-recognized

22   endpoint, the active control actually beat

1   gepirone, as I said, something we hardly ever see.

2   Those don't prove gepirone doesn't work.

3   Our concern is whether it gets in the way of

4   concluding that there is substantial evidence that

5   gepirone does, which matters to us. We do not want

6   to put into the world an antidepressant that

7   doesn't work. I don't know if that answers all of

8   your questions or not.

9   DR. PICKAR: Could we see a HAMD-17 and

10   HAMD-25? Nobody has shown one, and this is what

11   this whole thing is about, is that you've

12   re-analyzed the data -- I'm with you. I've done it

13   almost as long as you have.

14   But there was a change here. At least,

15   let's see what the issues are. I understand in

16   your briefing it's atypicality. The HAMD-17 is

17   very weighted to somatic things, as we've always

18   loved them to be, and certain drugs are going to

19   have the bigger somatic effect.

20   Other aspects of depression, I don't have it

21   off the top of my head what the differences are,

22   but I think we need that data, is to see what the

1   two is, and let the advisory board then

2   contemplate, a little bit at least, what the larger

3   significance of changing that is.

4   DR. YANG: This is Peiling Yang. Would you

5   like to see the distributions or the questionnaire?

6   DR. PICKAR: I'd like to see the

7   questionnaire. I want to see the 17 items and the

8   25 items.

9   DR. YANG: I have the 8 items appended to

10   the HAMD-17. The 8 items are in the backup

11   slide --

12   DR. TEMPLE: These are the additional items

13   in the HAMD-25. She's going to show you.

14   DR. YANG: Backup slide 13, number 13.

15   DR. PICKAR: Could we start with the

16   HAMD-17, and then see the additional ones?

17   DR. YANG: Does the sponsor have HAMD-17?

18   DR. PICKAR: This is so critical to this

19   whole conversation. Whether we should be or not,

20   it just is. Let's at least look at the data.

21   DR. TEMPLE: Peiling can show you the

22   additional questions. Whether we can find this

1   original 17, which are absolutely standard, I'm not

2   so sure. We'll see.

3   DR. YANG: These are the additional eight

4   questions assessing the reversed vegetative

5   symptoms. The first five are for the atypical

6   features.

7   DR. JENKINS: Dr. D'Agostino, this is

8   Dr. Jenkins. If we don't have the HAMD-17, we will

9   get those for you at lunch and try to have a slide

10   that we can show you after lunch.

11   DR. PICKAR: We really need to look at it.

12   Diurnal variation is one of the key elements of a

13   biological depression as long as I've been doing

14   it, and I'm an old man.

15   DR. JENKINS: Right. We'll try to do a

16   comparison chart.

17   DR. PICKAR: That's core. That's core to

18   the illness. Every bipolar depression, or folks

19   with bipolar depression, or even if it's

20   biological, tend to oversleep with hypersomnia.

21   That's part of the spectrum of serious illnesses.

22   It just fascinated me. And psychic retardation,

1   slowed motion, not able to be sharp --

2   DR. D'AGOSTINO: I think the point that

3   you'll get the 17 is after lunch, and then we can

4   see them as a totality.

5   Dr. Compagni?

6   DR. COMPAGNI-PORTIS: Dr. LaVange mentioned

7   something that I think is important, which is about

8   that the effectiveness varies. I'm not a

9   statistician, but when I looked at -- even if we

10   just look at the two trials that are positive, that

11   when you look at the dropouts, 97 were treated at

12   nine centers, but there's a small effect in only

13   three centers. That's on 007.

14   On 001, after dropouts, 74 were treated at

15   five centers, and there's an effect at only two. I

16   don't even know what my question is, but I think

17   that variability that you're bringing up is really

18   important because I don't know what's different at

19   these centers. I don't know if you have more to

20   say about if you have any thoughts about the

21   variability?

22   DR. LaVANGE: I actually haven't spent a lot

1   of time thinking about the within-trial variability

2   across centers, which I think is what you're

3   talking about, though we do often see center-to-

4   center variation, and we base our decisions on the

5   average effect, assuming the drug works better in

6   some centers than other centers and may not work at

7   all in some centers.

8   To some extent, what we're saying about

9   cross-trial variation could be applied within a

10   trial, but I don't think the results of the two

11   positive trials are being disputed, even within our

12   ranks. I'm not sure if that helps.

13   DR. COMPAGNI-PORTIS: I guess my point then

14   is just that it does seem like there may be a

15   population that this works for, but we don't know

16   who that is or why.

17   DR. TEMPLE: It's worth saying that possible

18   subset differences and things like that are present

19   for all drugs all the time, and they're very hard

20   to discern unless you try to -- unless you have a

21   clue and try to set up the things, so you stratify

22   or something like that. But that has never been

1   done, to my best knowledge, in depression trials.

2   DR. LaVANGE: Dr. D'Agostino, you had asked

3   earlier about multiplicity, and Dr. Temple

4   answered. Maybe I could just add another mention,

5   if that's okay.

6   DR. D'AGOSTINO: Please do.

7   DR. LaVANGE: I think multiplicity is a big

8   problem, and it happens if you determine assay

9   sensitivity based on a p-value. It's not as much

10   of a problem if you look at what the drugs do in

11   the studies, which is why I like the definition

12   that's in our non-inferiority guidance. Did the

13   active control do what you expected?

14   One of the things that's been troubling to

15   us is that there are some of the studies where the

16   four disputed studies -- the four studies where the

17   results are disputed, I don't know if the studies

18   are disputed -- that the active control actually

19   beat gepirone, and that that's something that we

20   never see. It's surprising.

21   But when I look at one of those studies,

22   which was -- I don't know if you have to pull up

1   the slide, but it's on slide 10 of Dr. Yang's

2   presentation -- the active control versus placebo,

3   there was -0.68. I just don't know what to do with

4   that. That active control did not do what it was

5   supposed to do.

6   With an effect that little, gepirone didn't

7   do what it was hoping to do either. It's actually

8   worse than placebo by one point. Are those two

9   occurrences noise in different directions? And you

10   put them together and you get a significant

11   difference, does that mean the study has assay

12   sensitivity?

13   I mean, I just don't know. I know that

14   Dr. Temple thinks it does because the study

15   differentiates two arms. But if I think about the

16   study being a test of whether active control did

17   anything, I have a problem with that study.

18   So even looking at the effects doesn't

19   always give you the answer that you need, but I

20   think it's better than just looking at the p-value

21   because I don't know what to compare the p-value

22   to. Is 05 still in play here if I've done four

1 different analyses? I don't know.

2   DR. D'AGOSTINO: Victor, are you still on

3 the phone?

4   DR. DE GRUTTOLA: Hello?

5   DR. D'AGOSTINO: Do you have any questions?

6   DR. DE GRUTTOLA: Yes. I wanted to ask on

7 question on analysis that [indiscernible] --

8   DR. D'AGOSTINO: Your voice is breaking up.

9 Can you understand what he's saying? We can't

10 understand, unfortunately, what you're saying.

11   DR. JENKINS: If you're using a

12 speakerphone, can you use the handset?

13   DR. DE GRUTTOLA: [Indiscernible.]

14   DR. D'AGOSTINO: You're still breaking up.

15 Maybe what we can do is over lunch, we can get your

16 question and get back, and maybe also find a better

17 mechanism to call in so you won't break up.

18   Dr. Follmann?

19   DR. DE GRUTTOLA: I will do that.

20   DR. FOLLMANN: Thanks. This is

21 Dean Follmann. In psychiatric drugs, we have this

22 concept of assay sensitivity, which is not commonly

1   used, I think, in other settings. And the language

2   and what we've seen today suggests that it's kind

3   of elusive or difficult to kind of get a trial that

4   will be good enough to -- you know, well-conducted

5   enough to show a benefit of a drug if it's

6   effective.

7   With that as sort of a background and sort

8   of a general question includes what do you do with

9   a couple of positive studies in a bunch of

10   so-called negative studies?

11   One of the things you talked about earlier

12   was that you want to approve drugs that will have

13   the purported meeting in the general population

14   basically. And it seems to me if you're sort of

15   twisting the dials in sort of non-replicative

16   manner to try and get a study that'll have assay

17   sensitivity, and you finally do that a couple of

18   times, and you get a significant result, does that

19   make it harder to think it'll have the effect in

20   the population where you're not arranging things, a

21   study of assay sensitivity?

22   To me, it seems a fair question. Doesn't

1   this make it harder to extrapolate to the so-called

2   real world when you now have to do a lot of effort

3   to try and get a study with assay sensitivity?

4   DR. D'AGOSTINO: Dr. Temple?

5   DR. TEMPLE: It's certainly a legitimate

6   worry, but I have to say that we have, for the most

7   part, not tended to dismiss a drug that had a lot

8   of failed studies. I'll give you the reason. The

9   people who know more about this can probably

10   address it.

11   We think that the circumstances under which

12   these trials are done tends to give you a very

13   large placebo effect. If you look at the results

14   of these trials, typically, the drug has an effect

15   of 12 and the placebo has an effect of 10 on

16   whatever Hamilton Depression Scale you're looking

17   at. So there is a huge change in the placebo

18   group.

19   There are a bunch of possible reasons for

20   that. One is that acute depression is cyclical. A

21   lot of people are getting better by themselves.

22   The other is that there is an environment that's

1   warm and nurturing, and there's a lot of

2   improvement, so we see very large placebo effects.

3   We have not reached the conclusion -- maybe

4   you want to challenge it -- that this means that

5   the drug won't work in real life. I have to tell

6   you one of the reassuring features is that in the

7   randomized withdrawal trials, the maintenance

8   trials which don't have those features that support

9   people, the drugs are almost always successful.

10   We've only seen two failed maintenance trials, one

11   with a drug that's approved for something else and

12   the other is gepirone.

13   That provides us some reassurance that these

14   drugs actually have an effect. But your question

15   is about the real-world environment and how that

16   compares to the trial setting are perfectly good

17   questions. But we haven't said it should make us

18   not want to approve antidepressants.

19   DR. D'AGOSTINO: Dr. Conley?

20   DR. CONLEY: Yes, thanks. A couple of

21   things. One is it seems to me as if I heard one

22   thing from the sponsor and different thing from the

1   FDA about three-arm studies. They had presented

2   data that suggested that you had like an 80,

3   90 percent failure rate in three arms. And Bob,

4   you were saying pretty unusual so that's one

5   question I have.

6   The second one is that it seems as if you're

7   conflating differential efficacy between drugs with

8   any efficacy at all, that you do have evidence of

9   efficacy from two positive studies but you have

10   these other ones that you're talking about as being

11   active comparators. But again, you've changed the

12   endpoint. You've done a lot of things that I would

13   say from a leveled playing field standpoint you

14   wouldn't allow a sponsor to do, so I'm trying to

15   speak from a general sponsor out in the field.

16   I do wonder if this is going to markedly

17   make it challenging to think about antidepressants,

18   which you had said you'd like us to think about, if

19   we worry about the number of secondary analyses

20   that aren't prespecified at all, kind of coming

21   back to bite you at the end of the day.

22   One from that sort of policy standpoint, do

1   we need to worry about a lot of new secondary

2   analyses? But then more importantly -- well, not

3   more importantly, just the other one, as I was just

4   really hearing two different things there. And I

5   wonder what you thought about that.

6   DR. TEMPLE: Dr. Yang showed you the rate of

7   failure in active control trials. It's not that

8   high. It's sort of like what we usually expect;

9   about half of all trials fail. She can give the

10   current numbers.

11   We all recognize the fact that using the

12   HAMD-17 instead of the primary endpoint poses a

13   multiplicity problem and that makes people nervous.

14   It makes us nervous if someone wanted to declare

15   drug effective on an endpoint that wasn't a

16   prespecified endpoint.

17   But remember, what we noted here is that

18   there are, yes indeed, two positive studies and

19   then a lot of studies that don't show anything.

20   And the question is how worried should we be about

21   that?

22   Our basic bias is that if a study is known

1   to be a truly failed study, it doesn't count

2   against you. We're not worried about the fact that

3   there -- if there were five failed studies, that

4   really doesn't make you worry about whether the two

5   positive studies are still good.

6   If there's a reason to think they weren't

7   failed studies and had the capacity to show

8   something -- and I realize there's a debate about

9   whether they did or not because of the change in

10   endpoint -- that makes you nervous about the

11   meaning of the two positive studies.

12   That's what's going on here. Nobody thinks

13   there's overwhelming evidence the drug doesn't

14   work. The question is whether there's sufficient

15   evidence to conclude that it does.

16   DR. CONLEY: But it worries me that if we

17   are in meetings with you all about planning how

18   we're going to do a study, that we might, after

19   submission, be held to a different standard because

20   you're going to begin to make up new analyses.

21   DR. TEMPLE: Well, we understand the

22   difficulties of new analyses. And frankly, when I

1   hear about somebody's meta-analysis, I'm

2   overwhelmed with the potential for doing a wide

3   variety of analyses until one of them comes out

4   right. So I worry about that too. We all do.

5   In this case, the contention is -- and

6   Peiling said this also -- we didn't think the

7   HAMD-17 was so far out. It's the endpoint on which

8   the known effective drugs were known to work in

9   depression. It was used as an endpoint in most of

10   the trials of gepirone, so it's not very far out.

11   We think it's perfectly reasonable to be

12   worried about the fact that the studies failed on

13   their primary endpoint and we gave them a different

14   endpoint. That's a legitimate concern, and I don't

15   have any doubt about it.

16   DR. CONLEY: I promise the last question

17   was -- actually, a statement, forgive me. But then

18   you've left me with the worry that the level

19   playing field is gone, that we will have new

20   analyses that we're not expecting to have happen

21   when we're judged for approval.

22   DR. TEMPLE: I don't think that's very

1 likely, and we don't do it very much. One reason

2 is you don't usually see such a large number of

3 what appear to be failed studies or

4 failed/negative. There are a number of negative

5 studies, too.

6   We were trying to see -- I mean, nobody

7 remembers this anymore, but early on in this, we

8 did a meta-analysis of all -- probably using

9 HAMD-17 -- of all of the studies that weren't

10 positive, you know, all 10 of them or something, to

11 see whether they had a lean. We were trying to see

12 if there was some basis for supporting those first

13 two studies, and it didn't show anything. And I

14 believe if you did Dr. Koch's analysis on the

15 studies excluding the two winners, that's exactly

16 what you get.

17   So all of the rest of these data don't seem

18 to show much, and that's what makes us nervous,

19 even though we acknowledge there are two positive

20 studies.

21   DR. D'AGOSTINO: On that note, we're going

22 to break for lunch. Members of the advisory

1   committee that have questions, we'll return at 1:00

2   to get back to the agenda. We'll have an open

3   hearing from 1:00 to 2:00, and after that, we'll

4   move back to the questions.

5   Let me give you a little spiel here about

6   what we're doing. We're going to come back at

7   1:00. Please take any personal belongings you may

8   want with you at that time.

9   Now, can we leave the computers here?

10   MS. BHATT: Yes.

11   DR. D'AGOSTINO: Panel members, please

12   remember that there should be no discussion of the

13   meeting topic during lunch amongst yourselves or

14   with any members of the audience. Thank you.

15   (Whereupon, at 12:19 p.m., a lunch recess

16   was taken.)

17

18

19

20

21

22

1   A F T E R N O O N   S E S S I O N

2   (1:00 p.m.)

3   **Open Public Hearing**

4   DR. D'AGOSTINO: I'd like to begin the

5   material. We're going to have the open public

6   hearing session now.

7   Both the Food and Drug Administration and

8   the public believe in a transparent process for

9   information-gathering and decision-making. To

10   ensure such transparency of the open public hearing

11   session of the advisory committee meeting, FDA

12   believes that it is important to understand the

13   context of an individual's presentation.

14   For this reason, FDA encourages you, the

15   open public hearing speakers, at the beginning of

16   your written or oral statement to advise the

17   committee of any financial relationships that you

18   may have with the sponsor, its product and, if

19   known, its direct competitors. For example, the

20   financial information may include the sponsor's

21   payment of your travel, lodging, or other expenses

22   in connection with your attendance at this meeting.

1   Likewise, the FDA encourages you, at the

2   beginning of your statement, to advise the

3   committee if you do not have any financial

4   relationships. If you choose not to address the

5   issue of financial relationships at the beginning

6   of your statement, it will not preclude you from

7   speaking.

8   The FDA and its committee places great

9   importance in the open public hearing process. The

10   insights and comments provided can help the agency

11   and its committee in their consideration of the

12   issues before them. With that said, in many

13   instances and for many topics, there will be a

14   variety of opinions.

15   One of our goals today is for this open

16   public hearing to be conducted in a fair and open

17   way where every participant is listened to

18   carefully and treated with dignity, courtesy, and

19   respect. Therefore, please speak only when

20   recognized by the chairperson. Thank you for your

21   cooperation.

22   We're going to start with speaker number 1.

1   Will speaker number 1 step to the podium and

2   introduce yourself? Please state your name and any

3   organization you are representing for the record.

4   DR. ZUCKERMAN: Thank you very much. I'm

5   Dr. Diana Zuckerman. I'm president of the National

6   Center for Health Research. Our center scrutinizes

7   research on medical treatments to see what works

8   and what doesn't. We do not accept funding from

9   pharmaceutical companies, so I have no conflicts of

10   interest today.

11   My training is relevant to this meeting, so

12   I just want to mention it quickly. I was trained

13   as a post doctoral fellow in psychiatric

14   epidemiology at Yale Medical School. And in fact

15   as a post doc over 30 years ago, I analyzed data

16   from a key study, one of the first studies to

17   establish the effectiveness of antidepressant

18   medications.

19   I've also been a faculty member at Vassar

20   and Yale and a researcher at Harvard. And I'm

21   currently a member of the Board of Directors of the

22   Reagan-Udall Foundation for the FDA, as well as the

1   Alliance for a Stronger FDA. These are two

2   nonprofit organizations that work to make sure FDA

3   has the resources it needs.

4   I have criticized the FDA in the past at

5   some of these advisory committee meetings, but

6   today, I want to say how impressed I am with the

7   very careful analyses that the FDA has done in a

8   very complicated set of data.

9   In the more than 30 years that I've been

10   looking at these kind of data and attending these

11   meetings, I have never seen so many failed studies

12   and negative studies, and certainly never seen such

13   a disproportionate number of them; and I think

14   that's an important thing to say.

15   In my remarks, I really want to focus on

16   meaningful benefit and the difference between

17   statistical significance and meaningful benefit to

18   patients because I think that's getting lost here

19   today.

20   We have a lot of very capable statisticians

21   and very capable clinicians on this panel. Much of

22   what I say, I think, will be information that you

1   know very, very well, but I want to bring together

2   those two perspectives and try to figure out what

3   we've got here in terms of information.

4   As you know, statistical significance is

5   necessary, but it is not sufficient in proving that

6   a drug has benefits. When studies are designed,

7   when sponsors design studies, the power, the

8   statistical power is very important. But you have

9   to make sure that your study is designed in such a

10   way that you can find the difference that's

11   statistically significant but also clinically

12   meaningful.

13   The statisticians on this panel know if your

14   study has enough patients, just about any tiny

15   difference can be statistically significant, but

16   that doesn't mean it's a meaningful benefit. And I

17   think that Dr. LaVange's analysis misses that point

18   of the meaningful benefit. Yes, there's some kind

19   of statistical significance in there in a

20   meta-analysis or in individual studies, but that

21   doesn't mean that patients are really benefiting.

22   Also, I want to mention I've worked in

1 Congress for a dozen years, and I would be the last

2 one to try to predict what they thought when they

3 wrote laws pertaining to the FDA. But I will say

4 that, at least for the members that I worked for,

5 when they talk about the FDA and when they think

6 about the FDA, what their goal is, is that the FDA

7 standards make sure that drugs are effective and

8 safe, and that they may not know the statistics and

9 they may not know the details of what that means,

10 but they do want meaningful benefit, not just

11 statistical significance. They want to make sure

12 that these drugs really benefit patients.

13   There are really three key issues as far as

14 I can see today in terms of the study design and

15 what the implications are for your votes and your

16 decision-making.

17   The first one is that, as you all know,

18 randomized, double-blind clinical trials are the

19 gold standard for clinical trials. So when a

20 sponsor comes, breaks the blind when the study is

21 done, and then changes the study population by

22 eliminating 32 patients or 40 patients after the

1   fact, that study cannot be considered an

2   appropriate analysis. It's not a legitimate

3   analysis to do that.

