FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

JOINT MEETING OF THE PSYCHOPHARMACOLOGIC DRUGS ADVISORY COMMITTEE (PDAC) AND THE DRUG SAFETY AND RISK MANAGEMENT ADVISORY COMMITTEE (DSaRM)

Wednesday, September 14, 2016

8:00 a.m. to 4:56 p.m.

FDA White Oak Campus

10903 New Hampshire Avenue

Building 31 Conference Center

The Great Room (Rm. 1503)

Silver Spring, Maryland

Meeting Roster

DESIGNATED FEDERAL OFFICER (Non-Voting)

Kalyani Bhatt, BS, MS

Division of Advisory Committee and Consultant Management

Office of Executive Programs, CDER, FDA

PSYCHOPHARMACOLOGIC DRUGS ADVISORY COMMITTEE MEMBERS (Voting)

Jess G. Fiedorowicz, MD, PhD

Associate Professor

Departments of Psychiatry, Epidemiology and Internal Medicine

University of Iowa Carver College of Medicine

Iowa City, Iowa

Rajesh Narendran, MD

Attending Psychiatrist

Re:solve Crisis Network

Western Psychiatric Institute and Clinics

Associate Professor in Radiology and Psychiatry

Psychiatric Molecular Imaging Program

University of Pittsburgh

Pittsburgh, Pennsylvania

David Pickar, MD

Adjunct Professor of Psychiatry

Johns Hopkins Medical School

Uniformed Services University of Health Sciences

Gabriel Sciences, LLC

Chevy Chase, Maryland

PSYCHOPHARMACOLOGIC DRUGS ADVISORY COMMITTEE MEMBER (Non-Voting)

Robert R. Conley, MD

(Industry Representative)

Global Development Leader

Pain and Core Therapeutics

Team and Distinguished Scholar

Eli Lilly and Company

Indianapolis, Indiana

DRUG SAFETY AND RISK MANAGEMENT ADVISORY COMMITTEE MEMBERS (Voting)

Kelly Besco, PharmD, FISMP, CPPS

(via telephone)

Health-System Medication Safety Coordinator

OhioHealth Pharmacy Services

Dublin, Ohio

Tobias Gerhard, PhD, RPh
Associate Professor
Rutgers University
Department of Pharmacy Practice and Administration
Ernest Mario School of Pharmacy
New Brunswick, New Jersey

Almut Winterstein, RPh, PhD, FISPE
Professor and Crisafi Chair
Pharmaceutical Outcomes & Policy
College of Pharmacy
University of Florida
Gainesville, Florida

TEMPORARY MEMBERS (Voting)

Dan Budnitz, MD, MPH (CAPT, USPHS)
Director, Medication Safety Program
Division of Healthcare Quality Promotion
National Center for Emerging and Zoonotic
Infectious Diseases
Centers for Disease Control and Prevention
Atlanta, Georgia

Scott S. Emerson, MD, PhD
Professor of Biostatistics
University Of Washington
Seattle, Washington

Terry Gillespie
(Patient Representative)
Westmont, Illinois

Sean Hennessy, PharmD, PhD
Professor of Epidemiology in Biostatistics and Epidemiology
Department: Biostatistics and Epidemiology
University of Pennsylvania
Philadelphia, Pennsylvania

Sonia Hernandez-Diaz, MD, DrPH
Professor of Epidemiology
Department of Epidemiology
Harvard School of Public Health
Boston, Massachusetts

Jennifer Higgins, PhD

(Acting Consumer Representative)

Director of Strategic Planning and Business Development

Center for Human Development

Springfield, Massachusetts

Stephen Marder, MD

Professor and Director

Section on Psychosis

Semel Institute for Neuroscience

University of California, Los Angeles

Los Angeles, California

Glen D. Morgan, PhD

Psychologist

Tobacco Control Research Branch

Behavioral Research Program

Division of Cancer Control and Population Sciences

National Cancer Institute

National Institutes of Health

Rockville, Maryland

Elaine H. Morrato, DrPH, MPH, CPH

Associate Dean for Public Health Practice

Associate Professor Department of Health Systems, Management and Policy

Colorado School of Public Health

University of Colorado Anschutz Medical Campus

Aurora, Colorado

Ruth M. Parker, MD

(Acting Chairperson)

Professor of Medicine, Pediatrics, and Public Health

Emory University School of Medicine

Atlanta, Georgia

Kenneth A. Perkins, PhD

Professor of Psychiatry

University of Pittsburgh

Pittsburgh, Pennsylvania

Rajiv N. Rimal, PhD

Professor and Chair

Department of Prevention and Community Health

Milken Institute School of Public Health

George Washington University

Washington, District of Columbia

Christianne Roumie, MD, MPH

Associate Professor Internal Medicine and Pediatrics
Institute for Medicine and Public Health
Vanderbilt University
Staff Physician
Veterans Affairs Tennessee Valley Healthcare System
Nashville, Tennessee

FDA PARTICIPANTS (Non-Voting)

Mary Thanh Hai, MD

Deputy Director
ODE-II, OND, CDER, FDA

Sharon Hertz, MD

Director
Division of Anesthesia, Analgesia and Addiction Products (DAAAP)
Office of Drug Evaluation II (ODE-II)
Office of New Drugs (OND), CDER, FDA

Celia Winchell, MD

Medical Team Leader

DAAAP, ODE-II, OND, CDER, FDA

Judith A. Racoosin, MD, MPH

Deputy Director for Safety

DAAAP, ODE-II, OND, CDER, FDA

Eugenio Andraca-Carrera, PhD

Reviewer, Division of Biometrics VII

Office of Translational Sciences

CDER, FDA

CAPT David Moeny, RPh, MPH, USPHS

Acting Director

Division of Epidemiology II

Office of Surveillance and Epidemiology

CDER, FDA

C O N T E N T S

AGENDA ITEM	PAGE
Call to Order and Introduction of Committee	
Ruth Parker, MD	15
Conflict of Interest Statement	
Kalyani Bhatt, BS, MS	20
FDA Introductory Remarks/Regulatory History	
Judith Racoosin, MD MPH	25
Applicant Presentations – Pfizer	
Introduction	
James Rusnak, MD, PhD	41
Evidence from Observational Studies	
Judith Prochaska, PhD, MPH	51
EAGLES Study Design, Investigator's Perspective on Study Conduct and on Treating Patients for Smoking Cessation	
Robert Anthenelli, MD	61
EAGLES Study Execution	
James Rusnak, MD, PhD	72
EAGLES Study Results I	
Cristina Russ, MD, PhD	81

C O N T E N T S (continued)

AGENDA ITEM	PAGE
Clinical Perspective on EAGLES Study Results	
A. Eden Evins, MD, MPH	100
Conclusions and Labeling Proposal	
James Rusnak, MD, PhD	116
Clarifying Questions to Applicant	121
FDA Presentations	
Clinical Review of the PMR Safety Outcome Trial	
Celia Winchell, MD	139
Statistical Review of the PMR Safety Outcome Trial	
Eugenio Andraca-Carrera, PhD	164
Sensitivity Analysis	
Celia Winchell, MD	181
Review of Observational Studies of Neuropsychiatric Events Associated with Smoking Cessation Products	
Chih-Ying Pratt, PhD	182
Clarifying Questions to FDA	194

C O N T E N T S (continued)

AGENDA ITEM	PAGE
Open Public Hearing	250
Charge to the Committee	
Judith Racoosin, MD MPH	304
Questions to the Committee and Discussion	307
Adjournment	424

P R O C E E D I N G S

(7:59 a.m.)

Call to Order

Introduction of Committee

DR. PARKER: Good morning. I'd like to remind everyone, if you would, to please silence your cell phones, smartphones, other devices if you haven't already done so. I'd also like to identify the FDA press contact, Michael Felberbaum. If you're here, would you please stand? He's the one not standing.

Okay. My name is Ruth Parker, and I'm the acting chairperson for today's meeting. I'll now call the Joint Meeting of the Psychopharmacologic Drugs Advisory Committee and the Drug Safety and Risk Management Advisory Committee to order. I'd like us to start by going around the table, and we'll include our one participant who is joining us by telephone today, and introduce ourselves.

Let's start with the FDA at the left side of the table there, and go around the room. And if you'll please state your name.

DR. THANH HAI: Good morning. I'm Dr. Mary Thanh Hai. I'm the deputy director in the Office of Drug Evaluation II.

DR. HERTZ: Sharon Hertz, director for the Division of Anesthesia, Analgesia, and Addiction Products.

DR. WINCHELL: Celia Winchell. I'm the medical team leader for addiction products in Dr. Hertz's division.

DR. ANDRACA-CARRERA: Eugenio Andraca-Carrera. I'm a statistical reviewer in the Office of Biostatistics.

CAPT MOENY: David Moeny, acting director for the Division of Epidemiology II.

DR. HENNESEY: Good morning. My name is Sean Hennesey, and I have a sensitive microphone. I do drug safety research at the University of Pennsylvania.

DR. RIMAL: Good morning. My name is Rajiv Rimal. I'm the professor and chair of the department at George Washington University.

DR. ROUMIE: Christine Roumie. I'm intern

medicine and general pediatrics. I also do drug safety research at the VA Medical Center in Nashville and at Vanderbilt University.