4   Of course, that gets into post hoc analyses.

5   That's just one of the most obvious examples of

6   post hoc analyses. But also, in several different

7   examples that you heard today, when the sponsor

8   felt that they didn't like the results of a trial,

9   it became a failed trial for reasons that they

10   described as not being appropriately designed. And

11   yet this was a trial that either the sponsor

12   designed themselves or agreed to conduct. To

13   change that to a failed study design after the fact

14   because they don't like the results is not a

15   legitimate way to look at the data.

16   The last methodological issue, which was

17   raised by Dr. Temple and others, is the one of

18   multiple comparisons. When you've got 13 studies,

19   even if you have a pretty decent significance level

20   equal to -- although not less than 0.01 in these

21   cases -- the chances of that happening just by

22   chance are much higher because there were 13

1 studies. You do enough studies, you're going to

2 have significant effects, especially when several

3 different measures of depression were being used.

4 The one other thing I want to say about

5 meaningful benefit is to keep in mind with

6 antidepressants, the placebo effect is very, very

7 substantial. The goal here is to find drugs and to

8 approve drugs that are better than placebo. If

9 that also means comparing them to other

10 antidepressants to make sure that the studies are

11 designed well and so on, then that's fine, too.

12 But you want drugs particularly that are

13 better than placebo and ideally, of course, drugs

14 that are also better than other antidepressants on

15 the market, although the latter is not the FDA

16 standard.

17 Although depression is very common, we have

18 to think about how serious a disease it is and how

19 important it is for patients to try new drugs that

20 are likely to work for them, not new drugs that

21 might possibly work for some small percentage of

22 patients that haven't been defined, might never be

1   defined.

2   If the company -- and I would encourage them

3   to do that -- if the company thinks that this drug

4   is effective for certain patients, they should

5   figure out how they can predict who those patients

6   are so that the drug can be approved just for those

7   kind of patients where the benefits are likely to

8   be real and meaningful and not just theoretical.

9   In my final conclusion, I just want to say

10   that those of us who've worked with depressed

11   patients and talked to depressed patients know that

12   depression is serious, that it can last for years,

13   if not decades. These are short-term studies. We

14   want to make sure that any drug that the FDA

15   approves for depression is an effective drug.

16   We sometimes are compelled to try to find a

17   reason to approve a new drug because we know that

18   there are patients out there who are hurting, who

19   need treatment. But it doesn't do patients any

20   favor to approve a new drug that is not going to

21   help them, where we have very strong reasons to

22   think that any tiny benefit it might possibly have

1   is not a meaningful benefit.

2   It hurts the FDA's reputation and it hurts

3   all of us who care about patients, as well as the

4   patients themselves, if we encourage the FDA to

5   lower their standards for approval. So I think

6   that the FDA's conclusions in their memoranda are

7   really very clear, that this is a drug that doesn't

8   meet an appropriate standard of effectiveness and

9   efficacy, and I encourage you to look at that

10   meaningful benefit and the overall picture when you

11   vote today. Thanks very much.

12   DR. D'AGOSTINO: Thank you. Will speaker

13   number 2 step to the podium and introduce yourself?

14   Please state your name and any organization you are

15   representing for the record.

16   MS. SORSCHER: I'm just waiting until my

17   PowerPoint's up. Thanks.

18   Good afternoon. My name is Sarah Sorscher,

19   and I am a researcher with Public Citizen's Health

20   Research Group. We are a research-based consumer

21   advocacy nonprofit that has worked in the field of

22   drug and device safety for over 40 years. I have

1   no conflicts of interest.

2   I'm here today because Public Citizen is

3   concerned that the approval of gepirone would

4   represent an unprecedented step backwards by

5   effectively weakening the FDA's standards for

6   approval of these drugs and possibly others.

7   The gepirone development program is notable

8   not only for the sheer number of failed and

9   negative trials but also for the direction of

10   treatment effect trends. This table includes the

11   nine completed efficacy trials that were considered

12   by the FDA in its efficacy review for gepirone ER,

13   but it excludes the three trials that were

14   terminated early and that all have agreed were

15   failed studies.

16   It also utilizes the effect sizes and the

17   p-values from the tables on pages 19 and 20 of the

18   FDA's briefing package, focusing on the preplanned

19   primary endpoints. And I should note that slide 49

20   of the sponsor's presentation, which did the

21   combined p-value analysis, actually presented much

22   lower p-values for these two positive trials for

1   reasons that are unclear. But these are the values

2   there were in the briefing package and on some of

3   the sponsor's slides.

4   Out of these nine completed efficacy trials,

5   four showed trends in the wrong direction with

6   greater observed improvement in the placebo group

7   compared with the group receiving gepirone. And

8   these appear in red at the top of the slide, the

9   four at the top of this line.

10   While two trials did achieve positive

11 results -- these are displayed in blue at the

12 bottom of the slide -- the probability remains, as

13 many have stated, that these positive findings

14 could be due to random chance.

15   Now, gepirone's sponsor has cited four other

16 drugs that have been approved by the FDA despite

17 large numbers of failed or negative trials and

18 suggesting that the FDA is somehow being unfair or

19 inconsistent by denying approval in this case; yet

20 the negative evidence in the gepirone development

21 program is unprecedented even in relation to these

22 drugs.

1   This table is derived from data published by

2   the FDA as part of the Action Package for Approval

3   of Celexa, first of the four to be approved. It

4   includes all of the phase 2 and 3 randomized,

5   placebo-controlled short-term efficacy trials for

6   depression that were considered by the FDA in its

7   efficacy review for that drug regardless of whether

8   they were considered negative or failed by the FDA

9   or the sponsor.

10   Now, some additional short-term trials were

11   not considered by the FDA for this application

12   because they did not involve subjects with

13   depression, lack-to-placebo control, or were

14   terminated early.

15   In this case, three of these trials failed

16   to achieve statistically significant results for at

17   least some doses. Yet, looking at treatment

18   effect, even the failed trials invariably involved

19   positive trends towards effectiveness for Celexa.

20   Two short-term trials showed positive

21   results -- and these are shown in blue again -- and

22   two additional positive long-term relapse

1   prevention trials, which are not depicted in this

2   table, also helps support approval in this case.

3   These are two more tables that were derived

4   from FDA approval packages for Cymbalta and

5   Pristiq, the next two drugs approved in the set of

6   four. Again, these tables contain all phase 2 and

7   3 placebo-controlled short-term depression trials

8   considered by FDA.

9   What's notable in these tables is that while

10   both drugs experienced some trials that failed to

11   achieve significant results, all of these trials,

12   again, displayed consistently positive trends. In

13   addition, each approval in this case included four

14   positive short-term studies where at least one of

15   the doses showed a significant effect rather than

16   two.

17   In fact, out of the four, only one, Viibryd,

18   showed any negative trends during clinical testing,

19   although in this case, it was only two trials that

20   showed negative trends as opposed to the four in

21   the development program for gepirone.

22   Now, Public Citizen is troubled by this

1   approval. Nevertheless, the two pivotal trials for

2   Viibryd included large enough enrollments to arrive

3   at very low p-values, providing some reassurance

4   that the positive results in this case were a true

5   effect and not a statistical fluke.

6   In addition to the unprecedented negative

7   short-term evidence, it's also highly unusual, as

8   the FDA has stated, to see failed relapsed

9   prevention trials like the failed trials submitted

10   in the gepirone application.

11   In fact, out of the 12 drugs cited in

12   today's briefing materials as having completed such

13   trials, failure occurred for only one drug, as I

14   believe Mr. Temple stated before, Fetzima, which

15   was approved in 2013.

16   Notably, that application also included four

17   adequate well-controlled positive short-term

18   efficacy studies, which allowed FDA to conclude

19   that the drug was effective despite the negative

20   finding in the relapse prevention trial.

21   We couldn't do this analysis for all

22   antidepressants approved by the FDA, but FDA did do

1   an analysis that appeared on slide 8 of Dr. Mathis'

2   presentation this morning. And it showed that

3   while there are a handful of drugs with four or

4   five negative/failed studies, none of them had more

5   than five. And for the programs that had four or

6   more failed studies or negative studies, they

7   invariably also had four or more positive studies.

8   So the balance of evidence has consistently been

9   stronger than what's been shown by this

10 application.

11   Finally, we've talked a lot today about the

12 fact that FDA statisticians have found that among

13 the four trials of gepirone that included an active

14 control arm, the active controls performed

15 consistently better than gepirone or placebo.

16 These trends remain troubling even acknowledging

17 that they did not reach statistical significance

18 using the preplanned analysis.

19   As has been stated, each of these multiple

20 failures in the gepirone clinical program is

21 exceptionally rare for an FDA-approved

22 antidepressant. They're especially notable in this

1 case from a drug that comes from a class in which

2 no drug has previously proven effective for

3 depression treatment.

4   If gepirone were approved, it would

5 represent a large step backwards creating an

6 example that will no doubt be used to pressure the

7 agency for approval of future new drugs in spite of

8 lack of substantial evidence of effectiveness as it

9 has been considered up to this date.

10   Depression is a serious condition and

11 requires effective treatment. Regardless of this

12 drug's safety, it should not be approved without

13 additional robust evidence of effectiveness.

14 Public Citizen urges you to vote today against

15 approval of this drug.

16   That's it. I have two minutes left, but I'm

17 going to give those up to the next speaker.

18   DR. D'AGOSTINO: Thank you. Will speaker

19 number 3 step to the podium and introduce yourself?

20 Please state your name and any organization you are

21 representing for the record.

22   DR. SALCEDO: I'm Dr. Beth Salcedo. I'm a

1   psychiatrist in an outpatient practice here in

2   Washington, DC. I'm speaking on behalf of the

3   Anxiety and Depression Association of America.

4   This is an association of clinicians, researchers,

5   and scientists whose mission is the prevention,

6   treatment, and cure of mood and anxiety disorders.

7   I'm a board member, and I've been a general member

8   for over 10 years.

9   In my professional life, I've seen a number

10   of great treatments be made available by the FDA

11   for treatment depressive disorders. I entered med

12   school the year after Prozac came to market. At

13   the same time, HIV positivity was considered a

14   death sentence. Since then, the prevention and

15   treatment of HIV and AIDS has completely turned

16   that around; yet the numbers of people dying from

17   depression remain unchanged.

18   Prozac was revolutionary. Finally, we had a

19   treatment that was both effective and tolerable

20   without the serious side effects like cardiac

21   toxicity that previous medications had.

22   I've seen multiple medications become

1   available since then for treatment of depression

2   and have used them all. The efforts on the part of

3   the FDA to bring new treatments to market is to be

4   applauded, and many have benefited from this.

5   Unfortunately, the treatments we have

6   available to us are not enough. We need to

7   continue to find new and better treatments so that

8   we can change the trajectory of depression in our

9   society.

10   Studies suggest that up to 17 percent of

11   U.S. adults will suffer depression at some point in

12   their lives. It's the most common cause of

13   disability for adults age 14-44 and results in

14   35,000 suicides a year. It's as costly to our

15   population as heart disease or AIDS, and it's

16   projected these costs will rise.

17   A study from the year 2000 suggested the

18   cost of depression to be $83 billion a year. If

19   you take into account other factors and 15 years

20   later, it's upwards of $200 billion a year. And

21   it's important to remember that depression often

22   begins in childhood or early adulthood, and it's

1   chronic and a recurring problem, so these costs

2   recur year after year.

3   Unfortunately, many people don't seek

4   treatment or don't get adequate treatment for a

5   multitude of reasons. The stigma of mental illness

6   is still a barrier for many, even those in the

7   healthcare profession. The treatments that we do

8   have can take time, effort, and energy that those

9   who suffer do not have to give.

10   While good psychotherapies are available,

11   not all who suffer from depression respond or have

12   access. And we know now that the majority of

13   people with depression initiate treatment with

14   their primary care provider who usually cannot

15   offer psychotherapy but can offer medication.

16   The medications we have benefit many but are

17   nowhere near being universally effective. Patients

18   will find it hard to wait the many weeks that may

19   take to achieve the benefit, and after weeks may

20   end up with a second or a third trial of medication

21   because of lack of effect.

22   Often, they will have at least some benefit

1   from medication, and then have to tolerate the

2   various side effects that often accompany the

3   medications. In fact, it's the knowledge of the

4   potential side effect burden of medications that

5   often prevents people from seeking them out as

6   possible treatments.

7   I've experienced many patients over the

8   years who do not respond until third, fourth, or

9   even 10th medication trial. There are more than

10   just a few roadblocks done for people seeking

11   effective treatment for depression.

12   It makes obvious economic sense why we need

13   to have so many effective treatments for

14   depression. It's also important to think about the

15   personal burden shouldered by patients and their

16   families living with this disease. My patients who

17   cannot achieve remission from depression suffer

18   every day of their lives. They describe it as

19   feeling like they're walking through quicksand all

20   day, every day. And everything that they try to

21   achieve takes way more effort than it should.

22   They cannot be fully present at work or at

1   home. Their relationships suffer. They can't

2   function as they feel they should at work or in

3   school, and they often feel they're failing as

4   parents, spouses, partners, or family members.

5   They're much more likely to divorce, achieve less

6   in school, and become unemployed.

7   It's all too often that patients tell me

8   that they would rather have cancer than depression.

9   It's common for them to think that death would be

10 preferable to living with depression, and many

11 contemplate suicide on a regular basis. The

12 suffering is chronic and debilitating, and the

13 ripple effect of that suffering is enormous.

14   We are in desperate need of more treatment

15 options for our patients. More better effective

16 treatments that are tolerable will lead to more

17 people seeking treatment and that hopefully will

18 finally realize for us a better treatment response

19 overall.

20   Please do consider this when you think about

21 the approval of new treatments for depression.

22 Thank you.

1   DR. D'AGOSTINO: Thank you.

2   Will speaker number 4 step to the podium and

3   introduce yourself? Again, please state your name

4   and any organization you are representing for the

5   record. Thank you.

6   DR. SIMON: Members of the advisory

7   committee, colleagues at the FDA, ladies and

8   gentlemen, I'm Dr. James Simon, clinical professor

9   of obstetrics and gynecology at the George

10   Washington University, School of Medicine in

11   Washington, D.C. I'm here today of my own

12   volition, at my own expense, and I have no

13   financial relationship with the sponsor whatsoever.

14   I'm not a psychiatrist, but I work with

15   large numbers of patients who use antidepressants,

16   anti-anxiety agents, and other mood stabilizing

17   agents, and I do so nearly every day. They see me

18   because I'm an international expert on sexual

19   dysfunction in women as are several others in the

20   room.

21   In this capacity, I see women who have

22   desire, arousal, and orgasmic dysfunction

1   temporally related to or caused by their SSRIs,

2   their SNRIs, their SPARIs, their SNRIs, and their

3   NDIs, or other of the burgeoning alphabet soup of,

4   let's call them mental health medications.

5   While these agents appear to adequately

6   address the patient's depression, anxiety, and/or

7   obsessive compulsive disorders, they commonly cause

8   new onset sexual dysfunction for which we have no

9   current approved antidotes.

10   Paroxetine and fluoxetine, used in two of

11   the three 3-armed trials discussed here, are of the

12   worst offenders in this regard. Off-label

13   approaches to treating these sexual side effects

14   abound. They range from drug holidays with the

15   inherent risk of suicide and relapse and the

16   addition or substitution of bupropion, buspirone,

17   or other treatments like alpha 2 antagonists, or

18   changes to tricyclics or MAOIs, all with their own

19   side effects and risks.

20   The additions of testosterone for desire

21   and/or the PDE5 inhibitors like Viagra, Levitra,

22   and Cialis round out the most common approaches

1   that I see in everyday life in the office, and none

2   of these are approved for use in women whatsoever.

3   Some of these techniques actually have small

4   studies documenting modest benefits, but additional

5   treatments for depression and anxiety with fewer or

6   no sexual side effects are desperately needed.

7   Gepirone appears to be somewhat different

8   than the other available antidepressants and

9   anti-anxiety agents with regard to sexual side

10 effects. Several papers presented at the

11 International Society for the Study of Women's

12 Sexual Health, an organization I am currently

13 secretary of, demonstrated fewer adverse events as

14 did the sponsor in their presentation as it related

15 to gepirone and its potential advantages in

16 avoiding sexual side effects.

17   Having choices is a good thing. There's no

18 antidepressant or anti-anxiety agent that works for

19 everyone, nor is there any single agent that is

20 optimal for everyone as the previous speaker so

21 eloquently documented.

22   None of these agents are side effect-free.

1   In our current approaches to treating depression

2   and anxiety disorders, the clinical focus has

3   actually shifted from efficacy to managing side

4   effects.

5   Although I trust in the judgment of the

6   FDA's advisory committee, I sincerely hope the

7   committee can view the medication in its entirety,

8   both the benefits and the side effects or risks,

9   remembering that one of the most basic of human

10   needs, healthy sexual function, is a common

11   casualty of antidepressant/anti-anxiety therapy.

12   Options with fewer or no sexual side effects are

13   desperately needed. Gepirone may fill that void.

14   Thank you for your attention.

15   DR. D'AGOSTINO: Thank you.

16   Will speaker number 5 step to the podium and

17   introduce yourself? Please state your name and any

18   organization you are representing for the record.

19   Thank you.

20   MS. BHATT: There's no speaker.

21   DR. D'AGOSTINO: There's no speaker

22   number 5?

1   DR. ISRAEL: My name is Steven Israel. I'm

2   a psychiatrist in private practice, been in

3   practice for 30 years, both in inpatient and

4   outpatient settings. And I have no financial

5   relationship with the sponsor.

6   As a psychiatrist who has worked with

7   patients for 30 years, I've witnessed the damage

8   done by persistent mood and anxiety disorders, not

9   only the dramatic and over, but the more subtle and

10   insidious.

11   They are the destruction of one's inner

12   sense of self-worth, missing the joy of being with

13   family and good friends, the tension and conflict

14   that can infect interpersonal relationships. All

15   these can lead to alienation, despair, a loss of

16   the possibilities of life, a sense of futility and

17   pessimism that pervades patient and family that can

18   cause hurt across generations.

19   In my experience, some medications certainly

20   have been life-saving. But for many patients, they

21   don't work well enough. They still suffer. They

22   experience intrusive negative thoughts, anxiety,

1   panic attacks, obsessions and compulsions, and we

2   feel driven to augment and augment the augmenter.

3   We run the risk of piling on side effects leading

4   to even morbidity.

5   As a psychiatrist, I feel compelled to try

6   off-label innovative strategies that have limited

7   likelihood of efficacy and increased risk of severe

8   adverse side effects like blood pressure medication

9   for mood cycling and dopamine agonists for

10 depression or sexual side effects.

11   Often, the patients feel the pain of side

12 effects is not worth the gain of the hope for

13 therapeutic benefit. They'll give up on medication

14 because of agitation, dysphoria, loss of libido,

15 weight gain, feeling like their emotions have been

16 suppressed. They may not be devastated so much by

17 the slings and arrows, but they feel robbed with

18 their peaks of happiness, their edge, with their

19 capacity for righteous indignation.

20   Sometimes patients will agree when I tell

21 them not to be afraid of the medication; it will

22 give them a chance to spend more of their time

1   being their best selves. But with some

2   justification, others won't take the medication

3   because they feel it robs them of essential

4   elements of their identity.

5   Our medications are still primitive and

6   thought to exert their affected sites somewhat

7   removed from where the actual problems reside. In

8   addition, our diagnostic scheme is similarly

9   primitive and can refer to a multiple of possible

10 underlying pathophysiologies.

11   It makes sense then that for many of our

12 patients, our medications are going to be partially

13 effective or cause the unintended side effects.

14 This would also argue for having available

15 medications that have variability in their

16 mechanism of action.

17   For the practicing psychiatrist, it might

18 require sequential trials of many medications even

19 from the same class but with differences in

20 specific mechanisms to find that optimal agent; or

21 it might require combining medications of

22 complimentary mechanisms to achieve the desired

1   results.

2   For example, one of my patients was driven

3   to attempt suicide by the thought-slowing

4   dysphoria-inducing effect of one atypical

5   neuroleptic, while another atypical was the magic

6   bullet that enabled her to resume a productive

7   life.

8   Another patient failed to respond to SSRIs

9   and all but one of the SNRIs. But it was the one

10   agent that had the strongest norepinephrine

11   relative to serotonin reuptake inhibition that made

12   all the difference. Still, another patient

13   experienced vengeful obsessions about his romantic

14   overtures as being rejected. Despite a high dose

15   of an SSRI and good dose of an atypical

16   neuroleptic, his anger would be triggered in social

17   situations causing fear and costing him jobs and

18   friends. Adding a 5-HT 1 partial agonist brought

19   these alarming symptoms under control, allowing the

20   patient to hold a job and expand his social

21   repertoire.

22   We have these situations in which current

1   agents don't work optimally and at which there's a

2   wide variability in individual response to

3   different agents. And we need a process that is

4   sensitive to the challenges facing practicing

5   psychiatrists and receptive to the introduction of

6   new agents and new indications.

7   I appreciate the opportunity to offer these

8   remarks today. It's an honor to participate in a

9   system that strives to protect and provide maximum

10 benefit for our patients. I'm confident we will

11 recognize the challenges facing psychopharmacology

12 today. The armamentaria will continue to expand

13 and offer us greater opportunity for therapeutic

14 success.

15   DR. D'AGOSTINO: Thank you. Will speaker

16 number 6 step to the podium and introduce yourself?

17 Please state your name and any organization you are

18 representing for the record. Thank you.

19   (Pause.)

20   DR. D'AGOSTINO: Six is trying to find a

21 parking space.

22   (Laughter.)

1   DR. D'AGOSTINO: We'll go to 7? Will

2   speaker number 7 step up to the podium and

3   introduce yourself? Please state your name and any

4   organization you are representing for the record.

5   Thank you.

6   DR. SALCEDO: Thank you. I'm reading a

7   statement by Kenneth J. Weiss, MD who is a clinical

8   trials investigator and clinical professor of

9   psychiatry at the University of Pennsylvania,

10   College of Medicine.

11   "It is my pleasure to supply an endorsement

12   for gepirone ER, a medication targeted for the

13   indication of major depressive disorder. As

14   medical director and principal investigator of

15   Delaware Valley Research Associates, I had an

16   opportunity to investigate psychotropic agents

17   under various conditions from phases 2 through 4,

18   including those with putative antidepressant

19   properties.

20   "Accordingly, I was involved in two gepirone

21   clinical trials conducted under randomized,

22   double-blind conditions for subjects with major

1   depressive disorder. It is well-known that

2   depression and the risk of suicide are major public

3   health issues.

4   "While it is true that many cases of

5   depression are treatable with non-pharmacologic

6   means, it is also true that the majority of

7   prescriptions for antidepressant medications are

8   written by primary care physicians and other

9   non-psychiatric prescribers unable to provide

10   psychotherapy. In other words, antidepressant

11   medications will remain the most prevalent

12   intervention. Thus, if patients fail on standard

13   treatments, unless there are alternatives, the risk

14   of morbidity and mortality rise. Alternatives are

15   essential.

16   "Before addressing gepirone, I should point

17   out that at the time of these trials, as well as

18   now, clinical antidepressant therapy has been

19   dominated by selected serotonin reuptake

20   inhibitors, SSRIs, in the same manner that

21   tricyclic compounds had done before.

22   "For patients intolerant of these classes of

1   drugs or who did not experience a robust response,

2   we were continuously hopeful for an alternative.

3   Having spoken with other gepirone investigators at

4   the time, it appeared there was a strong signal

5   toward to gepirone's efficacy in major depression.

6   "We did not study gepirone in

7   treatment-resistant depression. It was exciting to

8   employ a compound from the azapirone class given

9   its different mechanism of action and excellent

10   tolerability.

11   "Speaking from my site, we had a good

12   separation of gepirone from placebo, as well as

13   excellent tolerability. Though as an investigator

14   running a randomized clinical trial, I cannot speak

15   with certainty about efficacy. A group of us

16   published results in a peer-reviewed journal.

17   "We reported efficacy of gepirone ER against

18   placebo in the 8-week trial. Just as important,

19   side effects were mild and there was an indication

20   that compared to SSRIs, subjects were not troubled

21   by medication-induced sexual dysfunction.

22   "To summarize my experience with gepirone ER

1   in the broad context of treatment of depression, I

2   offer the following to the advisory committee:

3   "Current drugs treating depression have

4   little to offer in terms of novel mechanisms of

5   action. Gepirone ER has shown efficacy in treating

6   major depressive disorder at least over eight

7   weeks. The minimal side effects of gepirone ER

8   render it a standout among antidepressant choices.

9   "Because many persons with depression fail

10   on standard treatments or have troubling adverse

11   effects, gepirone ER would be an important

12   breakthrough product. It is, therefore, essential

13   that the FDA give careful consideration to proven

14   and safe antidepressant products so that citizens

15   with a disabling condition have the broadest range

16   of options."

17   DR. D'AGOSTINO: Thank you.

18   Will speaker number 8 step to the podium and

19   introduce yourself? Please state your name and any

20   organization you are representing for the record.

21   Thank you.

22   DR. AMSTERDAM: Yes. Thank you. My name is

1   Jay Amsterdam. I'm professor of psychiatry at the

2   University of Pennsylvania and Director of the

3   Depression Research Unit and have been at the

4   University of Pennsylvania engaged in

5   psychopharmacology research since completing my

6   NIMH post doctoral fellowships in

7   neuro-psychopharmacology back in 1977.

8   With the exception of travel expenses

9   provided to me to come to this meeting today from

10   the folks at ISS, I have no conflict of interest or

11   financial disclosures to make with the sponsor or

12   any other pharmaceutical company.

13   The comments that I'd like to make are

14   really adlib comments. I think that my greatest

15   strength that I can bring to the discussion is the

16   fact that I have been involved in gepirone research

17   trials since the early '80s, I estimate around

18   1983. And this has been the case up until as

19   recently as 2006, 2007. In fact, my site was one

20   of the prime sites for the 004, the actually

21   infamous 004 and 006 studies.

22   As I listened very carefully to the

incredibly elegant presentations this morning, both

pro and con, I come away with two somewhat

disturbing facts. One regards the 004 and 006

studies and related issues of assay sensitivity.

       If it had not been for the fact that these

were atypical studies with atypical patients, then

analyzed post hoc with an atypical analysis, it may

have been that we wouldn't be here today because

two studies that one might regard as failed were

now called negative with the evidence for

gepirone's effectiveness now pointing against their

two positive trials.

       I had the -- I don't know whether it was

good fortune and a privilege of being present at

Bristol -- I'm sorry -- at Organon when they were

organizing the 004 and 006 studies and can say now

with some pride today that I actually was at the

center and told them that I advise not to do the

studies.

       But on the other hand, you can see not only

didn't they listen to me but they went ahead and

did it. I also said that as an investigator, my

1   job is to test the null hypothesis and to show that

2   placebo doesn't work, so you should make me an

3   investigator.

4   Anyway, we did those studies and are the

5   data as we hear it today. What I would say is

6   this, that regarding the assay sensitivity, it is

7   somewhat troubling to me that, for example, in the

8   006 study, the reason that the FDA is now viewing

9   it as a negative rather than a failed trial is

10   because Paxil, paroxetine, appeared to

11   differentiate itself from placebo using the HAMD-17

12   in a post hoc analysis, and these were in patients

13   with atypical depression.

14   This is a real hornet's nest because the

15   symptoms of atypical depression do not really

16   appear on the HAMD-17, and the other thing is that

17   Paxil took 11 pivotal trials in order to get their

18   regulatory approval for their two positive trials.

19   So is that really the gold standard?

20   I put these comments out for the committee,

21   both pro and con, just to give you a perspective

22   from someone that had been in the trenches and

1   actually saw these patients. It's a difficult

2   issue that I'm sure will eventually be resolved

3   favorably.

4   I want to thank the committee very, very

5   much for this opportunity.

6   **Clarifying Questions to Sponsor or FDA**

7   DR. D'AGOSTINO: Thank you.

8   The open public hearing portion of the

9   meeting has now concluded, and we will no longer

10   take comments from the audience. The committee

11   will now turn its attention to address the task at

12   hand, careful consideration of the data before the

13   committee as well as the public comments. We'll

14   switch now back to the clarifying questions to the

15   sponsor or the FDA.

16   We had a number of people who had raised

17   their hand, and I didn't recognize them to try to

18   handle the lunch appropriately.

19   Dr. Mann, you would like to start?

20   DR. MANN: Thank you very much. I have a

21   question for Dr. Temple and Dr. LaVange. The

22   question is that as I understood your

1 presentations, you are attempting to address the

2 idea that there should be some penalty for going on

3 doing studies or conducting studies that are

4 negative or neutral potentially. At least, that

5 seems to be the thrust of it. That's my first

6 question.

7   What weighting should be given to the

8 studies that are not two of the pivotal studies,

9 the two pivotal positive studies? What weighting

10 should be given to those studies in making a

11 decision about gepirone?

12   DR. TEMPLE: Well, let me start. Let's be

13 clear. If we think a study did not have assay

14 sensitivity and was a failed study, we would not

15 count that against anybody. You don't expect too

16 many of them. They're disappointing, but they

17 don't show -- they don't convey anything negative

18 about you.

19   We do think that a study that has assay

20 sensitivity and is negative for the test drug, that

21 isn't lethal, but the question that I raised is if

22 you have two positive studies -- if you have two

1   studies in which gepirone looks effective

2   and -- depending on how you count it -- nine

3   studies in which it couldn't show any

4   effectiveness, you have to raise the question, what

5   is the chance the two positive studies is a chance

6   occurrence and not real?

7   That's easier if there's one out of four or

8   five. If there's two, then even two out of nine

9   has a low nominal p-value, however you want to

10   calculate it. But the more of those there are, the

11   more it cast out, in my mind anyway, as to the

12   likelihood that the two positive studies are real.

13   DR. MANN: But I sense that you are moving

14   towards a notion of some kind of offsetting metric.

15   I don't understand what that metric might be.

16   DR. TEMPLE: No. My version of it is very

17   simple. It's easiest for a one-study consideration

18   where everybody understands that if you do five

19   studies and test at a P of 0.05, the chance that a

20   drug with no effect all will be positive in one of

21   them is something in the neighborhood of

22   25 percent. That's why we like to have more than

1   one study showing effectiveness.

2   If there are two positive studies, the

3   chance that that result is due to chance is much

4   lower, much lower. We recognize that. But, for

5   example, if there are two out of 40, you'd

6   immediately say, Oh, well, that could be a chance

7   occurrence.

8   The question is where we are on that pattern

9   here. If it's two positives out of 14 or 12 or

10   something like that, there is a certain chance the

11   apparent positive result is a chance occurrence.

12   Mary Johnson calculated that a long time ago in an

13   early Fabre-Kramer submission.

14   The nominal p-value of the finding is still

15   less than 0.02, but that's not as good as we

16   usually expect for approving a drug. We expect two

17   positive studies and a lower overall p-value -- I'm

18   not going to say what it should be, two out of

19   three, two out of four, two out of five. Those are

20   all plenty good. When you get up to two out of

21   nine, you start to wonder whether this could be a

22   chance occurrence. That's the reasoning.

1   DR. MANN: Okay. Thank you. I have a

2   question.

3   DR. TEMPLE: Unless you asked both of us --

4   DR. MANN: No, I have a different question.

5   My question is this, and I know this may sound odd

6   coming from a non-statistician. Wouldn't one

7   approach for trying to achieve what Dr. Temple is

8   describing, which is a kind of melding of the data

9   from the positive studies and the negative

10   studies -- I'm not talking about the failed studies

11   but maybe the failed studies as well, but leave

12   them aside -- it would be to consolidate the entire

13   data set into a single patient level data set?

14   Then you can say, well, let's only take the

15   patients whose HAMD-17 is more than -- whatever you

16   like. Then let's take into account the drug dose.

17   Aren't you going to get a much more powerful

18   statistical approach if you do that rather than

19   look at a piecemeal by a study level kind of

20   meta-analysis and try to piece all these little

21   bits together which don't fit?

22   DR. LaVANGE: First, I'll just comment on

1 the metric question.

2   DR. D'AGOSTINO: Lisa, sorry, would you give

3 your name?

4   DR. LaVANGE: I'm sorry. Lisa LaVange,

5 director of Office of Biostatistics in CDER. Your

6 first question to both of us about the metric, I

7 actually would like to not think about a metric. I

8 don't like the counting trial approach, as I've

9 already mentioned for the reasons that I said.

10   I think that if we have a metric, you could

11 see a place where you have five negative studies

12 that doesn't -- you know, that's okay; six is not

13 okay. I mean I just can't imagine that we would

14 want to make decisions that way. I think we have

15 to look at all the evidence that's given us, and we

16 have to do the best we can.

17   I think that the meta-analyses, now to your

18 second question, are helpful for a variety of

19 reasons. They give some synthesis of what we've

20 seen across all the studies. They were presented

21 by the sponsor in the same way that we view them at

22 the FDA.

1   We look at the positive studies and the

2   negative studies and see how much impact the

3   negative study has on the effect that the two

4   positive studies had. And you saw it marched down

5   from minus 2 and a half to half of that when you

6   put the negative trials in. And if all 12 trials

7   were considered, including those that were

8   considered failures and shouldn't count against, as

9   Dr. Temple said, it goes down below 1 to something

10   like minus 0.7.

11   That's a good sensitivity analysis to see

12   just how negative the negative information is. But

13   I think we have to take all of that evidence into

14   account when we're making an approval decision.

15   With respect to the meta-analyses -- and

16   we're actually working on a guidance on

17   meta-analysis for safety, not efficacy -- we don't

18   normally accept meta-analyses for efficacy because

19   we don't believe you can replace the two

20   confirmatory studies that are required and that

21   have replication with a meta-analysis. And a big

22   reason is that replication is much more important

1   than just the very small p-value that it gives.

2   In this case, the meta-analyses are not

3   being used to stand in for the two studies. The

4   sponsor has the two studies. They're not being

5   used to stand in for a third study. They're being

6   used as a sensitivity analysis to see how much of

7   the effect is lost when you add in all of the other

8   information.

9   Yes, you can do a patient level

10   meta-analysis as long as you're careful to control

11   for the randomization scheme of each study;

12   otherwise, you can invoke some unusual findings

13   like the Simpson's paradox, for example.

14   The study level meta-analysis operated on

15   effect sizes that came from a patient level

16   analysis, so they were able to adjust for

17   covariates or do some other things with the ANCOVA

18   models. I'm not overly worried about too much loss

19   of information in those analyses. In fact, the

20   inverse weighted meta-analysis on the mean method

21   will give you very similar results than as the

22   patient level analysis. We're fine with either one

1   of those approaches.

2   But if you want to put all the patient level

3   data together and look at different subgroups of

4   patients as you suggested, yes, that could be very

5   helpful to see if the drug works. And I think the

6   sponsor presented meta-analyses where they took out

7   the less severe patients with the HAMD less than 20

8   criteria at baseline, so that was one such analysis

9   that you could do. But yes, they could be useful.

10   DR. TEMPLE: Just one --

11   DR. D'AGOSTINO: Related to this, Bob?

12   DR. TEMPLE: Just one more point. My

13   interest in the other studies is to help me

14   understand whether the two positive studies could

15   be a chance occurrence. When I do a meta-analysis

16   with the remaining studies, I don't include the two

17   positive studies. I already know they were

18   positive.

19   We did that a long time ago. Actually, it

20   was to see if we could salvage the drug. We did it

21   seven or eight years ago, and it didn't really show

22   anything. There wasn't even a lean.

1   Dr. Koch didn't show that result without the

2   two positive studies, but I know the mean effect

3   will be very, very modest if you count all of the

4   others. And he may have done that and knows what

5   it is.

6   It seems to be that's an important part of

7   the question. It's not irrelevant to know what

8   throwing them all together does. But remember, a

9   major interest was to see if, given all the failed

10   and negative studies, those two positive studies

11   could be a chance occurrence. One way to get some

12   insight into that is to see whether the others at

13   least lean, and they don't.

14   DR. D'AGOSTINO: Dr. Koch, do you want to

15   comment on that?

16   DR. KOCH: Yes, during lunch I asked

17   Dr. Johnson to apply the Fisher combined p-value

18   method because that's a spreadsheet calculation.

19   When we applied it to all 10 of the other studies,

20   excluding 01 and 07, and we focused on the

21   population that had the HAMD greater than or equal

22   to 20, we got a p-value of 0.24 one-sided. Now,

1   that's a lean. It's a weak lean, but it still is a

2   lean.

3   If we do the eight other studies, excluding

4   01, 04, 06, and 07, that one-sided p-value

5   decreases to 0.121 in that same population. And if

6   we additionally exclude 17, where 04, 06, and 17

7   are the ones which are in really different

8   objectives than the other studies, it drops to

9   0.072.

10   So depending upon how you choose to look at

11   it, there is a lean, and the lean, if you look at

12   all 10 other studies in the full overall population

13   without the exclusion of those less than 20, it's

14   still 0.31. So there's a lean; it's a weak lean, I

15   agree with that. But it's still lean. At least,

16   it's in the right direction and to some extent,

17   comfortably in the right direction.

18   DR. D'AGOSTINO: Dr. Koch, could you

19   remind -- what's the magnitude of the number of

20   people that were excluded relative to the --

21   DR. KOCH: Well, when you do the exclusion

22   of those patients with baseline less than 20 in the

1   04 and the 06 studies, because of the populations

2   they were designed for, you're excluding about

3   50 percent of the patients from 04 and 06.

4   Part of the difficulty of interpretation for

5   04 and 06 is that nearly half of their patients

6   were in this less sensitive population. The 17

7   study, I think, had a reasonable number of patients

8   under 20 as well. I don't have that number at my

9   fingertips, but I could get that later.

10   DR. D'AGOSTINO: I just wanted to get an

11 idea of the magnitude.

12   DR. KOCH: All of the other studies, it's a

13 very small number.

14   DR. D'AGOSTINO: Small number, yes.

15 Dr. Stein, do you have a comment or question?

16   DR. STEIN: Yes. I guess it's kind of along

17 the same lines as Dr. Mann's questions. They're

18 sort of questions for the agency. I'm just trying

19 to sort of understand -- the metrics is a good

20 word.

21   It seems there's agreement that there's two

22 positive trials, and my understanding is that has

1   sort of been the standard up until now for saying a

2   drug is effective. So I understand that there's

3   additional data available that now it's being

4   suggested we might want to look at in terms of

5   negative or failed trials, but there's no agreed

6   upon metric for doing that.

7   I don't understand the rationale for saying,

8   we've got two positive trials, but now let's look

9   at the non-positive trials to make sense of the

10   positive trials. It's kind of like if you're not

11   going with the two trials, you're using a new

12   metric, and where do you look?

13   You look in academia or you look at

14   something like when people want to ask the question

15   overall. I think you got to look overall. Across

16   all the gepirone randomized controlled trials that

17   have been done, does it look like it's effective?

18   You don't just look at the negative trials.

19   You look at all of them in their totality, and

20   you're now using a different metric. A

21   meta-analysis seems as good as any. It's what most

22   of the agencies in many other countries we use to

1   determine that.

2   As a journal editor, if somebody came to me

3   and said, we've done an analysis of all the drug X

4   trials, and there's two positive and five negative,

5   I'd sort of go, "Well, that doesn't tell me

6   anything. You've got the data. Go take a look at

7   it. Do a meta-analysis."

8   I've seen that. I haven't seen anybody

9   argue about the legitimacy of the meta-analysis

10   we've been presented with, and I've heard some

11   argument that that may be a good way to look at

12   things. When you look at them in their totality,

13   looks like the drug is effective.

14   Just to get back to the discussion of the

15   negative trials, that also seems, to me, kind of

16   like cherry-picking because my understanding is,

17   from what Dr. LaVange presented, is that in other

18   areas where you don't want to use placebo anymore,

19   those trials are typically done comparing to an

20   active comparator as a non-inferiority trial.

21   If one were to, now post hoc, judge these

22   trials in the context of non-inferiority, well,

1   then you have to go, well, what would the margin

2   have been and -- you know, I don't know. It

3   might've been two points or something. I think you

4   would've concluded that gepirone is not inferior to

5   the active drug because the difference is very,

6   very small.

7   I'm just not understanding the -- I don't

8   see what the problem is here, I guess is the short

9   way of saying it. I don't see what the debate

10   is -- why the debate is still going on.

11   DR. TEMPLE: Let me clarify your question.

12   In three of the trials where there was an active

13   control and placebo, gepirone was inferior to the

14   active drug.

15   DR. STEIN: But the margin was very, very

16   small.

17   DR. TEMPLE: No, it wasn't. No, it wasn't.

18   It was about 1 and a half, not so different from

19   the effect of the drug.