DR. FIEDOROWICZ: I'm Jess Fiedorowicz. I'm a physician scientist on the faculty at the University of Iowa and work with the Iowa City VA Health System.

DR. PICKAR: David Pickar, adjunct professor of psychiatry, Johns Hopkins; and former chief of experimental therapeutics branch, intramural research program, NIMH.

DR. BESCO: Good morning, everyone. This is Kelly Besco joining via phone today. I'm a health [indiscernible] pharmacist and medication safety officer for the OhioHealth Hospital, Columbus, Ohio.

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

MS. BHATT: Good morning. I'm Kalyani Bhatt. I'm with the Division of Advisory Committee Consultants Management.

DR. PARKER: Ruth Parker, professor of

medicine, pediatrics and public health at Emory University.

DR. GERHARD: Tobias Gerhard, pharmacoepidemiologist at Rutgers University.

DR. WINTERSTEIN: Good morning. I'm Almut Winterstein, professor and chair of pharmaceutical outcomes and policy at the University of Florida.

CAPT BUDNITZ: Dan Budnitz. I'm a medical officer with the medication safety program in the Division of Healthcare Quality Promotion at Centers for Disease Control and Prevention.

MS. GILLESPE: Good morning. I'm Terry Gillespe. I'm a consumer reviewer.

MS. HIGGINS: Jennifer Higgins. I'm the acting consumer representative.

DR. HERNANDEZ-DIAZ: Sonia Hernandez-Diaz, professor of epidemiology, Harvard School of Public Health in Boston.

DR. PERKINS: Professor Ken Perkins at University of Pittsburgh, and I do smoking cessation research.

DR. MORRATO: Good morning. Elaine Morrato.

I am an epidemiologist in the Department of Health Systems, Management and Policy, and associate dean for public health practice at the Colorado School of Public Health.

DR. MORGAN: Glen Morgan. I'm at Tobacco Control Research Branch, National Cancer Institute. Good morning.

DR. MARDER: Steve Marder. I'm a professor of psychiatry at the Semel Institute at UCLA.

DR. EMERSON: Scott Emerson, professor of biostatistics, University of Washington in Seattle.

DR. CONLEY: Good morning. I'm Rob Conley, the global development leader and distinguished scholar in neuroscience at Eli Lilly, and an adjunct professor in psychiatry at the University of Maryland.

DR. PARKER: Thank you.

For topics such as those being discussed at today's meeting, there are a variety of opinions, usually, some of which are quite strongly held. Our goal is that today's meeting will be a fair and open discussion of these issues, and those

individuals can express their views without interruption. Thus, as a gentle reminder, individuals will be allowed to speak into the record only if recognized by the chairperson, and we look forward to a productive meeting.

In the spirit of the Federal Advisory Committee Act and the Government in the Sunshine Act, we ask that the advisory committee members take care that their conversations about the topic at hand take place in the open forum of the meeting. We are aware that members of the media are anxious to speak with the FDA about these proceedings. However, FDA will refrain from discussing the details of this meeting with the media until its conclusion. Also, the committee is reminded to please refrain from discussing the meeting topics during breaks and at lunch. Thank you very much.

Now, I'll pass to Kalyani Bhatt, and ask her that she read the Conflict of Interest Statement.

Conflict of Interest Statement

MS. BHATT: Good morning. The Food and Drug

Administration is convening today's joint meeting of the Psychopharmacologic Drugs Advisory Committee and Drug Safety and Risk Management Advisory Committee under the authority of the Federal Advisory Committee Act, FACA, of 1972. With the exception of the industry representative, all members and temporary voting members of the committees are special government employees or regular federal employees from other agencies and are subject to federal conflict of interest laws and regulations.

The following information on the status of these committees' compliance with federal ethics and conflict of interest laws, covered by but not limited to those found at 18 USC Section 208, is being provided to participants in today's meeting and to the public. FDA has determined that members and temporary voting members of these committees are in compliance with federal ethics and conflict of interest laws.

Under 18 USC Section 208, Congress has authorized FDA to grant waivers to special

government employees and regular federal employees who have potential financial conflicts when it is determined that the agency's need for a special government employee's services outweighs his or her potential financial conflict of interest or when, in the interest of a regular federal employee, is not so substantial to be deemed likely to affect the integrity of the services which the government may expect from the employee.

Related to the discussion of today's meeting, members and temporary voting members of these committees have been screened for potential financial conflicts of interest of their own, as well as those imputed to them, including those of their spouses or minor children and, for purposes of 18 USC Section 208, their employers. These interests may include investments, consulting, expert witness testimony, contracts, grants, CRADAs, teaching, speaking, writing, patents and royalties, and primary employment.

Today's agenda involves discussion on a completed postmarket requirement randomized,

placebo-controlled trial of the neuropsychiatric effects of Chantix, varenicline; Zyban, bupropion; and nicotine replacement therapy, along with relevant published observational studies to determine whether the findings support changes to the product labeling. This is a particular matters meeting during which specific matters related to Chantix and Zyban will be discussed.

Based on the agenda for today's meeting and all financial interests reported by the committee members and temporary voting members, no conflict of interest waivers have been issued in connection with this meeting. For the record, we'd like to disclose that Ms. Kim Witczak is the consumer representative of the Psychopharmacologic Drugs Advisory Committee and has been recused from participating in this meeting.

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

With respect to the FDA's invited industry

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

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

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

DR. PARKER: Okay. We'll now proceed with the FDA introductory remarks presented by Dr. Racoosin, division director.

FDA Introductory Remarks/Regulatory History

Judith Racoosin

DR. RACOOSIN: Good morning. I'm Judy Racoosin, the deputy director for safety in the Division of Anesthesia, Analgesia, and Addiction Products. Today, I'll start by describing the regulatory history of neuropsychiatric adverse events with the smoking cessation drugs; describe the utilization of these products; briefly review the criteria for key sections of product labeling; and orient you to today's presentations and discussion topics.

We're going to be talking about three different smoking cessation products. The first nicotine replacement therapies were available by prescription only starting around the mid 1980s. Many other formulations were approved in the 1990s with the over-the-counter switch occurring in the mid to late 1990s.

Two nicotine replacement products are still only available by prescription, Nicotrol Inhaler and Nicotrol Nasal Spray. Zyban, which is a trade

name for bupropion, was approved for smoking cessation in May of 1997. The drug moiety bupropion had been previously approved in 1985 with the grade name Wellbutrin for major depressive disorder. Chantix, which is the trade name in the U.S. for varenicline, was approved in May of 2006.

In May 2007, the European Medicines Agency shared a concern with FDA about suicidality with varenicline about a year after FDA had approved the product. I'm going to describe a few sample cases to give you an idea of what was reported.

In this first case, a 36-year-old woman taking varenicline reported having experienced a complete personality change, including a violent temper going into unnecessary rage. She stated her brain felt it had been completely scrambled since about treatment day 14. The consumer believed her experience was not due to smoking cessation because she had given up smoking before and had never felt this way.

In another case, a 61-year-old man taking varenicline reported experiencing suicidal thoughts

approximately 1 week after starting treatment. Treatment was discontinued for 1 week during which those experiences resolved. He then resumed treatment. When the dose was increased to 1 milligram twice a day, he became depressed, and his wife told him his behavior is very aggressive. The patient discontinued varenicline due to these experiences. The suicidal thoughts, depression, and feeling like a zombie resolved, and the aggression persisted. It was not known if he had quick smoking at the time of these events.

Many of the cases that were submitted to FDA's Adverse Event Reporting System, or FAERS, feature the hallmarks of drug related events. For example, the onset of events was frequently shortly after the patient started taking the drug or when the patient titrated up to the full dose.

There were also examples of de-challenge in which the symptom went away when the drug was discontinued, and re-challenge in which the patient whose symptoms had resolved restarted the medication and had the symptom recur like in the

second case that I described.

Initially, there was the thought that these events were related to quitting smoking. However, although some of the symptoms, such as irritability and depressed mood, are symptoms that are associated with nicotine withdrawal, in many cases, the patient hadn't stopped smoking, so nicotine withdrawal didn't seem like a likely explanation. There are also a number of cases in which patients specifically articulated that he or she had quit smoking before and had not had theses experiences.

Finally, Chantix, a partial agonist at the nicotine receptor, can possibly cause nicotine withdrawal by displacing nicotine, a full agonist, at the receptor. We know from the situation with opioid dependence that displacement of an agonist by an antagonist or partial agonist can cause the onset of intense symptoms of withdrawal.

Once FDA had become aware of EMA's concerns about suicidality with varenicline, we evaluated adverse event reports that had been submitted to FAERS for varenicline with bupropion and nicotine

replacement therapy as comparators, as well as reviewing reports that Pfizer had submitted. As our evaluation of the cases progressed and our level of concern regarding the safety signal increased, the placement of labeling language about the association became more prominent, moving from adverse reactions to warnings and precautions.

Through the review process, we became aware that similar cases had been reported with bupropion, and ultimately a box warning was added to both products' labeling in July of 2009. I will describe a couple of sample cases with bupropion.

In this case, about 2 weeks after starting bupropion for smoking cessation, a 28-year-old woman experienced feeling emotional and having regular crying fits. The patient reported having threatened to kill herself, and stated that she didn't care if she lived. She had no previous history of depression documented.