20   DR. STEIN: And this is all sort of second

21   guessing. But if one had set up a non-inferiority

22   trial for gepirone versus an active comparator,

1   what margin would it have needed to fail by to

2   conclude that it actually is inferior?

3   DR. TEMPLE: Yes. That's a very good

4   question. That's why you never do equivalence

5   trials or non-inferiority trials with symptomatic

6   conditions because you can't say what the effect

7   is.

8   But I just want to say again, the effect

9   size difference, if you look at it, what it was

10   between gepirone and the active drug was almost the

11   usual difference that an effective antidepressant

12   has versus placebo, which is in the neighborhood of

13   2, 2 and a half. These were about 1 and a half.

14   That's not so small. To be inferior to an active

15   drug, I'm just telling you, if we've ever seen it,

16   we've seen it maybe once in one of these active

17   control trials. It's very unusual.

18   DR. D'AGOSTINO: Dr. Ionescu?

19   DR. IONESCU: Hi. I'm Dawn Ionescu. Just a

20   few brief questions for the sponsor. When the

21   patients were recruited for the study, was there an

22   external rater used for any of the metrics, or were

1   they all seen by the study doctor who was giving

2   the medication and rated by the same doctor?

3   DR. THASE: Michael Thase from the

4   University of Pennsylvania. The last study

5   enrolled the last randomized patient, circa 2005.

6   This was five years before the independent rater

7   standard became kind of the new gold standard to

8   ensure appropriate pre-treatment assessment.

9   DR. IONESCU: Okay. Thank you. One other

10   question. As far as the trials go, were all of the

11   trials in treatment-naïve patients or were some

12   trials in treatment-resistant depression?

13   DR. THASE: There were no studies of

14   treatment-resistant depression. Each study varied

15   in the proportion of patients that had to be taken

16   off an effective antidepressant. I would guess,

17   based on other experience, no more than a quarter

18   of the patients would be considered more

19   treatment-resistant, but that's a guess, not a

20   fact.

21   DR. IONESCU: Thanks.

22   DR. D'AGOSTINO: Victor, are you on the

1   phone? Do you have a question?

2   DR. DE GRUTTOLA: Yes. Can you hear me?

3   DR. D'AGOSTINO: Yes. Sounds clear.

4   DR. DE GRUTTOLA: Now, I'm going to mute.

5   Is this coming through? Hello?

6   DR. D'AGOSTINO: It's coming clear.

7   DR. DE GRUTTOLA: Okay. Thanks. I had a

8   comment and a couple of questions. First of all, I

9   think the reason we're perseverating on this set of

10 analyses is the ambiguity, as Dr. Temple said,

11 about whether we can make inference that the drug

12 is actually beneficial, which is different from

13 saying two of the nine studies were positive.

14   I think that an issue with the meta-analysis

15 is that although, when including all of the

16 studies, there was a relatively small p-value, not

17 impressively small given the amount of information

18 that was included, but you also saw an effect size

19 that got quite small as well. And I would like to

20 ask whether that is a concern, given that the

21 totality of the evidence doesn't appear to be

22 producing a robust effect.

1   Second, I wanted to comment on the analyses

2   that were -- the HAMD-17, which showed the

3   inferiority of the gepirone. Those were, as

4   pointed out, secondary analyses. But I think that

5   where secondary analyses are particularly

6   appropriate is where there's ambiguity about

7   results and their meaning. Those analyses seem to

8   me to be quite relevant.

9   But I had one question there about the

10   site-by-treatment interaction that was included in

11   one of the analyses by the sponsor but was not

12   included in the analysis by the FDA. So I just

13   wanted to ask whether that was a protocol-specified

14   analysis, again the site-by-treatment interaction,

15   or whether that was done post hoc.

16   DR. D'AGOSTINO: Yes.

17   DR. LaVANGE: Lisa LaVange, Office of

18   Biostatistics. I'll just take the last part of

19   that. The study, I think, you're referring to is

20   053. The prespecified analysis included the

21   site-by-treatment interaction, but it also included

22   a test of treatment effect that weighted the sites

1   equally. In that study, the sites were very

2   disparate in their sizes.

3   We felt like that analysis was a not fair

4   representation of what was happening in that study,

5   so our post hoc analysis took out the

6   site-by-treatment interaction. We could've kept

7   the site-by-treatment interaction in and just used

8   a contrast statement that weighted the effects at

9   each site proportional to the number of patients,

10   which would've given similar results probably with

11   a smaller residual variance, so it might've even

12   been better.

13   But we felt like in that case, the analysis

14   really wasn't showing what was going on in that

15   study, and it was prespecified, but the post hoc

16   analysis had a change of method.

17   I think in three other studies, we changed

18   the analysis method slightly, but I don't think it

19   had an impact on the results. The biggest impact

20   was when we change the endpoint from the HAMD-25 to

21   the HAMD-17.

22   DR. D'AGOSTINO: Thank you. Did you hear

1   that?

2   DR. DE GRUTTOLA: Yes. Thank you. That was

3   good. Thank you.

4   Then one other question, again, about the

5   effect size and concern about the fact that

6   although the p-value remained relatively small in

7   the meta-analysis, the effect size is really quite

8   small. Does that raise concerns with either the

9   sponsor or the FDA?

10   DR. D'AGOSTINO: Bob?

11   DR. TEMPLE: This is Dr. Temple. We have

12   somewhat skitzy attitudes here. In the two studies

13   that were nominally positive, the effect size was

14   in the usual range, so we were not particularly

15   worried about it.

16   It's perfectly obvious that in the studies

17   that didn't show an effect, either failed studies

18   or negative studies, the effect size is going to be

19   much smaller. We know that. So it's true when you

20   do the meta-analysis and throw all those studies in

21   with the two positives, the mean effect size

22   plunges.

1   That was not our main worry. My main worry

2   is whether the positive results are real or a

3   chance occurrence. So the fact that the negative

4   studies and the failed studies show a smaller

5   effect, to me, is neither here nor there because my

6   primary worry is whether there's an effect at all.

7   I don't think we take into account the fact

8   that negative studies obviously show a smaller

9   effect. We consider them negative, not supportive,

10   and so on. But I'm not sure how much to make of

11   the effect size.

12   Similarly for our failed study where you've

13   concluded that the environment or whatever it is

14   can't show in effective drugs that are known to be

15   active, we don't take into account the actual

16   putative measured effect size because we've

17   concluded the study didn't have a capacity to show

18   anything. So I don't know if that helps.

19   DR. D'AGOSTINO: Dr. LaVange?

20   DR. LaVANGE: To add to Bob's response and

21   also to add something, Professor Koch gave a

22   meta-analysis p-value using Fisher's method where

1   you operate on the p-value itself for the negative

2   studies without the positive studies. Bob just

3   mentioned the need, I think, to look at the

4   negative studies without the positive studies.

5   The sponsor did do this in their 2007

6   submission. Back then, they were looking at

7   five negative studies. Two of those, later, were

8   taken out of what was presented today because of

9   the different population. But with the five

10 negative studies at that time, the effect was about

11 minus 0.68 and the p-value was almost exactly what

12 came out of the Fisher's p-value.

13   Our reviewer also -- if you remember the

14 meta-analysis of the nine studies that I presented,

15 where the nine were picked because they all had

16 HAMD-17 as primary, regardless of any other

17 differences, that's a rough proxy for similarity or

18 homogeneity.

19   When you look at just the seven negative

20 studies from that meta-analysis, the treatment

21 effect again averaged with -- the inverse weighted

22 mean method was minus 0.67, which is almost exactly

1   the same. So that's two meta-analyses of the

2   negative studies giving you a considerably smaller

3   effect.

4   I will note, somewhat ironically, that

5   that's about the size of the active control versus

6   placebo in one of the studies that you consider to

7   be assay sensitive.

8   DR. TEMPLE: But did the analyses you just

9   described include the four failed studies?

10   DR. LaVANGE: These are the failed studies.

11   DR. TEMPLE: I'm sorry -- include the

12   four studies that we considered negative with assay

13   sensitivity. I don't think you included those.

14   DR. LaVANGE: Actually, it's a mixed bag.

15   It actually included some that did not have assay

16   sensitivity. It included some of the failed

17   trials. These were done at different times with

18   different criteria.

19   DR. TEMPLE: Okay. It's just worth

20   remembering -- I mean which ones you pick determine

21   how it comes out.

22   DR. LaVANGE: So for example, the 053 study

1   is included in those.

2   DR. TEMPLE: Okay. But the three

3   studies that we concluded -- some of us

4   concluded -- had assay sensitivity because of the

5   HAMD-17, those all went the wrong way. So those

6   would make the average effect smaller. So it depends

7   on which ones you include.

8   DR. LaVANGE: And I think we have that one

9   too, probably. I just haven't found it.

10   DR. D'AGOSTINO: Are there any other

11   questions? Why don't we go to the panel member?

12   Dr. Rudorfer?

13   DR. RUDORFER: Thank you. Question for the

14   agency. Just going back to the two atypical

15   depression studies, I guess I'm unclear why those

16   would be included at all in this body of work,

17   given that -- my understanding is that for these

18   folks to have met the criteria for atypical

19   depression, wouldn't other patients who met

20   criteria for MDD be excluded? And if the

21   indication being sought as MDD, this would

22   presumably be a subset of those patients.

1   Would these trials be eligible to be counted

2   even if positive for the indication of MDD?

3   DR. MATHIS: This is Mitch Mathis from

4   Psychiatry Products. We've always thought of

5   atypical features MDD to be as the DSM had defined

6   it along the way, as just another type of MDD. We

7   think there are probably people with atypical

8   features in most MDD trials just as they are

9   recruited from the general population.

10   At the time, there was no objection that I

11   can find where FDA said that it was not okay to

12   study patients with atypical depression. In fact,

13   it looks like the HAMD-25 was the instrument that

14   was chosen for that. It looked like, at that time,

15   I think that this drug might go down an atypical

16   depression pathway and then stopped perhaps after

17   those two trials.

18   DR. RUDORFER: Thanks. Just one follow-up

19   question. Should we be concerned about this

20   unusual path of the three sponsors, where clearly

21   there was on sponsor that, as you say, thought

22   about going down the atypical depression path, and

1 then the studies did not work out, and then they

2 passed the drug on to another sponsor?

3 Is it possible that this has inflated the

4 number of quick-and-dirty, so to speak, trials over

5 the years that we're seeing now?

6 DR. TEMPLE: It's hard to know. This is

7 Dr. Temple. It's hard to know what to make of

8 that. The trials don't look bad. I mean, we

9 thought a lot of them looked reasonably good. But

10 could the multiplicity of sponsors have introduced

11 error likelihood that we don't know about? I guess

12 it's possible.

13 It's not the first drug that has been passed

14 from one sponsor to another, so we don't ban it in

15 any sense. We try to look at the results. And

16 maybe that has something to do with why so many of

17 the other trials failed or were negative.

18 Again, I want to emphasize we're not saying

19 we know this drug doesn't work. We're just

20 expressing doubts about whether the two positive

21 studies are sufficient to provide the substantial

22 evidence we need. And as I said, we're only asking

1   for -- we only think, and we've said this over the

2   years, that one more study is what's necessary,

3   which is not a conclusion that the drug doesn't

4   work. It's that there isn't quite enough evidence

5   yet.

6   DR. D'AGOSTINO: Dr. Gogtay?

7   DR. GOGTAY: Nitin Gogtay, NIMH. I have a

8   question along the same lines. Given that it's

9   been such a long process, 12 trials, three

10   different sponsors, does FDA look systematically in

11   any patterns of site bias, site-specified bias of

12   positive versus negative trials?

13   DR. TEMPLE: Well, it's an interesting

14   question. The sites are all blinded as far as we

15   know, and we don't know anything to the contrary.

16   What would site bias mean in that case? It could

17   have to do with putting the right patients in or

18   putting a different bunch of patients in.

19   We haven't noticed anything -- to my best

20   knowledge. We should have the biostats people

21   comment too. We haven't noticed anything that made

22   any of the sites look funny. Sometimes on

1   inspection, we find a problem at a site, and then,

2   you know, we just toss those data or things like

3   that. We didn't see anything like that here.

4   What has everybody has noticed, the

5   sometimes the sites -- the two sites in a trial

6   behave, came out very differently. Well, we don't

7   know the explanation for that. But that's not

8   unprecedented in large outcome trials either.

9   Sometimes the effects in Australia aren't the same

10   as the effects in the U.S. or Canada. We don't

11   usually have any explanation for those things, and

12   we'd like to, but we don't usually know.

13   So we didn't see anything in here that made

14   one site look contaminated or anything like that.

15   We just looked at all the data, and we didn't have

16   anything that made us want to throw any of them

17   out.

18   DR. D'AGOSTINO: Any other comments on the

19   FDA? Dr. Pickar?

20   DR. PICKAR: As Dr. Temple said earlier, I'm

21   right on the same page, is looking at the data, are

22   we getting it right, is this drug an effective

drug? And that's a tough question. It really is.

I don't envy your position having to deal with

this.

On the other hand, having had three sponsors

to this, to do one good clinical trial, takes

offense, defense, and special teams, it's

remarkable that we got this far. I mean, it's just

a compliment, my God, to see where all these data

are coming.

Where does that leave us? I didn't get the

call, but I wish you hadn't done the post hoc

argument and just said, here it is, guys; this is a

toughie. What do you think? And let people kick

it around.

But you did. And, doggone it, it's a funny

thing to do. Rob Conley mentioned it. He's an

industry guy. I'm not an industry guy, and I'm

sure that gave you twinges. If you ever try to do

a post hoc after your primary endpoint wasn't met,

I don't know that you'd be well received. I just

wish that hadn't happened, but it did.

Everybody's gotten the 17 item. Dr. Yang,

1   could you up the items that were differentiated,

2   the 25 to the 17, please? Because in your

3   briefing, you say those additional items represent

4   atypical features or something like that? Isn't

5   that correct?

6   Can we just take a look at them -- or any

7   other explanation why post hoc, you would just drop

8   that one that you approved. I assume the HAMD-25

9   had your approval to start with at a primary

10   endpoint when they did the study. I'm guessing

11   that.

12   DR. YANG: Peiling Yang, FDA. It's back up

13   slide number 13.

14   DR. PICKAR: Okay. Let's take a look.

15   DR. TEMPLE: Can I just say something while

16   you're looking?

17   DR. PICKAR: Sure.

18   DR. TEMPLE: Nobody thought that using a

19   modified HAMD to emphasize things that were common

20   in the atypical syndrome was a bad thing because

21   they were looking to see if their drug had a

22   special benefit in those people. We didn't have

1   any trouble with that at all and would have thought

2   it was okay had it won.

3   What you're asking is how do you then dare

4   go back and look at the HAMD-17 for the rest of the

5   data? That all started actually when we did our

6   first meta-analysis of all of the trials that

7   didn't win to see if there was a good lean that we

8   could use to attribute benefit to the drug. And

9   the HAMD-17 was the one common denominator. It was

10   their primary endpoint in most of the trials, and

11   it was available for the additional trials.

12   So that's where it started. But the

13   rationale is what everybody has been saying. This

14   is a standard test for antidepressants.

15   DR. PICKAR: Hang on.

16   DR. TEMPLE: Okay.

17   DR. PICKAR: It was effective in 1960. This

18   is Max Hamilton's story, and it is very old. And

19   the reason people have done the MADRS and added to

20   it, because it doesn't, in fact, reflect larger

21   views of it.

22   Now, those symptoms, for example, psychic

retardation, motor retardation, those are of

serious depression. Are you kidding me? Diurnal

variation? It's fundamental. It's fundamental to

biological depression. Now, it may not have been

when Dr. Hamilton put the scale together, but it

just is.

In hypersomnia, I had mentioned it earlier,

there are certain syndromes that are hypersomnia.

I don't call them atypical. It's part of the

picture of depression. In fact, people, either

biology or the clinical aspects of bipolar disorder

are going to be hypersonic when they are treated

with antidepressants. It's a large part of what

people do.

So I didn't mind your -- it's just they're

not atypical in my mind, and they're post hoc. The

first thing I learned about the FDA was it's like

playing pool in that you got to call your shot, or

the ball comes back out.

So when you went to the post hoc -- oh, I

just wish you just left it alone and said, I don't

know whether to believe this. What do you guys

1   think? This felt like it was tampering after the

2   fact in way that just -- I hear your explanation.

3   I don't buy it. I do not buy it.

4   It would appear -- and please correct me.

5   You've got much smarter people than I, here. It

6   would appear that you saw something that favored

7   the idea, which was bothering you, is this drug

8   really effective? And you went to the -- the

9   Hamilton 17 has three items on whether you don't

10   sleep well, not too much sleep, do you get up

11   early, do you get up in the middle --

12   Tom Detre taught me that when I was a

13   medical student at Yale. That was a hundred years

14   ago. Did they have early morning awakening, or

15   middle of the night awakening or late awakening?

16   Is there a difference? No difference. It's an old

17   scale.

18   DR. TEMPLE: So you're saying you think the

19   HAMD-25 is probably a better measure of depression?

20   DR. PICKAR: Yes. And even if it's not,

21   that was what you approved.

22   DR. TEMPLE: Okay. Fair enough.

1   DR. PICKAR: Just laid it all out there, and

2   we'd have to struggle with --

3   DR. TEMPLE: All I'll do is say it again.

4   To look at all the studies together, you had to use

5   the HAMD-17 because that was the one that was

6   common to all of them, including, by the way, the

7   ones that won.

8   DR. PICKAR: I'd like you to say, "Advisory

9   board, forget our post hoc." Forget it. Shabatt

10   shalom.

11   DR. TEMPLE: Well, all right. If you --

12   DR. PICKAR: Okay, it's over. Now, let's

13   look at this as we were forced to look at it --

14   DR. D'AGOSTINO: I think the point has been

15   made.

16   DR. TEMPLE: Just let me say why it's

17   important. There's a difference in evaluating

18   results between whether a trial has failed, which

19   as I said before, it does not count against you in

20   any way, and whether it was negative. So we

21   thought -- maybe you don't agree and that's all

22   right. That's why we're here.

1   DR. D'AGOSTINO: We can pick this up when we

2   come to the particular questions.

3   DR. PICKAR: I'm all for you. I just like

4   to struggle with you on it.

5   DR. D'AGOSTINO: Dr. Goldberg, you have a

6   question?

7   DR. GOLDBERG: Judy Goldberg. Bob, you said

8   that you're asking sort of for another trial. But

9   if you were to design another trial today, how are

10   you going to design it? You're going to do less

11   observation carried forward; we can use the HAMD-17

12   versus 25; you're going to do mixed models and look

13   at the dropout over time in a different way.

14   So I'm still not clear that even if you got

15   one more positive trial here, would that

16   still -- all this other stuff is still there. So

17   now, what does it mean? It tips it a little bit,

18   but it's still -- if we take all of it, it's

19   nine -- what is it, 9 and 4, whatever. And it'll

20   become -- I'm sorry 2 and 9, so it'll become 3 and

21   9? I don't -- or possibly 2 and 10, which then

22   would go to the other side.

1   The best case, it could be 3, and I'm a

2   little unclear how that would be a solution. So

3   help me out here.

4   DR. TEMPLE: Well, I've tried to explain

5   this, but obviously haven't succeeded. There seems

6   like a reasonable chance, using the analyses that

7   Mary did years ago, the two positive trials could

8   be a chance occurrence. Okay? It could be that

9   even if the drug had no activity at all out of 10

10   trials or 12 trials, you could have two positives,

11   and there's a reasonable enough chance that that's

12   true. When you get to three, the chance goes down

13   to next to nothing. That's one thing.

14   How to do the trial, I'll leave to Mitch.

15   We are currently using modeling approaches to deal

16   with the dropouts. My own personal bias is you

17   should do 4-week trials instead of 8-week trials

18   because the dropouts all happen after four weeks.

19   Nobody's bought that to-date.

20   We don't use LOCF anymore because everybody

21   thinks that's not good enough. So the modeling

22   approaches that's been used -- maybe there's a

1   better way, but I'm not the one to offer a view of

2   that. I think we're only saying that they should

3   do another trial using whatever the current

4   standard is. And which measure they want to use

5   could depend on what population they pick, and it

6   would be okay with us if it won.

7   DR. D'AGOSTINO: I'm going to suggest we

8   move. We're going to now move to the charge to the

9   committee. Dr. Jenkins will now provide us with

10   the charge to the committee.