In another case, after about 1 month of bupropion treatment for smoking cessation, a 50-year-old man with a history of military service

and no prior PTSD experienced severe panic attacks, flushing, flashbacks, sleep loss, and as the physician reporting the case said, full-blown PTSD symptoms causing loss of work, and functioning, and self-confidence. The reporting physician noted the patient had no life-triggering events or stressors. The patient's symptoms persisted following discontinuation of bupropion, and he required medical treatment.

In addition to the labeling changes I described, FDA required that a risk evaluation and mitigation strategy, or REMS, be put into place to ensure the benefits of the drug outweighed the risks. The REMS consisted of a medication guide and a timetable for assessments to ensure that patients were adequately informed about the serious risk of neuropsychiatric adverse events. FDA also issued a postmarketing requirement for a clinical trial to assess the serious risk of neuropsychiatric adverse events with the smoking cessation drugs.

We recognize that spontaneous reports

generated the safety signal. However, we needed a clinical trial to systematically evaluate the risk of neuropsychiatric adverse events in a defined population of smoking cessation patients.

In June 2009, following FDA's completion of the evaluation of neuropsychiatric adverse events that had been reported with varenicline and bupropion, and internal discussion about the requirements of the trial design, FDA issued guidance about the PMR safety outcome trial design to Pfizer and GlaxoSmithKline.

First, it needed to be a large randomized, double-blind active and placebo-controlled trial. The treatment arms should include varenicline, bupropion, nicotine replacement therapy, and placebo. The trial should compare the risk of clinically significant neuropsychiatric adverse events, including but not limited to suicidality, and the trial should determine whether individuals with a prior history of psychiatric disorders were at a greater risk for such adverse events compared to individuals without prior history of psychiatric

disorders.

Finally, the trial needed to be sufficiently powered to adequately assess clinically significant neuropsychiatric adverse events within each treatment and each of the two subgroups, those with psychiatric history and those without. Dr. Celia Winchell will discuss further the protocol development in her talk later this morning.

FDA recognized that it would take several years for the sponsors to conduct the PMR trial, and so we sought other approaches to evaluate the issues, as did others in academia. FDA collaborated with our federal partners at the Veterans Administration and the Department of Defense to evaluate the risk of neuropsychiatric adverse events with varenicline using nicotine replacement therapy as a comparator.

In October of 2011, FDA summarized the results of these studies in a drug safety communication. Neither study found a difference in risk of neuropsychiatric hospitalizations between Chantix and nicotine replacement therapy. However,

both studies had a number of study design limitations, including only assessing neuropsychiatric events that resulted in hospitalization and not having a large enough sample size to detect rare adverse events. Later this morning, Dr. Natasha Pratt will discuss observational studies that examined the association of neuropsychiatric adverse events with smoking cessation drugs.

In April 2014, Pfizer submitted a labeling supplement seeking to remove the boxed warning from Chantix labeling. They asserted that more reliable data on neuropsychiatric safety of Chantix had become available, and these data did not support an association between treatment with Chantix and serious neuropsychiatric adverse events.

FDA sought the input of the Psychopharmacologic Drugs Advisory Committee and the Drug Safety and Risk Management Advisory Committee in considering this data. Some of you around the table today participated in that meeting.

Shortly before the October 2014 Chantix advisory committee meeting, a group of five consumer organizations submitted a citizen petition asking that FDA strengthen the Chantix boxed warning about neuropsychiatric adverse events. The consumer organizations included Consumer Reports; Institute for Safe Medication Practices; National Center for Health Research; National Physicians Alliance; and Public Citizen.

At the advisory committee, a majority of the committee agreed that more data were needed and recommended to retain the current boxed warning and reassess once the ongoing postmarketing safety outcome trial designed to capture serious neuropsychiatric adverse events was completed. Similarly, FDA decided to wait to respond to the citizen petition until we were able to review the results of the safety outcome trial.

Now, I'll move on to describe the current extent of utilization of smoking cessation products.

This graph shows the nationally estimated

number of bottles or packages of prescription and over-the-counter, or OTC, smoking cessation products sold from manufacturers to all channels of distribution in the U.S.

From 2011 through 2015, sales distribution data from manufacturers of prescription smoking cessation products, which is the line with the green triangles, remained relatively stable, and sales of OTC smoking cessation products, the line with the red squares, appeared to increase by about 24 percent from approximately 3.8 million packages or bottles in 2011 to 4.7 million in 2015.

However, the data source used to provide the over-the-counter sales data estimates a capture of approximately 50 percent of the entire OTC product market. Therefore, the OTC sales data shown are likely an underestimation of total OTC sales. Therefore, the market share and trends with OTC products should be interpreted with caution.

In terms of patient utilization data of prescription products, this graph shows the nationally estimated number of unique patients who

received a dispensed prescription for Chantix, Zyban, Nicotrol Inhaler, and Nicotrol Nasal Spray, through U.S. outpatient retails pharmacies from 2006 to 2015.

For Chantix, patients increased from 573,000 patients in 2006 to a peak of 3.9 million patients in 2007, before declining to 1.2 million patients in 2012 and remaining relatively steady thereafter. The decline in use beginning in 2007 coincides with the period that FDA started its evaluation of neuropsychiatric adverse events.

Utilization of the other products examined was low during this period, however, the graph underestimates the number of patients taking bupropion for smoking cessation. Although other bupropion products, such as Wellbutrin and generic equivalents, are not approved for smoking cessation, data that's not shown on this graph are suggested that bupropion products other than Zyban are also widely used for smoking cessation in the U.S.

Because we're going to ask you to consider

some product labeling issues today, I'd like to review the criteria for our warnings and precautions statement and the criteria for our boxed warning, FDA's strongest labeling warning.

Generally, a warnings and precautions statement is added to describe a serious or clinically significant adverse reaction that occurred with the drug or risks that are expected to occur. A warnings and precautions section should include a succinct discussion of the description of a topic and should include the following information if known: risk factors for the adverse reaction; the outcomes of the adverse reaction; and numerical estimate of risks or the adverse reaction rate; and steps that could be taken to prevent, monitor, and manage an adverse reaction.

A boxed warning is ordinarily used in the following situations. It may describe an adverse reaction that is so serious in proportion to potential benefit that it is essential it be considered in assessing the risks and benefits of a

drug; or there is a serious adverse reaction that can be prevented or reduced in severity or frequency by appropriate use of the drug; or a drug is approved with restrictions to assure safe use because the drug can only safely be used if distribution or use is restricted.

A boxed warning may also be used in other situations; for example, to highlight a warning that is especially important to a prescriber or for a drug that possesses risk-benefit considerations that are unique among drugs in a drug class.

Now moving on to what we'd like to accomplish today. Following the submission of the final report of the safety outcome trial, the sponsors, Pfizer and GlaxoSmithKline, submitted supplements with specific proposals.

The Pfizer supplement proposed that the boxed warning for neuropsychiatric adverse events be removed from Chantix labeling. Their labeling proposal retains the warning in Section 5.1 about neuropsychiatric adverse events with some changes to reflect the PMR trial safety outcome trial

results. GlaxoSmithKline supplement proposes that Zyban be released from the REMS requirement, but they will still maintain the Medication Guide.

Today, Pfizer will make the industry presentation. GlaxoSmithKline, though a recipient of the PMR and a co-sponsor of the trial, declined to participate in this advisory committee meeting. FDA will present our evaluation of the PMR safety outcome trial, including a presentation of the clinical review by Dr. Winchell, and a presentation of the statistical review by Dr. Andraca-Carrera.

FDA will also present our review of the published observational studies relating to smoking cessation products and neuropsychiatric adverse events. Dr. Pratt from the Division of Epidemiology will be making that presentation.

Following the industry and FDA presentations this morning and the open public hearing early this afternoon, we'll be asking you to opine on what you've heard today. Specifically, we'll ask you to consider the trial design and conduct and how they impact the trial results. We'll ask you to discuss

psychiatric history as a risk modifier for neuropsychiatric adverse events with smoking cessation drugs. And finally, we'll ask you to discuss the impact of the trial results and sensitivity analyses on smoking cessation product labeling.

Again, thank you for being here today to help FDA consider this important issue.

DR. PARKER: Thank you, Dr. Racoosin.

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

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

including equity interests and those based upon the outcome of the meeting.

Likewise, FDA encourages you at the beginning of your presentation to advise the committee if you do not have any such financial relationships. If you choose not to address this issue of financial relationships at the beginning of your presentation, it will not preclude you from speaking.

We will now proceed with Pfizer's presentations.

Applicant Presentation - James Rusnak

DR. RUSNAK: Good morning, Dr. Parker, panel members, members of the FDA, and the public. I'm Jim Rusnak, the chief development officer for cardiovascular and metabolic diseases at Pfizer, and we are pleased to be here today at this joint advisory committee meeting to share new and important data on varenicline.

EAGLES stands for evaluating adverse events in a global smoking cessation study. The EAGLES study was conducted to satisfy FDA postmarketing

requirements, issued for Pfizer and GlaxoSmithKline, related to varenicline and bupropion, respectively. The study was conducted by Pfizer in collaboration with GSK, and this presentation is provided on behalf of Pfizer, and reflects the views and opinions of Pfizer.

Smoking is the leading preventable cause of death and disease in the United States. Smoking causes nearly half a million deaths in the United States each year. The health benefits of smoking are immediate and substantial. Some people are able to quit on their own. Some are able to quit with behavioral counseling. The odds of quitting, however, are significantly improved with pharmacological smoking cessation therapy.

The accumulated body of evidence supporting this statement also supports guidelines, including those from the U.S. Public Health Service that indicates clinicians should encourage all individuals making a quit attempt to use both counseling and medication.

In the more than 50 years since the surgeon

general issued the first warnings on the hazards of smoking, there are only three FDA-approved pharmacological smoking aids available: varenicline, bupropion, and various forms of nicotine replacement therapy. Meta-analyses have shown that varenicline is substantially more efficacious than either bupropion or nicotine replacement therapy, an observation that is confirmed in the data that we will discuss today.

The good news is that the rates of tobacco use are declining amongst many segments of the population. There is however a disturbing outlier, smoking prevalence for people with mental illness. It is a crisis within a crisis for this patient population, and it is exactly these patients that need smoking cessational treatments the most.

As medicines advance through their development and life cycle, we continually learn more about their benefits and potential risks. These data emerge from randomized controlled trials, abbreviated as RCT here, postmarketing reports, and observational studies. As these data

emerge, we evolve our medical practices based upon the totality of evidence.

Each data source has its strengths and limitations. The collective body of these data from complimentary sources, none of which stands alone, allows for an overall assessment of benefit-risk. Large randomized blinded controlled studies are considered to be the highest level of evidence that can be obtained.

Looking at the original phase 3 clinical data that led to the approval of varenicline in 2006, no serious neuropsychiatric adverse events were identified. It is important to note that based on varenicline's mechanism of action, as well as its non-clinical in vitro and in vivo profile, neuropsychiatric adverse events would not be anticipated. Yet, after varenicline was approved, neuropsychiatric safety emerged as a question.

This signal was identified in 2007 through postmarketing reports of serious neuropsychiatric, or NPS, adverse events. These postmarketing reports led to new warnings in Chantix's labeling

in 2008 and a boxed warning in 2009. Along with these labeling updates, a postmarketing requirement for a large prospective trial to evaluate neuropsychiatric safety of varenicline was issued. This postmarketing requirement is the EAGLES study.

While the EAGLES study was ongoing, FDA convened a joint advisory committee meeting in 2014. This meeting evaluated new evidence from large observational studies and meta-analyses of randomized controlled trials to determine if these data were sufficient to remove the boxed warning.

The data discussed at that meeting identified no increased risk of serious neuropsychiatric events with varenicline compared to placebo. The largest of the meta-analyses discussed was a pooled analysis of 18 double-blind randomized placebo-controlled studies. This pooled analysis included over 8,000 patients, some of which had psychiatric conditions at baseline.

The results showed a similar incidence in common psychiatric events in patients treated with varenicline compared to patients treated with

placebo. Of these 18 studies, 5 of them assessed suicidal ideation and behavior with the Columbia-Suicide Severity Rating Scale, or C-SSRS, and a meta-analyses of these five studies, including nearly 2,000 patients, was conducted. The results showed no increase in the incidence of suicidal ideation and/or behavior in patients treated with varenicline compared to patients treated with placebo, with a risk ratio of 0.79.

These meta-analyses were added to Chantix's labeling in 2014. After discussing these data, the committee voted to reassess the need for the boxed warning after the completion of EAGLES, and this brings us to the purpose of today's meeting.

The EAGLES trial has completed. The EAGLES trial is the largest prospective randomized controlled trial of smoking cessation medications ever conducted. The EAGLES results did not show a significant increase in serious NPS events amongst varenicline treated patients when compared to either placebo or over-the-counter nicotine patch treated patients. With respect to efficacy,

varenicline was also more effective than placebo, nicotine patch, and bupropion in helping smokers achieve abstinence.

These data have substantially advanced our understanding of the benefit-risk profile for varenicline both for patients with and without mental illness. An update to Chantix labeling is warranted to accurately reflect product safety and efficacy profiles to allow patients and prescribers to make informed choices. Today, we will discuss the novel design of EAGLES, its rigorous conduct, data analyses, and outcomes from this important study.

The key findings of EAGLES are shown on this slide. First, serious NPS adverse events occur in patients attempting to quit smoking regardless of treatment allocation. Second, serious NPS events are more common in patients with a psychiatric history than without regardless of treatment allocation.

As you have read in the briefing materials, EAGLES has two main study cohorts, those with and

without a history of psychiatric disease. In the non-psychiatric cohort, the incidence of serious NPS adverse events was low overall, and there was a small numerical decrease for varenicline compared to placebo.

In the psychiatric cohort, EAGLES has defined an upper bound for the risk of serious NPS adverse events as well as characterized the nature of those adverse events. The incidence of NPS events showed a small numerical increase in varenicline versus placebo that was not statistically significant.

This numerical increase in NPS events was not driven by events that were serious adverse events. It was not driven by events that were adverse events of severe intensity. It was not driven by events that led to treatment discontinuation. And it was not driven by events that led to harm to self or to others.

The results of EAGLES corroborates and adds to the totality of evidence from this signal investigation that does not support an increased

risk of serious NPS adverse events with Chantix treatment compared to treatment with placebo or over-the-counter NRT patch.

With the EAGLES data, an update to Chantix labeling is warranted to accurately reflect the benefit and risk profile of this important treatment. Product labeling should accurately reflect product safety and efficacy to allow patients and prescribers to make appropriately informed choices about treatment. Current Chantix labeling contains both a boxed warning and a warning regarding serious NPS events reported in people treated with Chantix in the postmarketing experience.

As we will discuss today, the totality of scientific evidence from this signal investigation, including meta-analyses of randomized controlled trials, large observational studies, and the outcomes of EAGLES does not support an increased risk of serious NPS adverse events with Chantix treatment compared to treatment with placebo or over-the-counter NRT patch.

Varenicline is the most efficacious smoking cessational treatment option available. It is an important tool combating the public health crisis caused by cigarette smoking. The boxed warning in Chantix labeling does not accurately reflect the NPS safety profile of Chantix. Furthermore, the boxed warning has the potential to deter the appropriate use of Chantix. As such, Pfizer believes that the boxed warning should be removed.

Pfizer proposes to retain the warning regarding serious NPS events occurring in patients attempting to quit smoking in the warnings and precautions section of Chantix labeling, and to update this warning based upon EAGLES. Pfizer believes that such a warning would sufficiently alert prescribers to the possibility of these types of events may occur in smokers attempting to quit.

Smoking is the leading preventable cause of death and disease. Beyond the numbers, though, I would imagine that each one of us in this room would not have to look too far to a close relative, to a friend, perhaps even ourselves, someone who

has suffered the ill effects of smoking or has the ill effects of smoking looming on the horizon.

Chantix is the most efficacious pharmacological aid for smoking cessation. Today you will consider how to best reflect these data in Chantix labeling so patients can make appropriately informed choices. This slide shows the agenda for our presentation, and it is now my pleasure to invite Dr. Prochaska to the podium.

Applicant Presentation - Judith Prochaska

DR. PROCHASKA: Thank you, Dr. Rusnak.

Good morning, everyone. I am Judith Prochaska. I'm an associate professor of medicine at Stanford University. I am funded by the National Institutes of Health as a principal investigator on multiple tobacco treatment clinical trials, including treatment studies with smokers with mental illness.

I also have published on large population surveys, examining tobacco use in smokers with co-occurring disorders. I provide consultation to Pfizer, the National Institutes of Health, and to

work groups of the FDA. I have no financial interest in the outcome of this meeting.

Thank you for allowing me the opportunity to present the independent observational study data. As some of you were here at the last FDA advisory committee, you saw the presentation of the observational studies. Observational studies provide great information and have great strengths, but they also come with some particular limitations.

The committee wanted the limitations mitigated by the EAGLES trial, which you will see presented today. Additionally, since the 2014 meeting, there are three new observational studies that are available. I will present on all six observational studies here with regard to strengths and limitations.

Typically of large size, controlled observational or population-based studies can provide reliable estimates. Observational studies provide real-world data on use of a drug by actual patients and can be designed to test hypotheses

about specific safety signal.

Many of the observational studies of varenicline included smokers with and without mental illness. Hence, the estimates are likely to be more generalizable than randomized controlled trials. The studies examined varenicline in relation to a variety of comparators, including nicotine replacement therapy, bupropion, and a no-treatment period. Limitations of the observational studies include reliance on existing data sources that may not report on all safety outcomes of interest.

Observational studies also lack a randomized design. In clinical practice, assignment to treatment is not by chance. Hence, it is possible that patients prescribed varenicline have a lower preexisting risk of neuropsychiatric adverse events. This could happen if, for example, clinicians were reluctant to prescribe varenicline to smokers who had a history of psychiatric disease. Therefore, the studies assessed the extent of such possible bias and adjusted for it

statistically.

To address these limitations, the studies here used propensity score analysis to equate the groups on measured known confounders or have utilized the self-control design to control for confounders. For unknown confounders, sensitivity analyses can determine how large the differences would need to be to alter the study conclusions.