11   **Summary and Charge to the Committee**

12   DR. JENKINS: Before we get to the questions

13   that you have up on the slide there, can we go back

14   to Dr. Yang's slide, number 19, from her primary

15   presentation?

16   I think it's the best slide to give the

17   committee just an overview of the trials that we're

18   discussing and the differences between the

19   prespecified analysis done by the sponsor and the

20   analyses conducted by the review team leading up to

21   the three not-approvable actions. Slide 19?

22   MS. BHATT: Was this on Dr. Yang's

1   presentation?

2   DR. JENKINS: Dr. Yang's primary

3   presentation.

4   Okay. You, in your briefing books, you have

5   the individual versions of these slides in larger

6   separate panels. If you want to look at those,

7   it's slide 17 and 18. But slide 19 is a nice place

8   just to look at them all together.

9   Let's first focus on the left-hand panel.

10   Here, you've got all of the 12 short-term treatment

11   trials displayed, and you see the two positive

12   trials on the far left. And they were

13   statistically significant using HAMD-17 as a

14   prespecified endpoint.

15   You then have the three negative trials,

16   also used HAMD as the prespecified endpoint.

17   You'll note that the sponsor argues that study 08

18   and 02 are supportive because they have a trend

19   favoring gepirone in those studies.

20   I want to jump over to the far right first.

21   These are what have been referred to as the

22   uninterpretable trials. There were essentially

1   trials that were stopped early for business reasons

2   by Bristol-Myers Squibb, the first sponsor of the

3   application.

4   Then you get into the trials that are

5   labeled as being in dispute. You see that two of

6   those trials used the HAMD-25 as the primary

7   endpoint. One trial used the MADRS as its primary

8   endpoint. The fourth trial, 053, used HAMD-17.

9   You can see if you used the prespecified per

10   protocol analysis, what happens in those four

11   trials, none of the pairwise comparisons are

12   statistically significant, drug versus placebo,

13   active control versus placebo, and if you want to

14   go there, drug versus drug.

15   Then if you go over to the next panel, this

16   is the FDA analysis, and the only place it differs

17   is on the four trials that are labeled as being in

18   dispute because the positive trials, the negative

19   trials, and the uninterpretable trials all used

20   HAMD-17 as their primary endpoint to begin with.

21   So there was really no alteration of their analysis

22   when FDA tried to do its overall analysis using a

1   common endpoint.

2   You'll see that in slide 04, what changes is

3   that now, even though the active control does not

4   beat placebo even on this new analysis, there is

5   the finding that there's a nominally significant

6   p-value of gepirone versus active. The same is

7   true in study 006. Study 017, similar findings.

8   These are the things that Dr. Temple has

9   highlighted several times today that he thinks are

10   unusual and very important in determining that

11   these trials did, in fact, have assay sensitivity,

12   taking note of the fact though that the our primary

13   definition of assay sensitivity was active drug

14   versus placebo. And in these three instances, even

15   with the new analysis, they don't beat placebo.

16   The fourth study, 053 -- let me finish, Bob.

17   DR. TEMPLE: It does. Fluoxetine does.

18   DR. JENKINS: Bob.

19   DR. TEMPLE: Okay.

20   DR. JENKINS: Bob, you've acted on this

21   application three times. Now, I'm trying to get

22   the charge to the committee, so let me finish.

1   So 053, what changes there is, as

2   Dr. LaVange just cited a couple of times, there was

3   a change in the treatment-by-center interaction

4   analysis such that the active control now beats

5   placebo; gepirone does not.

6   This is the universal data you have in front

7   of you. The sponsors presented various sensitivity

8   meta-analyses. You've seen lots of different ways

9   to look at this data. But it comes down to the

10 question I presented at the beginning, have they

11 demonstrated substantial evidence of effectiveness

12 of gepirone in treatment of major depressive

13 disorder?

14   We've tried to show you what the historical

15 comparison for other drugs that have been approved,

16 that has positive studies, failed studies, and

17 negative studies. You've seen some different

18 presentations of that from the sponsor versus what

19 was presented by the FDA team. I think some of

20 that relates to what trials were accounted and

21 which trials were excluded based on whether they

22 may or may not felt to have had relevant doses

1   included in the study.

2   That's what you're really helping us to sort

3   through in this discussion today. If we can now go

4   back to the questions. As I said at the beginning,

5   the real focus of this dispute, the real focus of

6   this meeting, is on have they shown substantial

7   evidence of effectiveness?

8   That's why our first two questions are

9   discussion questions where we really would like to

10 hear you probe into these issues about how do you

11 integrate across a database of trials and interpret

12 two positive trials in the face of failed and

13 negative trials. And if you're going to have that

14 sort of cross trial comparison or integration, how

15 do you that? What's the appropriate metric?

16 You've heard several different approaches to do

17 that today.

18   The second discussion question asks you to

19 opine on the issue of assay sensitivity. You've

20 heard a lot of discussion about the use of the

21 active control in depression trials to try to judge

22 should the trial that fails for the test drug be

1   counted against them or just be thrown out as being

2   uninterpretable.

3   We ask you to discuss some of the issues

4   we've just been going over today about how should

5   you do the assay sensitivity analysis? Should we

6   limit ourselves to the primary endpoint in a

7   prospectively defined analysis plan unless, as

8   Dr. LaVange said, there was an obvious mistake in

9   that analysis plan? Or can you look at secondary

10   endpoints? Can you look at post hoc endpoints even

11   if the primary analysis failed to show a

12   difference?

13   Now, it struck me as I was looking at these

14   questions a little while ago that one could argue

15   you should discuss 2 before 1. But I'll leave that

16   up to Dr. D'Agostino. When we write these

17   questions, they look so smart to us when we write

18   them. And then we often get here, and you start

19   wondering if they make as much sense.

20   I have to say I look at them and thought it

21   might make more sense to do 2 on assay sensitivity

22   before you do 1 on how to integrate analyses across

1 trials because some of the methods of analysis to

2 integrate involve the counting methodology of

3 positive and negative trials.

4   Next slide, then you get into the standard

5 voting questions for a pre-approval review. We

6 want you first -- and obviously, number 3 is the

7 most important question, have they demonstrated a

8 substantial evidence of effectiveness for gepirone

9 in the treatment of major depressive disorder?

10   We'll be very interested, not only in your

11 vote, but after you vote, going around and having

12 each member tell us why you voted the way you

13 voted.

14   You've heard there's not much debate about

15 the safety profile, but to be complete, we give you

16 a question to ask if you think they've adequately

17 addressed evaluating the safety profile. And then

18 question number 5 is to integrate the benefits and

19 risks to see if you think it meets the recommended

20 standards for approval.

21   Then finally, at the end, we'll be very

22 interested in hearing any suggestions you may have

1   for additional trials that should be done before

2   approval if you don't think substantial evidence

3   has been demonstrated, after approval; if you think

4   it has but you still would like more information,

5   or some combination of those. This is a chance for

6   you to be academicians and opine on what additional

7   data you would like to see.

8   I'll leave it to you, Dr. D'Agostino,

9   whether you want to start with 2 or 1.

10   DR. D'AGOSTINO: Let me read the boilerplate

11   here, then we'll talk about the questions.

12   We will now proceed with the questions to

13   the committee and the panel discussions. I would

14   like to remind public observers that while this

15   meeting is open for public observation, public

16   attendees may not participate except at the

17   specific request of the panel.

18   We will be using an electronic system for

19   voting when we come to the vote. Once we begin the

20   vote, the buttons will start flashing and will

21   continue to flash even after you entered your vote.

22   Please press the button firmly that corresponds to

1   your vote. If you are unsure of your vote or you

2   wish to change your vote, you may press the

3   corresponding button until the vote is closed.

4   After everyone has completed the vote, the

5   vote will be locked in. The vote will then be

6   displayed on the screen. The DFO will read the

7   vote from the screen into the record. Next, we

8   will go around the room and each individual who

9   voted will state their name and vote into the

10   record. You can also state the reason why you

11   voted as you did if you want.

12   We will continue in the same manner until

13   all questions have been answered or discussed.

14   When we come to specific votes, I will remind you

15   of what I just read.

16   I think that -- well, first of all, I think

17   that I like actually the way the questions or the

18   discussions is laid out. My reason for it is that

19   the first discussion, one, is a broad set of

20   questions and discussion on the broad issues, where

21   the second one gets very focused on this particular

22   situation. So I would like to keep the discussion

1   questions the way you have them.

2   Shall I read them now? Discussion

3   question 1 --

4   DR. JENKINS: Dr. D'Agostino, can I just

5   make one more comment?

6   DR. D'AGOSTINO: Yes, please.

7   DR. JENKINS: On the slide that I had you

8   look at, slide 19, Dr. Temple has pointed out to me

9   that in the FDA analysis for slide 006 using the

10   HAMD-17, paroxetine did beat placebo. So there's

11   two of the four active controls in the FDA analysis

12   that beat placebo, and three situations where they

13   concluded that there was a nominal p-value for

14   active versus gepirone.

15   So that was my error in trying to look at

16   the red ring around the paroxetine circle.

17   **Questions to the Committee and Discussion**

18   DR. D'AGOSTINO: Is that understood by the

19   panel? Thank you.

20   Let's go back to the discussion questions

21   now. We're asked to discuss the following

22   question -- this is question 1A and B -- A. In the

1   situation where two positive adequate and

2   well-controlled trials have been completed, how

3   much and what type of negative evidence from other

4   negative or failed trials would it take to

5   undermine a finding of substantial evidence of

6   effectiveness?

7   Is the question clear? We could go around

8   the table and get some responses or get some

9   discussion on it, or do people want to start

10   responding on their own? If not, why don't we

11   start with Dr. Conley?

12   DR. CONLEY: So where there are two positive

13   and adequate well-controlled trials, we do think we

14   have evidence for efficacy. It would seem to me

15   that if we're looking at negative evidence, we

16   would want to look at trials that were truly the

17   same, so the inclusion criteria was the same, the

18   outcome measure was the same.

19   I honestly don't what the counting metric

20   is. I would agree I don't really like a counting

21   metric either. Maybe something that would give a

22   sensitivity analysis over a number of studies might

1   make you think that you may not have evidence of

2   efficacy. But at the minimum, I would think that

3   you would want to have literally the same type of

4   trials being compared against each other if you're

5   going to weigh a positive and negative.

6   DR. RUDORFER: Yes, I would agree with that.

7   MS. BHATT: Can you please state your name?

8   DR. CONLEY: I'm sorry. Rob Conley. Sorry.

9   DR. RUDORFER: I'm Matthew Rudorfer. I

10   think that one challenge is some sort of quality

11   measure for trials. We spoke about some such as

12   initial depression rating scale.

13   What I'm unclear about is whether sponsors,

14   especially if they believe they have a couple of

15   positive trials under their belt, might just

16   commission a few other quick-and-dirty trials just

17   to kind of see if there's a signal there, or as we

18   saw here, maybe look for a subtype of depression or

19   sometimes even go outside and just do a couple of

20   studies in anxiety disorders or eating disorders

21   just to kind of see and to test the waters.

22   It's not clear to me that those always meet

1   the same kind of rigorous quality standards that

2   we're looking for. And I guess the challenge then

3   becomes to have some proactive approach to ensure

4   that the studies that are going to be looked at

5   closely actually are up to snuff.

6   DR. GOGTAY: Nitin Gogtay. I agree with

7   what has been said. I don't know enough statistics

8   to be able to quantify how much negative or failed

9   trials would it take to undermine a substantial

10   evidence.

11   But what I worry about is one could go on

12   doing trials until they find the two positive

13   trials. The field that is already muddy about

14   antidepressant treatment, does it make sense to add

15   mud to the already muddy water?

16   DR. MANN: John Mann. I think that there

17   has to be some kind of consideration given to

18   trials beyond the two positive trials. I'm not

19   sure what the perfect answer is. I would propose a

20   solution.

21   If a company or sponsor comes forward with

22   two trials that are positive, but then they've done

1   X number of other trials, I would recommend pooling

2   the data from the other trials and treating them as

3   a third trial if the methodology is sufficiently

4   harmonious to be able to do that and to do what

5   I've been plugging away all day about which, is to

6   do a personal level coalesced single database

7   analysis.

8   Therefore, I would treat those data

9   potentially just as a significant as the two

10   positive trials and weigh them in the hopper, then

11   you've got 2 plus 1.

12   Another approach would be to take the data

13   and to divide it up so that the biggest of the two

14   positive trials is matched by a similar number or

15   more subjects. Then if they've got enough

16   subjects, then they can have two of those studies

17   matched up against their two positive trials and

18   see how they come out. I'm in favor of using all

19   of the data in some fashion or other.

20   DR. GOLDBERG: I guess there's no substitute

21   for two positive trials and the more the better. I

22   think it's an extremely complicated issue. If the

first two trials are positive, then this problem

isn't going to be there probably. If the first two

trials are not positive, it just goes on and on,

and that's the real problem.

I think that probably some sort of analyses

should be done where you do some random

reallocation and do some simulations to see how

stable the results are.

I think also that you need to do something

about your making sure that the base trials in your

program are as like as possible in areas like this,

so that you don't get into these post hoc issues.

I think that's probably one of the most important

things; it's these subtle differences in enrollment

and this -- I mean we haven't heard any discussion

today of the site-specific differences within the

trials and across the trials, all of which would be

obviated if the data were put into one big pot and

then overall analysis done on the individual data.

I think that there should be some guideline

developed, though, for a number of trials beyond

which it becomes moot. I think that's really the

1   dilemma here because there was no guideline.

2   I was in the pharmaceutical business and I

3   did develop some of these kinds of drugs. What you

4   saw is, You know, well, if we do six trials and two

5   are positive, it's going to be fine, and then what

6   happens if it's eight? And this is where we are.

7   I think those kinds of guidances need to be

8   developed for framing the problem.

9   DR. FOLLMANN: This is Dean Follmann. I

10   hadn't worked in psychiatric drugs until -- well, I

11   don't. Today was my first introduction to negative

12   trials, failed trials, assay sensitivity, so I'm

13   looking at all with kind of fresh eyes.

14   I'm more comfortable and familiar with the

15   setting where you have to two positive trials out

16   of two, you have replication, and you have a very

17   small p-value to combine them. You have a lot of

18   evidence of efficacy.

19   Here, we're in a different situation, but I

20   sort of harken back to that situation we're

21   familiar with, which is two of two basically. And

22   I try and think, how can I sort of modify that

1   thinking to the situation where we have failed

2   trials and negative studies?

3   One of the things we were introduced -- I

4   was introduced to here was the binomial method of

5   calculating a p-value, which I didn't really like

6   at first because I thought it was kind of

7   inefficient to just say, oh, it's positive or

8   negative. Yet, as time went by, I grew to like it

9   a little bit more. It sort of gets at, in my mind

10 kind, of the replication issue in a way that

11 meta-analysis doesn't.

12   If we have two positive studies, both

13 p-values less 0.025, we wouldn't do a

14 meta-analysis. If we had a positive and a negative

15 study, we would say start over or go again. We

16 wouldn't do a meta-analysis, combine the data, and

17 then see if it was small or not.

18   So I have some -- I've grown to like the

19 binomial method a little bit more than when I first

20 was introduced to it. I'd like to make small

21 comment on all that. Usually, the way we calculate

22 a p-value is to say, what's the probability of what

1   we've seen or something more extreme in favor of

2   the drug rather than what's the probability of what

3   we've seen?

4   What the FDA has done is: What's the

5   probability of two of 12 say? I think the

6   probability should be what's the probability of

7   2 of 12 or better, which is a standard way of

8   calculating a p-value.

9   This also can be viewed as a way of, okay,

10   we're in the arena where we're going to keep doing

11   studies until we get two positives. If that's sort

12   of the way -- if that's a reflection of sort of how

13   the drug development program is proceeding, this is

14   a way to calculate a p-value for that kind of

15   try-try-again kind of sampling. So I've grown to

16   like that a bit more, the binomial method of

17   calculating p-value. I still like meta-analysis.

18   I think it has its place.

19   In the other arenas, I like two p-values

20   less than 0.05 or that sort of kind of evidence.

21   It's quite rare I will be comfortable with weaker

22   levels of evidence. I think you can think of sort

1 of rare diseases as one place where you can only

2 have 20, 30, 100 in the people in the world to do

3 the study. Then maybe you can reduce the level of

4 evidence or if you're looking at it like in an

5 Ebola treatment study, or an HIV prevention study

6 where you can only do one and another one would be

7 unethical, sort of extreme, sort of unusual

8 circumstances where I can feel more comfortable

9 lowering the bar basically.

10   I don't see that here in psychiatry or these

11 drugs. There are a lot of patients that you can

12 enroll. You can certainly do a lot of studies, so

13 why can't you get strong evidence from these unless

14 the drug is not strong?

15   Another thing that's related to this issue

16 of like two positive and how many negative studies,

17 I sort of worry about whether -- if you have 2 of 8

18 or 2 of 10 or something like that, is it going to

19 be useful in the real world? The two might be a

20 selective and artificial representation of what

21 happens out there. I don't know psychiatry enough

22 to know if that's a real issue or not, but it

1   really gives me pause because it's unlike the

2   situations I'm familiar with.

3   I do kind of like the, I'll call it the

4   negative binomial method of assessing evidence. I

5   also like the meta-analysis method of assessing

6   evidence. I'll just have a couple of comments on

7   that.

8   It's nice, but in some ways, you have a

9   terrible burden when you do a meta-analysis because

10   you're potentially picking and choosing studies

11   when you know what the outcomes are. If you have

12   an individual study, this will be equivalent to

13   saying, okay, I want to look at the outcomes of all

14   the patients, know whether they were placebo or

15   treatment, and then decide who I want to have in

16   the final analysis.

17   We would never do anything like that for

18   individual patients in a trial to pick and choose,

19   and I think that same kind of thinking applies for

20   meta-analysis where you want to include all the

21   studies basically to protect against a tendency,

22   whether conscious or unconscious, to pick the right

1   sort of studies or not the right sort of studies

2   and so on. I think the strict remedy for that is

3   to include everything.

4   Finally, there's like a subtle issue, I

5   think, with the meta-analysis. Just as in a

6   regular clinical trial, if you at test statistic

7   over time and you stop early and you estimate the

8   treatment effect based on the evidence at that

9   point, it's likely to be biased basically. You're

10   likely to have stopped at random high.

11   If we're doing these kinds of studies where

12   we keep doing the study, and study again and again,

13   there's a similar kind of bias that can arise for

14   the meta-analysis because you're sort of stopping

15   when you've had a positive study.

16   For the statisticians, if you think about

17   geometric or negative binomial sampling, and you're

18   trying to estimate the probability of success,

19   that's a biased estimate. There's a similar kind

20   of thing going on here I think with the

21   meta-analysis.

22   I don't think it's a very substantial kind

1 of bias, but the try-try-again bias has a small, I

2 think, effect on meta-analysis here. So that's it.

3 DR. COMPAGNI-PORTIS: Natalie

4 Compagni-Portis. I want to refocus us on a

5 different part of this question, which is where it

6 says, what would it take to undermine a finding of

7 substantial evidence of effectiveness?

8 We have 20 years of data, I think, or more

9 and a lot of studies. We don't have any

10 consistency in these studies. I'm not a

11 statistician, but I'm a clinician and the patient

12 representative. And I don't see that we have

13 substantial evidence. In the documents, we keep

14 looking at robust findings, and I don't see that we

15 have robust findings. I want us to all remember to

16 focus on, is this something that would be of

17 significance to patients and would it change

18 clinical treatment?

19 I am, of course, disturbed by the reanalysis

20 and redefining on both parts of the sponsor and the

21 agency, but I think what does remain consistent is

22 that we don't have substantial evidence of

1   effectiveness.

2   MS. HIGGINS: Jennifer Higgins. I really

3   think we need to be looking at all the trials in

4   their totality. I can't disregard the negative and

5   failed trials from this analysis. I would have

6   preferred to have some sort of guidance regarding

7   counting or weighting of those trials if there was

8   any to be had.

9   DR. PICKAR: I agree virtually everything

10   that's been said and would love to see, as John

11   says, the totality of the data and the consistency

12   of the methodology met. This is the A and B. But

13   unfortunately, we're not. We're handed to look at

14   this.

15   It's a drug that has a different mechanism.

16   I've never used it. I learn from seeing how people

17   respond as well as from data, and I'm curious

18   whether its pattern response is different. I'm

19   fascinated that the other drugs beat it on the 17

20   but not on the 25. What's going on?