Another concern is that differential reporting may occur due to a lack of a blinded placebo-controlled design; that is primed by a boxed warning, patients prescribed varenicline or bupropion may be more likely than patients prescribed NRT to report neuropsychiatric adverse events, or patients may be observed more closely by their clinicians for changes in thoughts or behaviors.

Acknowledging these limitations, controlled observational studies are of stronger methodological rigor than case or postmarketing reports. This is because the denominator is known, and data collection methods are more systematic;

that is, adverse events are assessed consistently across exposure groups.

So now let's review the six observational studies that altogether provide information in over 300,000 smokers. The studies ranged in size of approximately 10,000 to nearly 70,000 patients treated with varenicline. The studies included patients with and without a history of psychiatric disease treated in routine clinical practice.

These studies were conducted in a broad selection of populations from primary care patients in the United Kingdom, the entire populations of Denmark and Sweden, the U.S. Military Health System, which includes active duty and retired military and their dependents, and the U.S. Veterans Administration, which includes U.S. veterans and eligible family members and survivors.

The design of the studies is broadly similar. They estimated the rate of occurrence of designated neuropsychiatric events in patients who have received a prescription of varenicline versus a comparator such as nicotine replacement therapy

or bupropion.

The first four studies shown here compared varenicline to NRT. These are the adjusted effects from analyses design to equate the patient groups unknown and in some cases unknown confounders. The 95 percent confidence intervals that do not include 1, indicate a significant group difference.

Here, Meyer reported a reduced risk of outpatient visits for neuropsychiatric events for varenicline relative to NRT. Thomas reported a reduced risk of antidepressant treatment for varenicline relative to NRT. Kotz reported a reduced risk for depression in fatal/non-fatal self-harm. The Cunningham study found no difference in rates of hospitalization or outpatient visits for 6 of 7 psychiatric diagnoses. However, a greater likelihood of outpatients' visits was found among those with schizophrenia.

The reasons for the visits are unknown and may be that smokers with schizophrenia treated with varenicline were monitored more frequently in outpatient visits relative to those treated with

NRT. The magnitude of the difference indicated 5 more visits per 100 years of treatment.

The study by Molero was unique in using a self-controlled analysis where each subject served as his or her own control in a longitudinal analysis, the comparison being treatment with varenicline versus a non-treatment period. The strength is a subject-matched design controlling for factors that do not change over time. The weakness is that the design cannot control for time-varying compounds, namely the experience of quitting smoking and nicotine withdrawal.

While most outcomes evaluated in the Molero study did not indicate a significant difference for the varenicline versus non-treatment period comparison, one difference was found specific to smokers with a history of psychiatric illness indicating treatment for mood or anxiety symptoms. Notably, depression and anxiety are characteristic of nicotine withdrawal. Nicotine withdrawal has been found to be more severe amongst smokers with a history of mental illness.

A sixth study by Pasternak found no difference in neuropsychiatric risks between varenicline and bupropion. Four of the observational studies reported on fatal and non-fatal self-harm. Here, the most serious events were studies, and they were extremely rare, even with these large sample sizes.

Kotz, one of the studies from the UK National Health Service, had the largest sample, nearly 160,000 smokers. As such, the study had the largest number of observed fatal and non-fatal self-harm events. The summary estimate indicated reduced risk of harm for varenicline compared to NRT. The Molero study also identified a sizable number of serious adverse events among the nearly 70,000 smokers observed.

Notably, the timing of event was unrelated to treatment with varenicline. The hazard ratio estimates were 1 or lower. The estimate for Kotz was statistically significant and indicated a reduced risk of fatal or non-fatal self-harm for varenicline relative to NRT.

Altogether, the six studies looked at a variety of outcomes and comparators in real-world settings. Multiple outcomes were assessed, and most were not significant. In most cases, the findings indicated no increased risk for varenicline relative to NRT, bupropion, or no treatment. Further, the most severe events of self-harm were extremely rare.

So how do we weigh this evidence? The observational studies offer important methodological advantages over postmarketing reports, the largest being you have a comparator group and a known denominator within a defined patient population. This allows you to understand whether the rates being observed are different from what would be expected among the populations of smokers attempting to quit.

This is crucial because smokers as a group are at greater risk for mental illness and suicidal behavior, and through the act of quitting smoking are likely to experience agitation, aggression, anxiety, and mood disorders due to nicotine

withdrawal.

In science, we distinguish between levels of evidence. The postmarketing reports are useful in providing an indicator of potential signal. But going a step further in scientific vigor and evidence, we now have six published, independently conducted observational studies of varenicline, neuropsychiatric with over 300,000 smokers. These studies were in the U.S., in the UK, in Denmark, and in Sweden.

The next more rigorous step, an empirical investigation with increased controls for bias, is a randomized, blinded, placebo-controlled trial. EAGLES was designed to estimate the potential safety risk of interest, and it sampled over 8,000 smokers, half with current or a history of mental illness. Notably, as delineated in the briefing document, the EAGLES findings are highly consistent with the observational study data, providing increased certainty of the neuropsychiatric safety of varenicline among diverse groups of smokers.

I will now turn over the podium to

Dr. Robert Anthenelli, who will provide the details of the EAGLES study data.

Applicant Presentation - Robert Anthenelli

DR. ANTHENELLI: Thank you, Dr. Prochaska, and good morning, everyone. I'm Robert Anthenelli. I am professor and executive vice chair of the Department of Psychiatry at the University of California, San Diego. I chaired the EAGLES steering committee and was the principal investigator on this study. In the spirit of disclosure, I provide consulting services to Pfizer, and my university has received funding for research studies from the sponsor. However, I have no financial interest in the outcome of this meeting.

As previously mentioned, postmarketing reports of serious neuropsychiatric adverse events in subjects treated with varenicline led to labeling revisions in a postmarketing requirement. EAGLES was designed to satisfy this postmarketing requirement.

The primary objectives of the study were to,

one, assess if there were differences in the risk of clinically significant neuropsychiatric adverse events in subjects treated with varenicline, bupropion, nicotine replacement therapy, or placebo, and two, determine whether individuals with prior history of psychiatric disorders are at greater risk for serious neuropsychiatric adverse events compared with individuals without such a history.

This study also had a main efficacy objective to compare smoking abstinence rates among the four treatment groups. In this presentation, I will describe the EAGLES study design and will share how my involvement with the EAGLES trials has helped my thinking evolve on smoking cessation treatment.

EAGLES was a randomized, double-blind, 24-week study that included four treatments: varenicline, bupropion, nicotine patch, and placebo. Subjects were treated for 12 weeks. Nicotine patch, an over-the-counter product, which does not carry warnings regarding serious

neuropsychiatric adverse events, was used as an active control. The target sample size was 8,000 subjects with 2,000 per treatment group balanced by a history of psychiatric disorder diagnosis. The primary comparisons were varenicline versus placebo and bupropion versus placebo.

Since the details of the study design were included in the sponsor's briefing document, I will highlight just a couple of points on the design. Treatment began on day zero, and subjects were encouraged to quit on day 8.

Subjects assigned to the nicotine patch group received placebo/varenicline and placebo/bupropion during their first week. Active nicotine replacement was started at the week 1 visit when subjects were asked to quit smoking in keeping with the manufacturer's recommendation.

The primary safety endpoint, which will be described shortly, was a composite of neuropsychiatric adverse events that occurred during the treatment period plus 30 days. The study included adult smokers. Subjects with

imminent suicidal risk or displaying self-injurious behaviors were excluded from the study.

All subjects were screened for axis 1 and 2 diagnoses using the DSM-IV-TR criteria based on the structured clinical interview for DSM-IV disorders, also known as the SCID. The SCID diagnosis was confirmed by a psychiatrist or a clinical psychologist.

Based on the SCID, subjects who had no current or past psychiatric diagnosis were included in the non-psychiatric cohort. Subjects who met criteria of either a current or lifetime diagnosis for one or more of the DSM-IV diagnoses and were clinically stable were included in the psychiatric cohort. Subjects in the psychiatric cohort were further stratified based on which of four categories their primary diagnosis fell: mood disorders, anxiety disorders, psychotic disorders, and personality disorders.

Now, prior to EAGLES, there was no precise definition or precedent for what constituted a clinically significant neuropsychiatric adverse

event. Therefore, Pfizer developed, with input from the FDA, a composite primary safety endpoint for the study. The composite endpoint included a broad range of serious neuropsychiatric adverse events, which were chosen because they reflected the type of events reported in the postmarketing experience and listed in the Chantix label.

Inclusion of only neuropsychiatric events of moderate to severe intensity was chosen to increase the specificity of the endpoint by excluding neuropsychiatric adverse events that were less clinically significant as well as events that were typically associated with the nicotine withdrawal syndrome.

The 16 components, which make up the composite primary endpoint, are shown on this next slide. The primary safety outcome measure was the percentage of subjects reporting at least one of the following neuropsychiatric adverse events during treatment and up to 30 days after the last dose.

To be included in the composite endpoint, 4

of the components, those more frequently reported with withdrawal symptoms anxiety, depression, feeling abnormal, and hostility, were rated as severe in intensity by the investigator. The 12 other components, listed on the right, were rated as either moderate or severe in intensity. The 16 components, which were agreed with the FDA, include 261 MedDRA preferred terms, examples of which are shown on this slide for just 6 of the 16 components.

EAGLES was sized to attain an adequate level of precision in the estimation of the risk difference in the NPS composite endpoint. Based on the assumption of a 3.5 percent neuropsychiatric adverse event rate and the placebo-treated non-psychiatric cohort, and the 7 percent neuropsychiatric event rate and the placebo-treated psychiatric cohort, a study of 8,000 subjects would provide an expected margin of error of plus or minus 1.9 percent for the non-psychiatric cohort, plus or minus 2.6 percent for the psychiatric cohort, and plus or minus 1.6 percent for the

overall study.

An independent data monitoring committee reviewed unblinded safety data every 4 months, and as per agreement with the FDA, interim analyses were conducted at 50 percent and 75 percent of available data to ensure that the target sample size was correct. At each of the two interim analyses, the data monitoring committee recommended to continue the study as planned, and therefore the sample size remained as originally estimated.

Key secondary safety endpoints included an analysis of the percentage of subjects with severe-only neuropsychiatric adverse events within the primary endpoint and an analysis of the individual components that make up the primary endpoint. In addition, three psychiatric rating scales were used.

The Columbia-Suicide Severity Rating Scale recommended by various agencies, including the FDA, was used to assess suicidal ideation and behaviors at every clinic visit. The Hospital Anxiety and Depression Scale, a validated self-rating

inventory, was used to measure anxiety and depression. And the Clinical Global Impression of Improvement Scale tool was used to rate the severity of psychiatric illness and change over time.

This slide shows the subject disposition by treatment group and by cohort. Approximately 1,000 subjects were entered into each treatment group and each cohort. Approximately 80 percent of the subjects in each cohort completed the study, which is a relatively high percentage for this type of trial.

The sites were trained to make every effort to retain subjects in the study. If a subject did not return for a scheduled visit, the site made phone calls to reach the subject, and if unsuccessful, sent a certified letter. If a subject discontinued treatment but was not lost to follow-up, the subject was encouraged to remain in the study off treatment.

The baseline characteristics of the subject population are shown in this slide by cohort.

Subject in both cohorts were moderately nicotine dependent based on the Fagerstrom score with the psychiatric cohort slightly more nicotine dependent than the non-psychiatric cohort.

Not unexpectedly, about one-third of subjects in the psychiatric cohort have a lifetime history of suicidal ideation, and 12 percent had previous suicidal behavior. This contrasts with only 5 percent of the non-psychiatric cohort who had suicidal ideation and less than 1 percent having suicidal behavior prior to entering the study.

The baseline characteristics of the psychiatric cohort review of the highest percentage had mood disorders as their primary diagnosis, and less than 1 percent had borderline personality disorder. About half of the subjects in this cohort were taking a concomitant psychiatric medication at baseline, and those are the study characteristics, sample characteristics.

Let me digress for a moment and share, my experience helping to design and conduct and

interpret the study's findings have influenced my thinking on treating smokers with and without psychiatric disorders.

I think EAGLES has turned out to be a landmark study that will help clinicians and smokers better evaluate the benefit-risk ratio of using smoking cessation medications. We now have important new information to help us sort through what's always been a challenge clinically, and that is how to disentangle medication side effects from other potential causes of smoking cessation related mental changes.

This diagnostic dilemma becomes more complex in smokers with psychiatric disorders. Working on EAGLES has also made me consider the potential consequences of the boxed warning in the labeling. I believe it affects how patients might accept an initial trial of the medication and their tolerance of possible side effects.

In my opinion, the attention to the safety risks associated with varenicline has led many smokers who might benefit from the medication to

stay clear of it or to quickly assume that any change in mental state that they experienced during the quit attempt is directly attributable to the medication versus other potential causes. This potential misattribution and rush to judgment about possible side effects ultimately affects adherence to drug and cessation outcomes.

Regarding the diagnostic dilemma, let me share two actual cases from the EAGLES trial. Shown here are vignettes of two patients with bipolar disorder, case A, a 57-year-old man, and case B, a 40-year-old woman. Both knew they needed to quit smoking because it was affecting their health. However, both had concerns about using the non-nicotine smoking cessation aids due to publicity about their potential side effects.

Early on in the trial, both subjects reported adverse events. Case A began experiencing changes in his sleep and mood, which he immediately felt must signal that he was taking varenicline. Case B experienced worsening depression and anxiety, but considered that the study medication

may not be the culprit.

Although I wasn't sure if their mood changes were due to study medication, tobacco withdrawal, an exacerbation of their illness, or other psychotropic medications they were taking, Case A stopped taking study medication on his own, continued in this trial until its end, but he never quit smoking. Case B agreed to a dosage reduction, quit smoking, and their psychiatrist added another atypical antipsychotic to her regimen, which improved her mood.

About a month ago, when I unblinded myself to the study results at our site, I learned the subjects' assignment. Case A was taking placebo and Case B was taking varenicline. Thank you this morning for your attention. With this background in mind, I'd like to now reintroduce Dr. Rusnak who will review the study's conduct.

Applicant Presentation - James Rusnak

DR. RUSNAK: Thank you, Dr. Anthenelli.

One of the questions, question 2, that is being posed to the committee today relates to

EAGLES data collection, adverse event coding, and application of the case definition on the ascertainment of the primary endpoint. It is in this context that I will now present some key aspects regarding the execution of EAGLES. In addition, you will also hear today some analyses and comments regarding the conduct and analysis of EAGLES, and we are prepared to answer the committee's questions regarding these matters.

EAGLES was designed to capture a unique, complex, and subjective endpoint. EAGLES was executed as designed and captured this endpoint. Additional measures were taken in the study to ensure data quality, which I will briefly describe. The study protocol included tools aimed at the standardization of the collection of NPS events. As a result, there was a wealth of information regarding NPS safety collected in this study on which to base conclusions.

Mental health professionals were required to be affiliated with each site to confirm the SCID diagnoses and evaluate adverse events of interest

associated with the primary endpoint. External medical professionals were used to help ensure psychiatric patients were properly diagnosed prior to randomization. They were also used to provide training on the SCID and the neuropsychiatric adverse event interview.

NPS adverse events required additional attention, and a multi-pronged approach to the ascertainment of these events was incorporated in the protocol. NPS adverse events were captured by any of the following means: volunteered adverse event reporting, which is the routine method for the collection of adverse events. In addition, EAGLES augmented adverse collection by actively soliciting events using the NAEI; collecting proxy reports, and through the C-SSRS.

Only events that were deemed to be adverse events by the investigator were reported as such. The NAEI and proxy reporting are special attributes of EAGLES that will be further described in the next two slides.

The Neuropsychiatric Adverse Event Interview

was developed by Pfizer in partnership with academic collaborators and was originally used in a study using varenicline in a patient population with depression. Prior to use in EAGLES, an additional clinical study further refined this interview in a patient population that fit the inclusion criteria of EAGLES.

Per study protocol, if a subject has a positive response to any item in this interview, a determination was made by the investigator as to whether it met the criteria for an adverse event. The interview was intended to enhance the primary endpoint collection, not replace volunteered reporting of adverse events. The same can also be said related to proxy reporting for the collection of NPS adverse events.

The contact card the study participants received is shown here. Patients were encouraged to share their participation in EAGLES with their professional and personal acquaintances and ask them to call their study doctor on their behalf should they potentially display any of the listed

neuropsychiatric events. Given that the primary endpoint rests on the collection of adverse events, it is necessary to understand this process in detail.

This slide depicts the investigator's key role in adverse event reporting and the strengths of ascertainment of the primary endpoint. Investigator verbatim terms are the foundation of adverse event reporting in EAGLES. One hundred percent of all adverse events collected had an investigator verbatim term.

Severity is also assessed by the investigator using their clinical judgment. Adverse event reporting begins with the identification of a medical event. The medical event could come from a variety of sources, such as the patient, a proxy reporter, laboratory results, response to questionnaires, or investigator observations.

Once a medical event is identified, the investigator then determines if that event meets criteria for AE reporting. The investigator

describes the event in concise medical terminology. That investigator terminology is then coded in MedDRA, and two things happen.

First, the MedDRA preferred term is reported in the general adverse event reporting, and secondly, that investigator terminology coded is queried to determine, one, if it meets any one of the 261 preferred terms in the primary NPS composite endpoint; and if so, if it met severity criteria. If the answers to those questions are yes, then that MedDRA coded preferred term is a primary NPS composite endpoint.

Importantly, it is these investigator verbatim terms, not subject verbatim terms, that ultimately code through MedDRA the primary composite NPS endpoint. Subject verbatim terms, meaning what the patient voiced as a symptom to the investigator, were naturally used by the investigator in their medical assessment of the potential adverse event. The subject verbatim terms were collected, which is atypical and required site education, but these subject verbatim

terms were not further utilized in aggregate data analysis.

Overall, the acquisition of subject verbatim terms for adverse event reporting was very high, and in the over 24,000 adverse events reported, the subject verbatim term was available in 98.9 percent.

The benefits of using investigator preferred terms however are twofold. First, it utilizes the patient's reported symptoms, meaning the subject verbatim term with the addition of real-time clinical observation for any physical or behavioral science exhibited by the patient, and the benefits of medical assessment by a trained professional of the patient verbatim description in its full context.