21   Is this activated? I could ask you five

22   questions. When I was running my own laboratory, I

1   had a lot of things asked to the fellows doing

2   this. There's a lot of unanswered questions here,

3   and unfortunately, I don't have that in front of

4   me.

5   But it smells a little bit like it's a

6   different kind of drug. I know that, as Steve

7   presented pharmacologically. And I suspect its

8   clinical responsivity is different, but that's a

9   guess.

10   But we're using old rules because we're old

11   people, and it's an old science, so the data is not

12   overwhelming. But I got to tell you, coming out

13   with two positive studies with a drug with a new

14   mechanism is not chopped liver either, and it's

15   very hard to do it. It's very hard to do it. And

16   I'm not talking in terms of just work. I'm just

17   talking about a signal from a drug, particularly in

18   current depression studies.

19   So I'm still digesting those positive trials

20   rather than getting sick to my stomach about the

21   negative ones.

22   DR. STEIN: Hi. Murray Stein. I'll just

1   reiterate some of the things I said before. The

2   two positive trials and their positive with the

3   kinds of effect sizes, I think everybody agrees we

4   typically see in recent antidepressant trials, so

5   decent effect sizes.

6   The idea that we would then look at the

7   negative trials to try and convince ourselves those

8   two trials aren't sufficient evidence of the drug's

9   efficacy just sort of flies in the face of reason

10   to me. If we're going to look at the negative

11   trials using a different metric -- and you can't

12   count trials, right? You can't positive and

13   negative because they've all got different effect

14   sizes and they have different numbers of subjects

15   in them. So they're not all equally meaningful in

16   terms of the information they provide.

17   One way to level the playing field and look

18   at all of the information is with meta-analysis.

19   And as I said before, I'm convinced in looking at

20   that, and it does provide some idea of the drug

21   placebo difference to the effect size that this

22   drug is squarely where other antidepressants that

1   are currently marketed are.

2   I think it's been raised before, but we

3   don't want to forget that it's not that there

4   weren't some people who did really well on this

5   drug and other people who did not well on the drug

6   at all. And that's the state of the art of

7   treating depression right now.

8   It'd be wonderful if we knew who give which

9   drug to. We don't know that. For me, having

10   another drug on the market that maybe has a

11   different mechanism -- it does have a different

12   mechanism of action. Whether or not that matters

13   or not, I don't know if it's going to help some

14   people that other drugs currently marketed don't

15   help, but it might. And I don't have any sense

16   that it's going to be any worse than anything out

17   there.

18   So I appreciate looking at the negative

19   trials. I don't think you can exclude them, but

20   you got to look at all the trials together, not

21   just look at the negative trials, to un-convince

22   yourself that the two positive trials are

1   substantial evidence of effectiveness.

2   DR. NARENDRAN: Raj Narendran. I think the

3   two positive trials give some confidence, but on

4   the other end, I would argue when you

5   have -- you're telling us 50 percent of trials in

6   depression fail. So if it's two positive, two

7   negative, I feel very good. Two positive and six

8   negative, I don't feel as good. And then it's 10,

9   it causes a lot more pause. What's going on?

10   At that point, I think you are obligated to

11   go in and look at why these negative studies are

12   negative and why are they failing, because this is

13   going to go into the market, and people are going

14   to get prescribed, and major depression is not

15   without consequence. People commit suicide.

16   What the FDA did to understand this data set

17   I think was very reasonable. The HAMD, for

18   example, the 17 was their outcome measure in 9 out

19   of 12 trials. And I would feel more confident, I

20   think, if the negative data is informative like you

21   did in the vilazodone article.

22   Even when you have negative trials, the

1   change is in the right direction, and there's also

2   a trend level p indicating -- which hasn't reached

3   significance but is in the right direction and you

4   get p-values like that. So I think it's important

5   to kind of understand the negative data as well in

6   the context of two positive trials.

7   To synthesize it, I think the meta-analysis

8   is reasonable, but again, I have a problem with

9   when you say we're going to just pick the people

10   with HAMD over 20 and ignore the people with the

11   lesser score when you included them in a trial. If

12   you're going to do a meta-analysis, then you have

13   to pool everything in irrespective of what you

14   think because that is sort of a post hoc as well.

15   That's my thoughts.

16   DR. IONESCU: Dawn Ionescu, psychiatrist.

17   Just reflecting on all the comments from my

18   colleagues and everybody today, thinking about the

19   two-positive study requirements from 1962 is our

20   current status quo.

21   As psychiatrists, that's gotten us to the

22   point of approving medications that work for, at

1   best, 60 percent of patients, but there's still 30

2   to 40 percent of patients that still need better

3   treatments. And we now have 18 to 19

4   antidepressants approved, as we've heard today.

5   So I think with most things in life, there's

6   not a black and white answer to this first

7   question. There's something in the gray. I'm not

8   a statistician. I don't know exactly how many

9   studies would need to be done to show more negative

10   over positive versus the other way around.

11   But the one thing that keeps coming back to

12   me as a psychiatrist is, first, do no harm. And

13   approving a medication like this that has so many

14   negative trials with only two positive trials,

15   would that potentially cause our psychiatric

16   colleagues, that don't necessarily work in academic

17   centers like many of us do, to prescribe a

18   medication that doesn't work quite as well as other

19   medications that we currently have approved? I'm

20   not sure of the answer to that.

21   With all of the data that's been presented

22   today, I'm looking at the negative trials, and I

1   can't ignore them. I just want to make sure that

2   we're doing the right things for our patients

3   because we do have to go back and explain these

4   medications to our patients, and we have to do

5   what's in their best interest.

6   As I don't know the answer to this question,

7   I keep thinking, first, do no harm.

8   DR. D'AGOSTINO: Dr. De Gruttola, are you

9   still on the line and you have an opinion?

10   DR. DE GRUTTOLA: Yes. Can you hear me?

11   DR. D'AGOSTINO: Yes, we can.

12   DR. DE GRUTTOLA: I certainly agree that

13 it's important to look at the totality of the

14 evidence, not just the positive studies. And when

15 you're talking about an effect size, also look at

16 the totality of the evidence. So I do think that

17 the meta-analyses and the analyses done by FDA have

18 been very useful.

19   One point I want to come back to, there was

20 a lot of missing data in the studies. The FDA

21 representative pointed out that they used a

22 non-completer equals failure, which is described as

1   conservative. That can be conservative, but not

2   necessarily because, for example, if taking in the

3   treated arm dropouts differentially because of side

4   effects and if getting side effects -- intolerance

5   to side effects is related to the amount of

6   depression, then it can be anti-conservative.

7   I think, especially in a setting where

8   there's lack of clarity about the efficacy, a

9   deeper consideration of missing data is relevant.

10   I think Dr. Temple said that in the meta-analysis,

11   he was concerned only about the p-value, not the

12   effect size. And I understand his point. But when

13   you have this level of uncertainty about the drug

14   and have questions about what impact of missing

15   data and what impact of choosing certain

16   populations to be represented in meta-analysis,

17   when those concerns arise, it certainly would be

18   more reassuring to see more robust overall effect

19   sizes as not just in the positive study.

20   So I completely agree with the necessity to

21   take into account the totality of the evidence, not

22   just the positive and negative study. I think

1   there's been a lot of discussion on the approaches

2   out there.

3   DR. D'AGOSTINO: Incorporating my opinion

4   and also where I think is a summary of what has

5   been said is we're talking about having two

6   positive studies. If we have very strong positive

7   studies, good effect size, well-run studies versus

8   having two positive studies that just sort of made

9   the mark and are not completely clear, we would

10   have different ideas in terms of total evidence

11   that -- negative evidence that would sway our

12   opinion from going with just the two studies.

13   I think the type of issues we would have

14   would be that is the negative studies or are there

15   negative studies coming from similar trials,

16   similar designs, similar outcome measures, showing

17   there's consistency in terms of design and

18   implementation -- are they the ones that are

19   producing the negative results?

20   Then the totality of all the evidence, all

21   the studies you have, are we seeing a robustness as

22   you go across the studies. So you've got these two

1   very positive studies, and then you've got a number

2   of studies, unfortunately, that have been similar

3   designs and what you have you but aren't showing

4   the same results. And when you will go beyond

5   those similar types of studies with simple design

6   endpoints, you're seeing this robustness in a

7   negative fashion.

8   I think things like that would -- that type

9   of affair would be very bothersome for us and

10   wouldn't lead us to say that the negative evidence

11   gets to be very convincing.

12   You could also mix in that a meta-analysis.

13   And with the meta-analysis, trying to do it at a

14   binomial method or person level, depending on the

15   quality of the data and depending on the different

16   designs and what have you, you could try both. But

17   the idea would be that would the meta-analysis

18   convince you that the negativity is overwhelming,

19   the two positive studies.

20   If the original study wasn't solid, you have

21   two positive studies but questionable studies, then

22   I think you'd look at the similar trials, you look

1   at the meta-analysis, you look at the consistency,

2   but you would be more likely to be concerned about

3   the negativity if the two so-called positive

4   studies weren't as positive as one would like.

5   I think in the explanation of the

6   negativity, from the sense I get from what we're

7   talking about, we would want to avoid reanalyzing

8   the data with changing the endpoints and doing

9   different analyses and sort of beating the data to

10 death until you find -- until it confesses, as one

11 of my colleagues used to say -- in a negative

12 fashion.

13   But doing it in a reasonable fashion in

14 terms of similar trials, good meta-analysis,

15 consistency in the other trials, robustness, I

16 think it's this type of accumulated evidence that

17 would lead us to say that the negativity is very

18 concerning and possibly could overwhelm the two

19 positive trials.

20   Now, I gave my opinion, but also I was

21 trying to summarize. If I'm off the mark in terms

22 of the summary, please let me know, and I'll throw

1   you off the committee.

2   (Laughter.)

3   DR. D'AGOSTINO: Does that sort of capture

4   what we're talking about?

5   (No audible response.)

6   DR. D'AGOSTINO: We have a break right now,

7   and why don't we take it and come back

8   around -- what time is it now -- let's say about

9   3:17? Thank you very much.

10   (Whereupon, at 3:02 p.m., a recess was

11   taken.)

12   DR. D'AGOSTINO: Let's have the meeting back

13   in order please. We discussed discussion

14   question 1. There were two parts in that

15   discussion question. Our answers and hopefully my

16   summary sort of captured both the A and the B part

17   of that. Is there any problem with that? If not,

18   then why don't we go right to question 2?

19   Question 2 is, please discuss your views on

20   ways to evaluate clinical trials for assay

21   sensitivity. Please consider the following

22   questions in your discussion: Is the primary

1   endpoint for efficacy prospectively defined in the

2   protocol the only meaningful way to evaluate assay

3   sensitivity? And B is, can post hoc analyses of

4   other efficacy endpoints or use of other analysis

5   methods contribute to the determination of assay

6   sensitivity?

7   We can probably do like we did with the

8   first one, if we feel reasonable to blend the

9   responses for both A and B. Why don't we start on

10   the other end?

11   Victor, are you on the phone?

12   DR. DE GRUTTOLA: Hello? Can you hear me?

13   DR. D'AGOSTINO: Did you hear me read the

14   question? We're on discussion question 2.

15   DR. DE GRUTTOLA: Yes, I did. Sorry. I had

16   muted. Yes, I would comment that it depends what

17   the goals are. Evaluating assay sensitivity is a

18   different goal than confirming something that is

19   effective. As I understand it, assay sensitivity

20   is that you want to demonstrate that in the

21   population that you're studying, would it have been

22   possible to see a treatment effect, not that you

1   want to confirm that that treatment effect should

2   be a basis to approve that drug. I think it's

3   reasonable to consider other endpoints, as was

4   done, which I guess would be response to B.

5   DR. D'AGOSTINO: Just to make sure I

6   understood, you think the post hoc type of analysis

7   and using different endpoints to see the robustness

8   is appropriate?

9   DR. DE GRUTTOLA: Yes, I thought it was

10   appropriate for the context that it was used in

11   this discussion.

12   DR. D'AGOSTINO: Thank you.

13   Please state your name before you give the

14   response.

15   DR. IONESCU: Dawn Ionescu. I definitely

16   think that there's value in post hoc analyses but

17   my one concern is with post hoc analyses, we often

18   only see what we want to see. There's a positive

19   publication bias. Everyone in this room knows

20   that, and we don't often see the negative trials.

21   So when we're balancing post hoc analyses,

22   it's really hard to truly balance them because I

1   think that we don't see a lot of data sometimes.

2   But with that being said, I still think they're

3   certainly useful.

4   As far as primary endpoints go, I think this

5   was the big change that happened with

6   ClinicalTrials.gov, is stating our hypotheses

7   before we start out on a trial. We can look back

8   on trials in thousands of ways and find different

9   results. And I do think it's very important to

10   keep the primary endpoint that was defined at the

11   beginning of the protocol in mind.

12   With that being said, again, as I said, I

13   think post hoc analyses are important, but to have

14   a prospective trial with a hypotheses that we came

15   up with at the beginning and then to show that is

16   either statistically significantly different or

17   not, in this case with medications, is important to

18   keep in mind.

19   DR. NARENDRAN: Raj Narendran. I guess the

20   key question -- I mean, another way of framing of

21   these questions is the test drug versus the active

22   comparator, which is the assay sensitivity, should

1   they be evaluated on the same metric?

2   For a test drug, which we don't know

3   anything about, it makes sense to say it has to fit

4   the primary endpoint. On the other hand, if we're

5   talking about fluoxetine or paroxetine, we know

6   these drugs work. It's really the question, for

7   the assay sensitivity, I think it's perfectly fine

8   to use post hoc analyses. And I don't think it

9   should be restricted to the primary endpoint such

10   as the test drug.

11   We're talking about psychiatry where these

12   are rating scale-driven. This isn't like a blood

13   pressure value to say, okay, we're going to stick

14   with systolic blood pressure and mean arterial

15   pressure, and all of a sudden say we're going to

16   change the measure to life expectancy. That's not

17   what we're talking about. Psychiatry is inherently

18   difficult.

19   So I think from a statistical point of view,

20   yes, you have to be puritan. You guys say primary

21   endpoint is the only way to go. But I don't think

22   that is applicable when we know a drug already

1   works. I think it's perfectly fine to use post hoc

2   analyses.

3   DR. STEIN: Hi. Murray Stein. Yes, I think

4   the whole idea of assay sensitivity, it's really

5   just a metaphor. This isn't an assay. We're not

6   measuring anything. We're asking people how they

7   feel and if their symptoms have gotten better. We

8   know very little about things like test/retest and

9   all sorts of things. This is far from anything

10   like any other assay that we would want to

11   consider.

12   So I'm not sure exactly what the question

13   is. I think the question is, in terms of the

14   primary endpoint, I think trials need primary

15   endpoints, and that needs to be the basis on which

16   they're judged primarily as being positive or

17   negative, and I think that has been done here.

18   But I think we've seen today that looking at

19   other endpoints and other approaches to data

20   analysis can be useful to tell us something beyond

21   yes or no on the primary endpoint.

22   DR. PICKAR: I actually agree with Murray's

1 comment on that, and I think the notion of a

2 primary endpoint, calling it before is a very

3 important thing in this kind of drug evaluation

4 process.

5   Post hoc analysis, Dawn commented on the

6 potential of it getting a little squirmy, that's

7 what bothers me. What are your post hocs? We did

8 one today. We talk about it. There could be three

9 or four others, which can give information. That's

10 the way we as clinical investigators used to live,

11 but not the way the FDA used to live. We were

12 instructed that. So I'd be very careful about

13 going to the post hoc thing.

14   Murray is entirely right. It's a metaphor

15 called assay sensitivity in something like that.

16 So my instinct is to stick with your primary

17 endpoint, understanding that that post hoc, you're

18 going to ruin an awful lot that's important. But

19 when you got to keep score, I'd be careful.

20   MS. HIGGINS: Jennifer Higgins. I agree

21 that we really need the primary endpoints and to

22 keep those in sight. Dawn raised an interesting

1   question earlier about the rater reliability, and I

2   give that some thought, too, when I think about the

3   primary endpoint and how much weight we give it,

4   just a question I have.

5   DR. COMPAGNI-PORTIS: Natalie

6   Compagni-Portis. I think what Raj said is really

7   important, though I look at it a little

8   differently. I think given how subjective what

9   we're looking at is, I think words matter. And how

10   we ask and what we ask needs to stay even more

11   consistent in a way because of that because it's

12   not just a number that we're dealing with.

13   I think good science dictates that we don't

14   go back and cherry-pick and change what we said

15   we're looking at. As Dawn said, that we have to

16   stick with the initial question, that that's what

17   we've all learned to do.

18   I think the post hoc analysis can contribute

19   to what we're looking at, but as people have

20   pointed, it really does muddy the water. You can

21   keep looking and looking until you find what you

22   want to find, and that's a challenge with this.

1   DR. FOLLMANN: Dean Follmann. I wanted to

2   have a couple comments on assay sensitivity kind of

3   in general before I go to A and B.

4   I think it was alluded to that it sort of

5   comes from the laboratory where you're doing

6   something in the laboratory of a positive control,

7   a negative control. If the positive control

8   doesn't show a positive signal, you throw out the

9   plate and you basically do a do-over. And I think

10   in the laboratory, that's a fine thing to do. And

11   I think it's predicated on the idea that the

12   positive control is almost always going to be

13   positive unless something catastrophic happens.

14   That's not at all the situation here where

15   maybe 50 percent of the time, or greater than that,

16   the active control doesn't beat placebo. So that

17   should give us pause.

18   Particularly, something that Lisa mentioned,

19   that the operating characteristics of this kind of

20   procedure haven't been evaluated. When I looked at

21   this, I thought, wow, they're really getting

22   multiple chances to give a positive study when if

1   new beat placebo. But if new doesn't beat placebo

2   and active doesn't beat placebo, you get a do-over.

3   So you can keep trying again, not counting these

4   failed studies in some sense, which will inflate

5   your type 1 error rate.

6   So I was curious about how much type 1 error

7   rate inflation can there be. And it's a bit

8   difficult because to do this properly, you'll have

9   to know what benefit does the active have. But if

10   the active has a benefit such that about half the

11   time, it'll result in a positive study, the type 1

12   error rate goes from about 0.025 to close to 0.05.

13   If, in fact, the active is useless, the

14   type 1 error rate is about 50 percent, not 0.05,

15   but 50 percent. So it depends enormously on how

16   well the active is, truly; it's something you don't

17   really know.

18   So I think you need to evaluate in the

19   operating characteristics of this sequential

20   testing procedure, find out what they are, and then

21   sort of calibrate appropriately. I think the

22   calibration is difficult, if not impossible because

1 you have to sort of know what the true effect of

2 the active is, which you don't. So I don't know

3 that there's a good simple path forward on this.

4   A couple of things I thought of was, one,

5 you could legitimately take all of the data in a

6 blinded way and then say, gee, I think the trial

7 had assay sensitivity, yes or no. Maybe you could

8 look at the lumped benefit, the dropout, I don't

9 know, certain things. But that would be a fair way

10 to throw away bad studies and keep good studies

11 without putting your thumb on the scale and

12 inflating the type 1 error rate like a completely

13 blinded thing.

14   Another thing that you could do would be to

15 say, okay, I'm going to have a different rule for

16 assay sensitivity. I'm going to have a large

17 placebo group. Half the placebo group, I'll test

18 against active. If that's positive, then I'll test

19 against new. If that's negative, I'll call it a

20 failed study. And that would be a fair way to

21 throw out studies as well. But if your new drug

22 wins against placebo and active didn't, well, you

1   can't count it.

2   So I think there are some serious

3   ramifications to the operating characteristics of

4   this assay sensitivity kind of maneuver, which I

5   can understand where it came from, but there are

6   consequences to it. Anyway, that's sort of some

7   general comments about assay sensitivity.

8   In regards to A and B, I think, basically,

9   you really want to prespecify what you mean by

10   assay sensitivity. What is the rule for discarding

11   a study, yes or no, and then evaluate its operating

12   characteristics and maybe adjust things so it has

13   proper type 1 error rate.

14   Part B, as many people have mentioned, I'm

15   not really in favor post hoc analyses unless

16   something sort of dramatic or substantial has

17   happened. Sometimes we might discard a primary

18   endpoint in a study if we see there's a huge

19   mortality effect and so on. So that kind of thing

20   is sort of a license. I think if something

21   exceptional is happening, then maybe you could look

22   at post hoc analyses, but as a general rule, I

1   wouldn't favor it.