This process is in contrast to coding the subject verbatim term that would simply rely on the judgment of a code or using the subject's verbatim term who has no contact with the patient. Secondly, it facilitates accurate coding by encouraging the use of appropriate and unambiguous

medical terminology and also affords the opportunity to provide medical diagnoses of symptoms expressed by the subject.

For example, the ambiguous subject verbatim of, quote, "I've not been feeling myself for a few days," end quote, could be taken many ways. In postmarketing reports, this subject verbatim may likely code to feeling abnormal. However, in the context of a clinical study with the oversight of an investigator, additional clinical insights and solicitation of information can be undertaken.

In this case, the investigator reported nausea and dypsnea as the adverse event terms that best characterized why the patient had not been feeling themselves for the past few days. As in clinical practice, this process was based upon the investigator's clinical judgment.

Naturally, there will be investigator-to-investigator variability in their clinical judgment. This variability however is spread across all treatment groups evenly due to the randomized, blinded treatment allocation.

Conclusions can be effectively drawn on the observed relative rates of events. Moreover, sensitivity analyses can and have been performed to assess the robustness of the primary NPS composite endpoint. These sensitivity analyses tested a variety of potential confounding factors, and each of these sensitivity analyses has supported the primary analyses.

One final point on the study execution merits clarification. The primary endpoint of EAGLES is defined based on all treatment-emergent adverse events irrespective of relatedness or causality assessment.

We have already heard from Dr. Anthenelli today two clinical vignettes that underscore the difficulties in ascribing relatedness versus the many factors that come from attempts to quit smoking, versus the patient's underlying medical condition or concomitant medications. This difficulty in ascribing causality is perhaps best exemplified amongst placebo-treated patients in EAGLES.

As you will see in our blinded study, NPS adverse events are also associated with placebo treatment at comparable levels to active treatment both in frequency and relatedness. This difficulty is in part why we're here today. We are here because postmarketing reports identified a safety signal, but there is no control group, and reporting is subject to bias.

Interpretation of that signal required the conduct of a randomized, blinded trial to provide the level of evidence to support or refute the signal. Through the strengths of the blind to remove bias and through the interpretation of the relative frequency of all-cause, treatment-emergent NPS adverse events, we can understand whether this safety signal from postmarketing reports has been supported or has been refuted with this rigorously conducted trial just described.

I would now like to invite Dr. Russ to present the EAGLES data.

Applicant Review - Cristina Russ

DR. RUSS: Good morning. My name is

Cristina Russ. I am a medical director in the varenicline team with Pfizer.

This is the order in which the study results will be presented. We will start with a key outcome of the EAGLES study, the primary neuropsychiatric adverse event composite endpoint. The observed incidence of the endpoint for the overall study population is shown on the left, the non-psychiatric cohort is shown in the middle, and the psychiatric cohort on the right.

The vertical axis shows percent of subjects with at least one event meeting the prespecified criteria for the primary endpoint. The incidence is based on adverse events reported by the investigators regardless if considered treatment related or not. The incidence is similar across treatment arms in the overall study population, around 4 percent. The incidence in the non-psychiatric cohort ranges from 1.3 percent for varenicline, the blue bar, to 2.5 percent for NRT, the purple bar. In the psychiatric cohort, it ranges from 4.9 percent for placebo, orange, to

6.7 percent for bupropion, green.

This data confirmed that subjects experienced serious neuropsychiatric adverse events when attempting to quit smoking regardless of treatment, including in the placebo arm. The data also reflect a higher incidence in the psychiatric cohort across all treatment arms. The data suggest that the results may defer by cohort when comparing varenicline and placebo.

The statistical analysis is shown on the next slide. The risk differences at 95 percent confidence intervals for each active treatment versus placebo for the overall study population are shown on this plot. The vertical line through zero indicates no difference. The point estimate for the risk difference for varenicline versus placebo is very close to zero.

Risk differences for each cohort separately will be shown next, the non-psychiatric cohort on the top and the psychiatric cohort on the bottom. In the non-psychiatric cohort, the risk differences for active versus placebo are close to or lower

than zero, showing a small numerical decrease. Associated 95 percent confidence intervals are below or include zero. In the psychiatric cohort, the risk differences are higher than zero, showing a small numerical increase. The differences are not statistically significant, and 95 percent confidence intervals include zero.

The small numerical decrease in the non-psychiatric cohort and the small numerical increase in the psychiatric cohort seen for varenicline versus placebo have the same magnitude, between 1 and 2 percent. More precisely, the risk differences are minus 1.28 and plus 1.59, respectively.

The primary endpoint was the result of a careful balance between specificity and sensitivity, and was based on the define selection of terms and severity ratings, however, we did conduct further analysis to better understand the endpoint. We will start with a sensitivity analysis that expands the neuropsychiatric endpoint with the intent to minimize the impact of potential

variability in classification of events by investigators. It was conducted after the briefing document was submitted. The results were consistent with a prespecified primary analysis and with a sensitivity analysis, included by the FDA in their briefing document.

The expanded neuropsychiatric endpoint includes all subjects meeting the prespecified primary endpoint. It also includes additional subjects identified by a blinded clinical review of cases of neuropsychiatric worsening captured by the psychiatric scales or by the mental health evaluation.

The intent is to capture any relevant event potentially missed. Also, it includes all subjects with moderate events included in the components: depression, anxiety, hostility, and feeling abnormal. These components were included in the primary prespecified endpoint but only if they were rated severe by the investigators.

The reason behind the prespecified criteria was to minimize noise based on nicotine withdrawal.

This addition to the expanded endpoint equalizes the threshold of severity required across all components in order to be counted for the endpoints, so minimizes the impact of the variability in the classification of events by investigators.

Last, the expanded endpoint also includes all subjects with a moderate or severe adverse event of irritability. Irritability was not part of the primary endpoint, but was now added because some events of anger could be judged to be irritability.

The incidence of the expanded neuropsychiatric endpoint is shown in contrast with a primary prespecified endpoint. It is highlighted. It is approximately 5 percent in the non-psychiatric cohort and approximately 12 to 13 percent in the psychiatric cohort. The expansion appears proportional across treatment arms.

If we look at the data broken down by category, we can see that the main contributor is the addition of moderate events of the prespecified

components, particularly depression and anxiety. The cases identified by the clinical review added fewer than 5 subjects per treatment arm in each cohort. Irritability added fewer than 8 subjects.

The difference between varenicline and placebo in the sensitivity analysis is consistent with a prespecified analysis, and the statistical analysis confirms this conclusion. The stable list of risk differences and 95 percent confidence intervals were all comparisons between active and placebo, for the sensitivity analysis and for the prespecified analysis. The risk differences for varenicline versus placebo are highlighted. They are very consistent between the two analyses.

We will now review an analysis that narrows the endpoint to the most severe events, or serious adverse events, or events leading to treatment discontinuation. While the total number of subjects reporting any neuropsychiatric event in the non-psychiatric cohort in the primary prespecified endpoint are shown on the left, the subset of subjects with at least one event rated as

severe in intensity by the investigator is shown on the right.

This subset was a prespecified secondary analysis. Please note that we switched from percent to number of subjects on the vertical axis, and the data is shown descriptively.

As seen in the first set of bars, a total of 13 subjects met the primary endpoint for varenicline, 22 for bupropion, 25 for NRT, and 24 for placebo. On the right, we see one varenicline subject with an event rated as severe by the investigator and 5 such subjects for placebo. We can also add serious adverse events, such as life-threatening or leading to hospitalizations, in events that led to treatment discontinuation to this analysis.

We see the combined data for all the steps of events in the endpoint now on the right. The results show a lower number of subjects treated with varenicline with events that were rated severe, or were SAEs, or events that led to treatment discontinuation when compared to placebo.

The same analysis is presented for the psychiatric cohort; on the left, totals; on the right, subjects with events rates severe, 14 for varenicline and 13 for placebo. Again, we added SAEs and events leading to treatment discontinuation.

The combined data is now shown on the right. We see 26 subjects for varenicline and 23 subjects for placebo. The denominator is around 1,000 subjects per treatment arm. So therefore, the small numerical difference observed between varenicline and placebo for the entire primary endpoint in the psychiatric cohort was not driven by events rated by investigators as severe, or serious adverse events, or events that led to treatment discontinuation, but rather by moderate events.

We will now review the components of the prespecified endpoint with a more in-depth review of events that have the potential to or resulted in harm to others, or to self. The 16 components of the primary end point for the non-psychiatric

cohort are now presented. The graph shows the number of subjects with events by component. Subjects could be counted in one or multiple components, depending on the terms reported by the investigators.

Some of the components are a part of the same syndrome, such as anxiety and panic, or hostility and aggression, and could potentially be seen more as a continuum than as highly distinct categories. The most frequent type of event is agitation. It includes moderate and severe events. Agitation is the only component for which the statistical analysis could be conducted, and it showed the results that are similar for varenicline and placebo. For all the other components, there are fewer than 5 subjects per treatment arm.

The 16 components are shown for the psychiatric cohort now in the decreasing order of frequency for varenicline. Agitation is the most frequent component also in this cohort. The statistical analysis does not show significant differences between varenicline and placebo.