2   DR. GOLDBERG: I think Dean raised some

3   really good points, and I agree with what he said.

4   I think my basic view is assay sensitivity is a

5   little bit flaky at best. My gut reaction is that

6   you really should stick to what the metric was for

7   the approval of that first drug so that at least

8   you know that it worked like roughly 50 percent of

9   the time with that metric. And then, you can do

10   post hoc if you need to. But I think it's a very

11   fraught with problem area, and I'm not sure what it

12   really means.

13   What will we do if we have the new drug

14   working in the trial but the old drug not, even

15   though the drug was approved on that endpoint? I

16   mean that's going to happen. It has to happen.

17   What do you do? You throw the trial out? I

18   don't think so. I think it's a very problematic

19   thing, and maybe the notion of assay sensitivity

20   should be more in the background or there should be

21   another way to figure out how to incorporate it in

22   a more meaningful way.

1   DR. MANN: I think that the -- first of all,

2   the selection of the active comparator is really

3   important to get off on the right foot, and the

4   primary endpoint should match the strength of the

5   active comparator.

6   I like the idea that you can throw out

7   failed studies. I think that's helpful. At least,

8   it removes some of the confusing data from the

9   scene. But I'm also cognizant of the point that

10   Raj raised, which is this is a different drug with

11   a different mechanism of action. We don't have

12   another antidepressant that acts like this.

13   So you could have a primary endpoint that

14   really doesn't match this drug's action. Part of

15   the problem may be you're not assessing it

16   properly, in which case, you may go to other types

17   of post hoc analyses. And I think that's a good

18   thing to do. That's how we progress

19   scientifically. But when it comes to approvals at

20   the regulatory level, I think all that tells you,

21   as any post hoc discovery tells you, that you now

22   have to do the real study all over again, properly,

1   and so the clock starts again.

2   DR. GOGTAY: I agree that the primary

3   endpoint for efficacy should be -- it's critically

4   important, but at least in this context, when it

5   comes to the approval of a compound that's going to

6   be given to real people with real effects, to me,

7   at least, but not being a statistician, the

8   post hoc analysis, presented both by the sponsor as

9   well as FDA, has actually quite helpful in our

10   trying to understand the concerns in a more

11   informed way, because I'd much rather be safer than

12   sorry.

13   DR. RUDORFER: Matthew Rudorfer. I agree

14   with John's point about the mechanism of action in

15   trying to identify the correct comparator. I think

16   the primary endpoint is most important, but I'm

17   reminded that still, even with all the data we've

18   seen, the efficacy measures still only get us so

19   far so that in the real world, we're concerned also

20   about effectiveness, what difference a drug would

21   make in a person's life.

22   Hamilton didn't have to worry about that

1 back in the tricyclic era. But again, measures of

2 functioning don't appear anywhere in our database.

3 You can report that question 1 on Hamilton is

4 improved because your mood is better. But if

5 you're still lying in bed and unable to get up and

6 go to work or school, that drug still might not be

7 terribly helpful for you, which leads me to

8 conclude that post hoc analyses could have a role.

9 I'm thinking in particular, maybe not today

10 but one day, we'll do a better job of being able to

11 kind of disaggregate the various measures that go

12 in to our rating scales, so that if a drug, for

13 instance, is helpful against anhedonia but maybe

14 that gets diluted in a whole Hamilton rating scale,

15 that still might be worth attending to. Thank you.

16 DR. CONLEY: Rob Conley. I agree with a lot

17 of what I've heard. David, what you said I wished

18 also, that it kind of hadn't gone down the post hoc

19 way because I really do think that's a problem with

20 looking at efficacy, and I think it's something

21 that we've gotten kind of not only used.

22 But there's a good reason to stick to

1   primary endpoints. I think post hoc analyses are

2   extremely useful for exploratory analysis, which is

3   what I usually hear coming the other way from the

4   table as an industry person. And of course we do

5   use them that way, all the time.

6   With that said, I understand that this is

7   very hard, and what you all have tried to do is to

8   try to figure something out when you had a real

9   question in front of you. So I don't want this

10 just to be heard as critical. It's not really easy

11 to figure this out.

12   You've clearly said that you've been worried

13 about efficacy with this for several times, so the

14 question really is what to do. To me, the

15 challenge -- and I know that I often get frustrated

16 when I read meta-analyses; that you see we search

17 the literature and we found 500 articles, and we

18 applied selection criteria, and then finally get

19 down to four or something like that. But there's a

20 reason for that.

21   I think that the critique I have in thinking

22 about this is it felt like there was a real problem

1 between even defining what was a negative study

2 versus a failed study, what's the inclusion

3 criteria for the population, things that in

4 meta-analyses, as a journal editor, I might've just

5 said why don't you kick that particular thing out

6 of the analysis?

7   I think there was a real strive in here to

8 include all the data that you could, which sounds

9 good. I mean, in many ways, I'd love to be able to

10 do individual-level data that pools things

11 together, but as you well know, particularly doing

12 clinical trials for many years, if you don't have

13 similar inclusion criteria and outcome measures,

14 that's just really hard, and that's, I think, what

15 you're faced with.

16   I don't think there's an easy answer to this

17 because you are faced with a data set that's over a

18 decade long as it were, and things have changed

19 over time. But there is a reason we have the

20 standards we have. There is attention to this

21 because, to me, a lot of what I see behind this

22 question is truly almost more of an academic

1   question.

2   Would I like to look at other things in the

3   primary endpoint and look at post hoc analysis? Of

4   course, I would, and I do it all the time to try to

5   figure out what to do next. But when we have the,

6   as it were, regulatory tension of does stuff work

7   or not, I think there's a reason we have to kind of

8   stick to the rules, and they're hard at times.

9   I really do worry with this stuff that it

10   does have a different mechanism of action, that

11   there may be a population, which albeit has not

12   been found yet, that's the right population for it,

13   even though you see it in the general population,

14   you know, when you see a positive effect in two

15   studies.

16   But that's one of the things that we've

17   looked at in psychiatry for a long time. In kind

18   of echoing what you said, Dawn, I mean, I've also

19   been here for a long time and still treat patients,

20   and I still feel frustrated that a lot of them

21   don't get better. So there's a real tension about

22   do you want something new that might work in a

1   different way versus something that doesn't work?

2   And I'm not pretending I know the right answer to

3   that either. But I do think there's a reason we

4   have the level of evidence we have.

5   So in general, not in favor of post hoc

6   analyses, but that's why.

7   DR. D'AGOSTINO: Let me try and give my

8   views and summarize the views of the committee. I

9   think we're saying that the primary endpoint -- if

10 you're going to be concerned about the assay

11 sensitivity, the primary endpoint is primary in the

12 sense that this is what the trial was supposed to

13 be dealing with and once you look at that.

14   That doesn't exclude the possibility of

15 ad hoc analyses, but they're secondary exploratory

16 in presenting, and I think if I'm hearing

17 correctly, that they should be presented as such

18 and not suddenly take precedent over what would've

19 been the primary endpoint.

20   I think we're also concerned with the nature

21 of the comparator, that we're talking about assay

22 sensitivity and we're looking at the active versus

1   placebo and the treatment group under investigation

2   versus placebo. When you select a comparator for

3   this assay sensitivity, it has to have some

4   meaningfulness to the treatment that you're looking

5   at in terms of evaluating, and I think that we're

6   trying to make that point.

7   I'm concerned when I find the active doesn't

8   beat the placebo, but it does beat the treatment,

9   how to interpret that. I'm not so sure we have

10   anybody around the table that's given an answer to

11   looking at that and saying, yes, we've got a

12   positive study or a negative study. It is

13   something that came up and certainly needs to be

14   considered here.

15   I don't hear anybody going to ad hoc

16   endpoint, and this type of analysis, active beats

17   out the treatment but doesn't beat out the placebo,

18   is somehow another way of extracting great

19   information.

20   I think also, the point that was raised by

21   Dean, that if you have these assay sensitivity

22   studies and you're getting multiple failures, you

1   can say the failures you just throw aside as

2   opposed to being negative. But if you keep doing

3   them enough, are you actually affecting the error

4   rate of your investigation on the drug?

5   Any comments or changes to what I just said?

6   DR. JENKINS: Yes. I'd really appreciate if

7   you could ask the committee members to opine

8   further on the issue about active beating test drug

9   in the situation where the active doesn't beat

10 placebo.

11   For the time being, take away the issue of

12 the post hoc analyses in this case. Just think

13 about primary analysis, primary methodology, active

14 doesn't beat placebo, test doesn't beat placebo,

15 but active beats test. How do you interpret that

16 in deciding assay sensitivity?

17   DR. PICKAR: That's a failed study. That's

18 a failed study, and you got to live with it. I

19 mean, I'm sorry. It's hard. It's a hard job we're

20 in, but that's a failed study.

21   DR. MANN: It might not be failed. It might

22 just mean that the test drug is making people

1   worse.   That would be completely consistent with

2   that finding.

3   DR. PICKAR: It could be. But if it's not

4   worse than placebo, if there's no difference from

5   placebo, just from each other, if neither drug

6   differed from placebo.

7   DR. CONLEY: We've actually looked at that

8   internally but with our phase 2 and phase 3

9   psychiatric studies, because we do see that at

10   times. We really have seen, unfortunately, very

11   poor replication of the test drug being positive

12   later on. It sort of doesn't matter.

13   We don't believe that something that's

14   already on the market may not work, but I think it

15   goes back to it just feels like a failed study. It

16   doesn't feel like we've been able to extract much

17   information out of it. We've had the little metric

18   that nothing from nothing is still nothing.

19   If they're all under placebo, I don't know;

20   I really don't actually know what that means from

21   looking at a lot of them. I know you still try to

22   gain information from it, but that's one where if

they're below placebo, we count them as all truly

failed.

DR. FOLLMANN: I mean, it's a different kind

of failure. I think when you're -- the original

concept was that, for whatever reason, the clinical

trial infrastructure couldn't show anything. I

mean, I think that was the original idea for assay

sensitivity when you compare active versus placebo.

Here, you have something different, and it

could be maybe the active is kind of good and the

new is actually kind of worse than placebo. I

don't know if that's plausible or not, but that

would be the simple interpretation of the data.

To me, it's a different kind of failure, I

guess, than when nothing at all happens. As I was

mentioning before, if you're going to play assay

sensitivity, sequential testing, and decision, you

should really try and codify it up front and

understand the performance characteristics of it.

DR. NARENDRAN: Raj. I don't know if you

could -- I kind of disagree with that you could

just call it failed studies in a simplest fashion.

1   Actually, it is more concerning if the active drug

2   beats the test drug. And if it happens repeatedly,

3   I would more worry like what John Mann said, maybe

4   it's making them worse. It would actually even

5   raise more concern. I wouldn't really discount it

6   as a failed trial per se.

7   DR. D'AGOSTINO: Let me --

8   DR. PICKAR: One second. But that's where

9   neither of them are different from placebo. We're

10   not talking about anybody being --

11   DR. NARENDRAN: I know but still -- but an

12   active comparator, which you know it works already,

13   and it was decided based on that because it's on

14   the market, and it beats your test drug, that's a

15   concern because this drug is going to into humans,

16   given to patients, and it can make them worse.

17   I don't know if I can just discount that

18   data and say, okay, let's shelf this, this is a

19   failed trial, and let's go with the positive stuff.

20   I think, actually, it's more concerning. And I do

21   agree with the division's original analysis, that

22   when it happens three different times, it raises

1   serious concerns, I think.

2   DR. D'AGOSTINO: Let me ask the question

3   that was, I think, hinted at in some of the

4   material I read, maybe not. But say you have a

5   study and you have a treatment, placebo active.

6   You look at the treatment versus placebo. If it

7   wins, that's the end of the game. If it doesn't

8   win, you then go to looking at this sensitivity of

9   the active -- let me say -- I may have missed

10   the -- if the treatment versus placebo, the

11   treatment wins, you stop. If the treatment doesn't

12   beat the placebo, then you look to see if the study

13   had assay sensitivity.

14   DR. PICKAR: That's fair.

15   DR. D'AGOSTINO: Are we happy or satisfied

16   with that as a reasonable procedure? Because what

17   is the type 1 error that's floating around there,

18   where you're only looking at the active versus

19   placebo in the presence of a negative treatment

20   effect?

21   DR. FOLLMANN: I don't mean to play devil's

22   advocate. You could quit doing these active

1   studies and just say compare new to placebo. If

2   you win, great; if you don't, it's a negative

3   study. You know, that's what all the other areas

4   do. You could, you know, maybe come out with

5   better drugs or better trials or something instead

6   of trying to have this other do-over kind of

7   mechanism, which is a different way to think about

8   it.

9   I don't mean to be provocative, really, but

10   it's a different way to describe this landscape.

11   So maybe we could entertain not doing those kinds

12   of studies anymore, and you don't have that

13   problem.

14   DR. PICKAR: If you're not asking the

15   question of whether the new drug is superior or

16   different from existing drug, it's not your

17   question. You're just trying to separate from

18   placebo. And your old drug uses purely to see if

19   you're carrying out the trial in a reasonable way.

20   It's not about comparing it with the treatment

21   drug. I don't think that's the intent, is it?

22   That's what you were saying, Dean, that the

1   use of the established drug was to establish or to

2   speak to the soundness of the testing.

3   DR. FOLLMANN: Right. I think it's

4   problematic for reasons I was talking about earlier

5   in a single -- you know, one thing to consider is

6   just not having the active control, just to new

7   versus placebo. If you win, you've demonstrated

8   it's better than placebo, which is I guess what the

9   FDA cares about, really.

10   You're not so much interested in amongst

11   approved drugs or good drugs, what is the better

12   ones and the worst ones. That's not your remit,

13   really, I don't think. If you want to do these

14   assay sensitivity studies, they're problematic, I

15   think. They lead to a host of problems.

16   DR. PICKAR: The use of the active drug --

17   DR. FOLLMANN: Yes.

18   DR. PICKAR: -- it's an interesting one.

19   It's assay sensitivity. But if your test drug is

20   negative, but you can't show that a proven drug is

21   positive either, then maybe you're doing something

22   wrong in this study.

1   DR. FOLLMANN: Yes, I understand.

2   DR. PICKAR: It takes that out of play.

3   DR. FOLLMANN: I understand that, and I

4 think it came from a reasonable kind of thing, that

5 sometimes the plates are bad in the laboratory.

6 Sometimes in study are bad here for no good reason,

7 really. So why can't we come up with a mechanism

8 to allow for that?

9   But the problem is, if you are doing the way

10 you describe where you get two shots and do-overs,

11 basically, it really inflates the type 1 error

12 rate. I think if we do try and do corrections for

13 it, they might be so onerous, you conclude it's not

14 really worth doing.

15   DR. NARENDRAN: Can they do like placebo

16 run-in studies to just compare with placebo, I mea,

17 just to -- I don't know. Is that standard

18 to -- could just have like everybody go on placebo

19 for two weeks, and then just drop the placebo

20 responders, and then put the active drug on a

21 blinded fashion, randomize them after that? It

22 seems like there must be better designs.

1   DR. FOLLMANN: Right. That would be a

2   design where you give everyone placebo for a period

3   of time, and then you randomize to stay on placebo

4   versus getting the new drug.

5   DR. D'AGOSTINO: Bob, did you have a

6   question or a comment?

7   DR. TEMPLE: Just a comment. This is a

8   distinctly a novel problem. Remember, the reason

9   to include the active control was to avoid having a

10   study that couldn't tell anything be considered a

11   negative study, and that's what it usually does.

12   The only other case that I know of actually

13   is a trial NIMH did on St. John's wort, where the

14   active drug didn't beat placebo, but it did beat

15   St. John's wort, so you can draw conclusions you

16   want from that.

17   This sort of thing is very unusual, and I

18   wouldn't throw out the idea of having an active

19   control, although I grant you all the difficulties

20   of interpretation.

21   DR. D'AGOSTINO: I raised the question

22   because you can run into this sort of a novelty

1   type of thing that's not what you were using the

2   active control for. You were trying to get the

3   sensitivity of the study, so then you can make a

4   meaningful comparison with the treatment and the

5   placebo.

6   DR. TEMPLE: Right. Usually it's to

7   distinguish a failed study from a negative study,

8   and that's good for people. It's a perfectly good

9   question, whether if the active drug beats your new

10   drug, whether that's the same as having assay

11   sensitivity. That's what you're discussing. I'm

12   not trying to intervene with that. This problem is

13   not common.

14   DR. D'AGOSTINO: We're going to now move on

15   to the voting questions. Hopefully, we've had

16   enough discussion where we have our minds all set

17   for the vote.

18   I do want to remind the members of the

19   committee that we will have the question, the vote.

20   And you will give a yes, no, or abstain,

21   which -- the speaker in front of you. When all

22   people have made their vote, we will be told that,

1   and then we will go around the table with the

2   reasons for the vote.

3   Question 3 is a vote. Has the sponsor

4   provided substantial evidence of effectiveness for

5   gepirone extended-release in the treatment of major

6   depressive disorder? Yes, no, or abstain?

7   Is the vote machine ready to take the vote?

8   MS. BHATT: Yes.

9   (Vote taken.)

10   MS. BHATT: We are entering

11   Dr. De Gruttola's vote, so that's why it took a few

12   minutes. Now, I'm going to give the results.

13   The voting results, yes is 4, no is 9,

14   abstain is zero. No voting is zero. These are the

15   results of who has voted, what their answer is.

16   DR. D'AGOSTINO: We'll go around the table

17   now asking people why they voted as such. Please

18   give your name and your vote, and then explanation

19   or justification. We'll start with you.

20   DR. RUDORFER: Yes. Matthew Rudorfer. I

21   voted yes because of the two positive trials, and I

22   was concerned that there was so much noise in what

1   else we reviewed.

2   DR. GOGTAY: Nitin Gogtay. I voted no

3   because I'm not convinced about the data overall

4   that I've seen. Especially, I am concerned about

5   the negative trials.

6   DR. MANN: John Mann. I voted no for the

7   same kinds of reasons.

8   DR. GOLDBERG: Judy Goldberg. I almost

9   voted yes, but as I've reflected on all of these

10   sensitivity analyses and all of the ins and outs,

11   there are too many opportunities for the thing to

12   move. It's old data; it's old methods of analysis.

13   There are analytic problems, and I think it's not

14   convincing enough to say that it's irrefutable.

15   DR. FOLLMANN: My name is Dean Follmann. I

16   voted no. I wanted sort of level of evidence that

17   was substantial. Whether I do the negative

18   binomial counting method of calculating of p-value

19   or the meta-analysis, where you use all the studies

20   and all the data and not excluding patients, the

21   p-values, the level of evidence just didn't seem

22   persuasive to me at all.

1   I also like independent replication, so the

2   withdrawal study was appealing to me. But that was

3   not showing a benefit, so that factored into my

4   thinking as well.

5   Also, something I'd mentioned earlier, with

6   so many failed studies and kind of weak evidence,

7   I'm not sure it would work in the more realistic

8   world, when you're trying to describe it to people

9   and outside of the clinical trial environment that

10 these studies were done. So I have skepticism that

11 it would work in the real world.

12   DR. COMPAGNI-PORTIS: Natalie

13 Compagni-Portis. I voted no. I didn't think the

14 evidence was substantial or robust, or was going to

15 be a significant advance for patients. And I was

16 definitely impacted by the great number of negative

17 results.

18   MS. HIGGINS: Jennifer Higgins. I voted no

19 because I believe the negative evidence really

20 undermines the effectiveness.

21   DR. PICKAR: Dave Pickar. I voted yes. I

22 was impressed by those two studies, and I was

1   particularly thrilled to send this back to

2   Dr. Jenkins and to see some yes votes so you

3   considered -- weigh these things. It's a tough

4   question, but I voted yes as my sum. So it's going

5   to be to you, sir.

6   DR. D'AGOSTINO: D'Agostino. I also voted

7   no -- excuse me, yes. And my reason was that I

8   think that there's a tremendous number of studies,

9   unusual number of studies, that are classified as

10   either negative or failed. But I think that the

11   two positive studies are impressive.

12   I think the type of analyses that we saw the

13   sponsor produce with the meta-analysis, the

14   sensitivity analysis, that has a lot of validity to

15   it. And I think it's at a point where one doesn't

16   necessarily have to run to a new study but look at

17   the existing data in the framework that we've done

18   today, and start off with the yes from us so that

19   it moves to a higher level with clarity.

20   I think as far as a maintenance study, I was

21   assuming that they didn't have to produce that for

22   the approval, that that could be something that

1 could come later on.

2   DR. STEIN: Murray Stein. I voted yes

3 because of the two positive, what used to be called

4 pivotal trials. And then in whatever sensitivity

5 analyses I saw, if you want to call them that, I

6 was most persuaded by the meta-analyses, both that

7 the sponsor did and I think that the FDA also

8 produced. I also saw clear and compelling evidence

9 that the drug worked better than placebo.

10   Should I say anything about safety, or is

11 that a separate question?

12   DR. D'AGOSTINO: That's the next question.

13   DR. STEIN: Next question, okay.

14   DR. NARENDRAN: Raj Narendran. I voted no

15 because I thought the number of failed and negative

16 trials kind of undermined the two positive trials

17 and it didn't really meet the threshold for

18 substantial evidence.

19   DR. IONESCU: Dawn Ionescu, and I voted no.

20 First of all, I can appreciate the amount of time,

21 money, and effort that the sponsor went in to doing

22 these trials.

1   As a depression researcher, I want nothing

2   more than new medications to treat my patients

3   with. However, we currently have medications that

4   work.

5   As Dr. Mathis pointed out in his

6   introductory slides, slide 8, there's many

7   medications that we have to treat depression for

8   patients with first episode depression. As I

9   mentioned earlier, we can treat about 50 to 60

10   percent of those patients with one to two

11   medication trials.

12   We also know that the longer that a patient

13   takes to get better from their depression, the less

14   likely they are to respond to further medication

15   trials.

16   A few points I just want to make about this

17   new medication. There was a lot of excitement

18   mentioned over its new mechanism. I wish I could

19   be a bit more positive, but I see it at the end of

20   the day as another serotonergic modulator. We're

21   not necessarily looking at changing effecting

22   circuits. We're not necessarily looking at

1   changing genetics. This is affecting the serotonin

2   system as many of our current medications do.

3   There was a lot of excitement also -- many

4   people had mentioned the potential for treating

5   patients' depression, which also may in turn treat

6   their suicidal thinking. However, there was no

7   evidence presented for that. And as we learned,

8   these trials were done prior to us really looking

9   at suicide.

10   Third, there was a lot of excitement over

11   the potential for this medication to decrease

12   patients' sexual dysfunction that comes along with

13   many of our other serotonergic agents. But I ask

14   the question, at the end of the day, if this

15   medication isn't going to treat their depression,

16   what's the point of giving them a medication that

17   will also not cause as much sexual dysfunction?

18   Finally, I understand that the two positive

19   trials meet the 1962 standard for approving

20   medications through the FDA. However, in light of

21   all the negative trials and data that we saw, I

22   think the bigger question not necessarily for this

1   group today but for our field of psychiatry in

2   general is, do these standards need to evolve in a

3   similar way that our medications need to evolve for

4   our patients?

5   DR. D'AGOSTINO: Vic, did you have a

6   comment?

7   DR. DE GRUTTOLA: Yes. I would just say

8   that I voted no because of the amount of negative

9   evidence.

10   DR. D'AGOSTINO: Thank you. We'll move on

11   to the next vote. Has the sponsor adequately

12   characterized the safety profile of gepirone ER in

13   the treatment of MDD?

14   (Vote taken.)

15   MS. BHATT: The voting results are yes, 11;

16   no, 2; abstain, zero; no voting, zero.

17   DR. D'AGOSTINO: Victor, do you want to

18   start with your reason for your vote?

19   DR. DE GRUTTOLA: Yes. I thought there has

20   been, obviously, extensive study of the drug and

21   didn't note that there were significant safety

22   concerns that would preclude its use.

1   DR. D'AGOSTINO: Thank you.

2   DR. IONESCU: Dawn Ionescu. The primary

3   reason I voted no was because there's no data on

4   the suicide increase or decrease from this

5   medication. I remember reading in the documents

6   that we received prior to this meeting that if

7   approved, this medication would have a similar

8   warning that other serotonergic agents currently

9   have.

10   However, in light of learning today that

11   these studies were done prior to doing what we

12   consider the standard of suicide rating now, I

13   would be concerned prior to prescribing this

14   medication without that data.

15   DR. NARENDRAN: Raj Narendran. I voted yes.

16   I didn't see any concerns that was presented by the

17   FDA in the briefing.

18   DR. STEIN: I'm Murray Stein. I didn't see

19   any concerns about safety and was actually

20   impressed by the safety profile, including the

21   relative absence of sexual side effects, which I

22   think could be a substantial benefit to patients.

1   There are few antidepressants on the market with a

2   similar side effect profile.

3   DR. D'AGOSTINO: Ralph D'Agostino. I voted

4   yes. I think the previous comments capture very

5   well the safety issues. They don't seem to be at

6   all the issue before us actually.

7   DR. PICKAR: I agree with what Murray and

8   Ralph just said.

9   MS. HIGGINS: Jennifer Higgins. I voted

10   yes. I didn't feel like the data demonstrated any

11   safety concerns.

12   DR. COMPAGNI-PORTIS: Natalie

13   Compagni-Portis. I voted no. I agree with what

14   Dawn is saying about the concerns about

15   suicidality. I also think we have short-term data

16   on a long-term drug that people would be taking for

17   many years, so there could be other safety issues.

18   In terms of the sexual side effects or lack

19   thereof, even though I know that's been swirling

20   around, I think we don't really have the data about

21   that either, yet. And I think even the sponsor

22   said that they weren't making that claim, yet. But

1 I might be wrong about that.

2   DR. FOLLMANN: I'm Dean Follmann. I voted

3 yes. I really have nothing to add beyond the other

4 comments that have voted yes.

5   DR. GOLDBERG: Judy Goldberg. I voted yes.

6 I have nothing to add.

7   DR. MANN: John Mann. I voted yes. I

8 wasn't as concerned about the suicidal effects

9 because I think that the effects of antidepressants

10 in adults reduce suicide risk in proportion to the

11 degree of improvement in depression. So I thought

12 that the risk in this agent is probably related to

13 how good it is for depression, and that was the

14 subject of the previous vote.

15   DR. GOGTAY: Nitin Gogtay. I voted yes, no

16 safety concerns.

17   DR. RUDORFER: Matthew Rudorfer. I also

18 voted yes. Also, I had no safety concerns.

19   DR. D'AGOSTINO: Given the previous two

20 votes, does it make sense to vote on number 5?

21   DR. JENKINS: Yes. I would still like to

22 have you vote on number 5.

1   DR. D'AGOSTINO: Okay. Do the available

2   data support a favorable risk-benefit profile for

3   gepirone ER to support approval? Yes or no?

4   (Vote taken.)

5   MS. BHATT: One more person needs to vote.

6   (Pause.)

7   The voting results, yes is 4; no is 9;

8   abstain is zero; no voting is zero.

9   DR. D'AGOSTINO: So we got a surprise vote.

10   Do we want to run around the table just to see if

11   it's consistent, and please state your name again

12   so we can get it into the record.

13   DR. RUDORFER: Matthew Rudorfer. I voted

14   yes. I think this drug could offer another option

15   for some patients. I think the impressive safety

16   rating is important because I think the clinical

17   marketplace has a good way of sorting things out

18   after the fact. If, in fact, the drug proves to be

19   helpful for some patients and not others, I think

20   that will become clear.

21   Also, if I may color outside the lines a

22   little bit, I'd love to see an augmentation study

1   similar to what we've seen with buspirone with this

2   drug. Thank you.

3   DR. GOGTAY: Nitin Gogtay. I voted no for

4   the same reasons I explained earlier. I'm not

5   convinced about the benefits despite the novelty of

6   the mechanism.

7   DR. MANN: John Mann. I have nothing to add

8   to my previous comments.

9   DR. GOLDBERG: Judy Goldberg. Nothing to

10 add.

11   DR. FOLLMANN: Dean Follmann. I voted no.

12 Nothing to add.

13   DR. COMPAGNI-PORTIS: Natalie

14 Compagni-Portis. I voted no. I think we most

15 definitely need new and varied options for

16 patients, but I think patients need and deserve

17 treatments that are truly effective and safe and

18 that are no worse than available options or no

19 better than placebo. So I didn't find the evidence

20 compelling on the benefit-risk ratio.

21   MS. HIGGINS: Jennifer Higgins. I voted no.

22 I'd like to see further study.

1   DR. PICKAR: I voted yes. I think its

2 efficacy, as I mentioned before, was substantial in

3 that regard. It's really very modest in terms of

4 side effects, which are a big problem in the

5 treatment, so that was my yes. And it would be

6 very handy to have an antidepressant with these few

7 side effects.

8   DR. D'AGOSTINO: Ralph D'Agostino. I voted

9 yes to be consistent with my previous two votes. I

10 thought there was efficacy. There is a need for

11 maintenance study, but there was efficacy on the

12 short-term studies, and there is clear safety.

13   DR. STEIN: Murray Stein. Nothing to add to

14 my previous comments.

15   DR. NARENDRAN: Raj Narendran. I have

16 nothing to add to my previous comments. I voted

17 no.

18   DR. IONESCU: Dawn Ionescu. I voted no, and

19 I have nothing to add to my previous comments.

20   DR. D'AGOSTINO: Victor, could you state

21 your name and give your reason?

22   DR. DE GRUTTOLA: Yes. Victor De Gruttola.

1   No, for reasons stated.

2   DR. D'AGOSTINO: Now, we have a discussion

3   question. What, if any, additional studies are

4   needed pre- or post-approval to address outstanding

5   issues, for example, additional effectiveness

6   studies, an additional randomized withdrawal

7   maintenance study, and so forth and so on?

8   I think the question here is in terms

9   of -- the way I read it, those who voted no for

10 efficacy were going to be suggesting some efficacy

11 studies. I'd like to go around the room and see

12 what people say about that. I'm wondering if these

13 studies, we're going to be talking about the

14 inclusion of a positive control for assay

15 sensitivity and would they just be short-term

16 studies.

17   It was quite about the maintenance trial

18 that we definitely need that. Again, it's sort of

19 a jump off in terms of asking individuals their

20 opinions in terms of what other studies might be

21 needed.

22   Matthew, do you want to start off?

1   DR. RUDORFER: Well, I personally don't

2   think further study is needed before marketing

3   approval, so I might be the wrong person to ask.

4   I'm all for postmarketing surveillance and

5   long-term maintenance studies. And I think, as has

6   been pointed out, which is a very good point in

7   general, that certainly long-term effectiveness and

8   long-term safety can never be demonstrated in

9   eight-weeks trials.

10   So I think that as is true for all

11   antidepressants on the market, that I think

12   longer-term studies would certainly be helpful.

13   DR. GOGTAY: Nitin Gogtay. When I first

14   started reading this, I actually was excited to see

15   a novelty of the mechanism. Although I wasn't

16   convinced about the evidence that was presented, I

17   do think that this mechanism is worth exploring

18   further with a proper study done.

19   It should be done to substantiate the two

20   positive studies that have shown some evidence for

21   this. I'm actually surprised that the last study

22   that was done was in 2005. And by now, in the last

1   10 years, why another trial was not done already.

2   So I do support this. I actually like the idea of

3   an active control.

4   DR. MANN: I'm also intrigued by this drug

5   or this kind of medication, 5-HT 1A agonist. I

6   think it has a lot of potential. I don't really

7   understand exactly why this medication doesn't seem

8   more effective.

9   I would do an occupancy study with a PET

10   scanner. I'd take blood levels, and I'd map the

11   blood levels on to the occupancy in the brain. You

12   can do that pretty easily. Then I'd do another

13   proper clinical trial using 2016 standards and draw

14   blood to check the blood levels, and see if that

15   influences the outcome. Thanks.

16   DR. GOLDBERG: I would propose that they do

17   use a placebo control but no active comparator.

18   From the evidence that we saw, it looks like the

19   drug works more -- is more effective at the higher

20   doses. And I think they need to get there and make

21   sure that the patients stay on it in those doses.

22   So they need to track that, and then using

1   2016 standards for the analysis and for the

2   inclusion. The inclusion should match up, to some

3   extent, to what was done before so that we can see

4   how it fits; same thing for the endpoints. But I

5   agree that there should be some correlates to show

6   what's going on within other domains than just

7   those clinical questionnaires.

8   DR. FOLLMANN: Yes, I think there should be

9   additional studies done, probably one or two like a

10   maintenance trial possibly and an additional

11   effectiveness trial. You have a lot of data

12   actually from all these studies. There's been

13   comments about, gee, this drug should work in some

14   patients, so you could try and identify those

15   patients using the data you have.

16   The sponsor showed some slides where they

17   looked at sort of baseline score and showed that

18   there's a bigger benefit for people who had worse

19   scores at baseline. That could be partly

20   artificial because there's a statistical thing

21   going on there where you're correlating a baseline

22   variable with a change, which induces correlation.

1   Nonetheless, there still might be a subgroup

2   that would have -- you know, for whom the drug

3   worked, just beyond those who have the worst scores

4   at baseline. You talk about it works in someone,

5   and the burden is really to identify the group it

6   would work in.

7   Also, we talked a little bit about a

8   different kind of trial where you'd give everyone

9   placebo for a while and then randomize maybe to

10   continue placebo or getting a new drug. A

11   variation on that would be to see who responds, not

12   so much to placebo, and then in those, maybe you

13   would get a bigger drug effect. This identified

14   group that doesn't have such a great placebo

15   effect, maybe they would have a bigger drug effect.

16   Now, there would be issues, I suppose, with

17   marketing. There's population of people who didn't

18   respond so well to placebo but maybe worked well

19   with the drug. But nonetheless, it's a different

20   design to consider, and I guess just part of

21   discussion, so I bring that up, too.

22   DR. COMPAGNI-PORTIS: I just want to echo

1   what other people said, that I'd like to see

2   additional studies that would help us stratify what

3   patients might be helped if some are -- so more

4   data on the patients and studies with consistent

5   endpoints and consistent study protocols.

6   MS. HIGGINS: I really think we do need to

7   look at some subgroup analysis, and I'm

8   particularly passionate about elders. I know that

9   depression is the most common disorder among

10   elders, and I'd like to see some age analysis done.

11   I'd also like to see larger Ns and longer trials,

12   longer duration. I would like to see greater

13   control of dropouts.

14   DR. PICKAR: There are a lot of studies to

15   be done and to learn about. John, you were talking

16   about mechanisms and all that, but it's useless. I

17   mean let's get real. You're in a fantasy land.

18   The drug needs to be approved. What drug is going

19   to get it approved? There's not endless money.

20   I know Dr. Jenkins offered to do this study

21   and pay for it from his budget, but I think he was

22   just mumbling there for a second. I'm teasing, but

1   I'm not.

2   So the task here for this -- if there's any

3   future in gepirone -- is how do you get it on the

4   market? And if it doesn't pass muster, it doesn't

5   pass muster, so be it.

6   I think any thinking about future trials

7   should be focused on that issue, and what the pros

8   want to see there, and how can it be done in an

9   efficient way. Then there's a bunch of academic

10   things to be done for sure. And if we send it to

11   an intramural program, it'll stay there forever and

12   never get anywhere.

13   This is about getting drugs into people and

14   getting them out and approved. What needs to be

15   done? So this is consultation with those folks,

16   and I'd love to hear exactly what they think about

17   it. That's who's going to keep score.

18   DR. STEIN: Murray Stein. Yes, I think I

19   already said that I think that there have been

20   enough trials done already. I'm convinced that

21   there were two positive trials.

22   If there were more trials done, I think the

1   question would be, what would the purpose be?

2   If the purpose was to further increase the

3   strength of the evidence, then I think the placebo

4   lead-in trials have been tried in depression. And

5   they aren't helpful in large part because people

6   don't just get better in the first couple of weeks

7   on placebo; they get better over time on placebo,

8   so that doesn't really help you a lot.

9   But from what I've heard, probably doing a

10   study that would have two different doses of the

11   drug, something that would be considered inadequate

12   and then adequate, so maybe that's 20 and 60, in a

13   group of patients that had high enough severity

14   that maybe there could be an effect shown would be

15   the way to go.

16   DR. NARENDRAN: Raj Narendran. I pretty

17   much echoed Dr. John Mann's comments. I have

18   nothing more to add.

19   DR. IONESCU: Dawn Ionescu. Just thinking

20   about utility, I don't know if more trials are

21   necessarily needed for this medication. However,

22   if the sponsor decides to go with another trial, I

1   would recommend using something like an external

2   rater system as we talked about earlier, which

3   would definitely add to the previous trials that

4   have been done.

5   As all of us know that have done clinical

6   trials, one of the biggest problems is the

7   heterogeneity of patients in trials. Also,

8   oftentimes, sites are valued based on randomizing

9   patients, not necessarily the best patients but

10   patients. I'm not accusing the sponsor of doing

11   that or anything like that from the previous

12   studies. However, moving forward, I think having

13   an external rater system may help decrease that

14   heterogeneity of patients.

15   Also, to echo what Dr. Mann mentioned, I'll

16   just say the "B" word, biomarkers, looking to find

17   out which of those patients that may have

18   responded, what they look like either from a

19   clinical profile or, as mentioned by my colleagues,

20   a more biological profile.

21   Finally, of course, just to echo the common

22   theme of patients, advocate groups, as well as

1   doctors wanting more treatments for suicidal

2   thinking and suicidal ideation, maybe doing a trial

3   that involves that if the sponsor pursues more

4   trials in this area.

5   DR. D'AGOSTINO: Victor, do you have a

6   comment?

7   DR. DE GRUTTOLA: I echo many of the

8   previous comments. First, start by trying to

9   identify patients who appear to have benefited from

10   treatment and also patients who, to the extent

11   predictors can be found, who will comply with

12   treatments and remain in study, it would be good to

13   increase enrollment of such people.

14   I would also agree that placebo-controlled

15   trials make the most sense. I don't think the

16   assay sensitivity issue appears to be paramount,

17   that maybe to have another solid trial that

18   demonstrates benefit over placebo.

19   DR. D'AGOSTINO: In trying to summarize, I

20   think what we're seeing here is potentially two

21   levels here. If the drug is not ready for approval

22   on the basis of what it has now, then an efficacy

1   trial would have to be put together. At least one

2   efficacy trial would have to be put together that's

3   broad enough and bring enough representation of

4   patients and test for effectiveness and continued

5   safety.

6   The committee was raising a number of other

7   potential designs and what have you, and issues

8   such as the heterogeneity and looking for the

9   biomarkers that could identify the types of

10   individuals who might respond to the drug. And

11   that would be, I think, a different level, possibly

12   a different level type of study.

13   It seemed like the idea of, again, the

14   subgroups, the biomarkers, would be something that

15   would be of importance to look at as you move

16   along. The maintenance study obviously has to come

17   up. If I heard correctly, there wasn't

18   overwhelming support for the notion of assay

19   sensitivity being built in to these studies.

20   Any comments, questions? Should I adjourn

21   the meeting?

22   DR. JENKINS: If I could just make a couple

1   of comments?

2   DR. D'AGOSTINO: Please do.

3   DR. JENKINS: This is John Jenkins. I'd

4   like to thank you, Dr. D'Agostino, for your

5   excellent job serving as chair today. Special

6   thanks to Dr. De Gruttola for serving remotely.

7   You clearly are sick. I don't think you're just

8   phoning it in. I think you're clearly sick, so we

9   really appreciate your --

10   DR. D'AGOSTINO: He has a machine that makes

11   that cough.

12   (Laughter.)

13   DR. JENKINS: Okay.

14   I really want to thank the committee. I

15   think it's been a really robust and helpful

16   discussion. I think the discussion emphasizes why

17   this has been a challenging issue.

18   We heard varying opinions from committee

19   members as we went around the table, and you've

20   given me a lot to think about. So thank you again

21   for your willingness to serve, and I wish you all a

22   safe trip home.

Adjournment

DR. D'AGOSTINO: I do want to thank you for asking me to chair this meeting. I haven't chaired an advisory committee meeting, it must be, 15, 20 years, since the Non-Prescription Drug Advisory Committee that I had the honor chair.

I want to thank the committee. I thought the questions were very pointed. I thought the discussions from the presentations of the FDA and the sponsor were very good. And I thought the committee was really very keyed into the major issues. And hopefully, we've given you useful information.

With that said, I will now adjourn the meeting. Panel members, please remember to drop off your name badge at the registration table on your way out so they can be recycled. I have about 27 of these name tags home, so I'm going to drop mine off for sure this time. Thank you again.

(Whereupon, at 4:21 p.m., the meeting was adjourned.)