There were no ones for hostility, homicidal ideation, or suicide as can be seen on the right. The rest of the components showed differences between varenicline and placebo of 3 subjects or fewer, except for aggression, the second component shown on the graph. As mentioned, we conducted a more in-depth review of aggression and suicide related events, as they could result in harm to self or other, and we still start with aggression.

The table shows the number of subjects who met the primary endpoint due to events mapped to this component to aggression in each cohort. There were 3 subjects for both varenicline and placebo in the non-psychiatric cohort and 14 versus 8, respectively, in the psychiatric cohort. The tables also show the number of subjects with events in this component that were rated as severe, that were SAEs, or led to permanent treatment discontinuation, and this number was of 2 subjects or fewer per treatment arm and similar for varenicline and placebo.

We also did a qualitative review of these

cases. We reviewed the data for the component aggression subject by subject. The majority of cases were verbal aggressions or feeling of anger. All the events in this component resulted in a physical act of aggression -- so other than verbal against people or objects are summarized in the table by using verbatims from subjects. The varenicline cases did not involve other people but throwing objects. The most severe case resulting in significant harm to another person was in the placebo arm.

The three components in the primary endpoint, including suicidal ideation, suicidal behavior, and completed suicide are shown for the non-psychiatric cohort now. There were no cases mapping to these components in the varenicline arm. You can see the number of subjects in the other treatment arms, which were mapped to these components, and the short description of the suicide behaviors. This included an NRT subject who cut wrist and the completed suicide in a placebo subject who jumped from a monument.

In the psychiatric cohort, for suicidal ideation, there were 5 subjects for varenicline, 2 for bupropion, 4 for NRT, and 2 for placebo with events in this component. The highlighted row now shows number of subjects with suicidal behavior: varenicline, one subject who cut wrist 20 days after last dose of medication, case considered by the investigator non-suicidal, but this case is included in the primary endpoint; bupropion, one subject inhaled gas from a cigarette lighter; placebo, one subject took an overdose of psychotropic medication on day 9. There were no completed suicides in the psychiatric cohort.

So hence, the review of suicide behavior did not reveal more severe cases for varenicline.

While the neuropsychiatric endpoint in the study is novel, the psychiatric scales that will be reviewed now are broadly used in clinical trials. Their outcome did not show an increased neuropsychiatric risk for varenicline, important for the triangulation of evidence. We will continue the review of the suicide related events

and their enhanced collection through the Columbia scale.

The table shows positive answers on the Columbia-Suicide Severity Rating Scale during treatment plus 30 days. As expected based on lifetime history, there are fewer reports in the non-psychiatric cohort on the left than in the psychiatric cohort on the right.

The highlighted row shows the cases of suicide behavior. There is one varenicline subject in this row in the psychiatric cohort. The subject heard voices to jump in front of a bus and was included in the primary endpoint for hallucinations.

All the other suicide behaviors captured by Columbia scale and now shown in the highlighted row are included in a primary endpoint in the component suicide behavior and were described earlier, with the exception of a placebo subject who took 4 bottles of study medication. This was considered by the investigator non-suicidal and reported as an overdose. The subject was though included in the

expanded sensitivity analysis in the placebo arm.

The row highlighted now shows subject with yes answers for suicidal ideation. Positive answers are presented regardless of the investigator's interpretation of the answers and the adverse event reporting. We see similar numbers for varenicline and placebo: non-psychiatric cohort 9 and 7 subjects; psychiatric cohort, 29 and 26.

The most concerning ideations that have the highest predictive value for suicide behavior are the ideations with some intent and/or specific plans, a 4 or 5 on the scale. There were no such ideations reported as the most severe type of event on the Columbia scale for varenicline. For placebo, there were 2, and for NRT, 1 in the psychiatric cohort.

In summary, the total number of subjects with positive answers for any ideation, passive or active, and/or behavior was similar for varenicline and placebo in the non-psychiatric cohort, 9 and 8 subjects; and psychiatric cohort, 30 versus 28.

We will now review the last two scales, the Hospital Anxiety and Depression Scale and the Clinical Global Impression of Improvement. The analysis presented in the briefing document for these two scales show very similar data for the average weekly scores for all treatment arms. Additional analysis will now be presented that assess the worsening of severity and also shows similar outcomes for varenicline and placebo.

This graph is for the HADS Anxiety subscale. It shows percent of subjects with an increase in category of severity at any time during treatment, plus 30 days versus baseline, non-psychiatric cohort on the top, psychiatric on the bottom. The bars on the left show any worsening of category, while the bars on the right are the subset for an increase from a score below 11 to 11 or higher, so a shift in the most severe category.

We see a high percentage in the psychiatric cohort versus non-psychiatric cohort across all treatment arms, as for other outcomes. Most importantly, we see similar or lower percentages

for varenicline versus placebo. The other HADS subscale for depression follows a similar pattern; again, similar results, differences of less than 1 percent for varenicline versus placebo or versus NRT.

The last scale, the Clinical Global Impression of Improvement, again, non-psychiatric on the top, psychiatric on the bottom, the graph shows percent of subjects with a worsening of their status as reflected by this scale versus baseline at any time during treatment plus 30 days. The categories are minimally worse, much worse, or very much worse, and results look very similar within 0.5 percent difference for varenicline versus placebo in both cohorts.

We will now briefly review the efficacy outcomes. Varenicline did prove to be the most effective therapy tested, confirming previous studies and meta-analysis. The results shown are for the main efficacy endpoint for the non-psychiatric cohort. The graph on the left indicates the percent of subjects reaching

continuance abstinence during the last 4 weeks of treatment, weeks 9 through 12. Odds rate [ph] shows 95 percent confidence intervals shown on the right.

Varenicline achieved the highest abstinence of 38 percent. Odds rat shows varenicline versus placebo are 4, and versus bupropion and versus NRT, 1.7; 95 percent confidence intervals are narrow showing robust effects.

To account for relapse expected to occur during the non-treatment follow-up, a long-term abstinence rate was also prespecified, and this continues abstinence rates for weeks 9 to 24 now shown in the second bar chart. The differences between varenicline and NRT or varenicline and bupropion continued to be statistically and clinically significant.

The results in the psychiatric cohort showed the same ranking of effectiveness. The abstinence rates are lower than for the non-psychiatric cohort in all treatment arms. The continued abstinence rate for varenicline for weeks 9 to 12 is

29 percent, odds ratio for varenicline versus placebo greater than 3.

There was no interaction between treatment and cohort, meaning that the effect of the treatment was not dependent on psychiatric history, and both cohorts do benefit from active treatment. So varenicline shows significantly higher effectiveness than NRT and bupropion in both cohorts at the end of the treatment, and the difference remains statistically significant through the non-treatment phase.

EAGLES provided the first head-to-heard comparison between the three approved smoking cessation pharmacotherapies in the large placebo-controlled trial. The results did not show an increased risk of neuropsychiatric adverse events in the composite primary endpoint in the overall study population for varenicline versus placebo or versus NRT.

In all treatment arms, including placebo, the incidence of the primary endpoint was higher in the psychiatric cohort than in the non-psychiatric

cohort. When looking at the risk difference for varenicline versus placebo, we do see in the non-psychiatric cohort a small numerical decrease for varenicline, in the psychiatric cohort, a small numerical increase. However, the difference did not reach statistical significance and was not driven by events that were rated as severe, or were serious adverse events, or events that led to treatment discontinuation, or resulted in harm to self or other. A sensitivity analysis that expanded the endpoint was consistent with a primary analysis.

The outcomes of the psychiatric scales did not show an increased neuropsychiatric risk for varenicline versus placebo or versus NRT. Varenicline was shown to be the most efficacious treatment in both cohorts.

Thank you. I would like now to introduce Dr. Eden Evins to share her views on the clinical implications of these results.

Applicant Presentation - Eden Evins

DR. EVINS: Thank you, Dr. Russ.

Good morning. My name is Eden Evins. I'm pleased to provide a clinical perspective on EAGLES outcomes. I'm a psychiatrist. I serve as director for the Center for Addiction Medicine at the Massachusetts General Hospital and as the Cox family associate professor of psychiatry in the field of addiction medicine at Harvard Medical School. I treat smokers with and without serious mental illness, and I teach medical students, residents, and practicing clinicians about tobacco addiction and smoking cessation treatment.

I was an investigator in the EAGLES trial, and I've conducted as principal investigator nine additional randomized controlled trials of smoking cessation treatment, and those were schizophrenia, schizoaffective disorder, bipolar disorder, and major depressive disorder, as well as those without a mental illness.

By way of disclosure, I provide consulting services to Pfizer, to NIDA, and to various universities and institutes for grant review, and I've been compensated through my university for my

work as an investigator for the EAGLES trial, but I have no financial interest in the outcome of this meeting.

I started this line of work during my fellowship when, in the same week in the schizophrenia program at the Mass General Hospital, one of my patients who smoked 3 packs a day died from sudden cardiac death at age 40. And in the same week, another patient of mine, also a heavy smoker in his 40's, began to attend the clinic with portal oxygen that he needed just to walk around because of his severe emphysema. I'd like to share with you why as a clinician, teacher, and researcher I think that the EAGLES trial results are extremely important.

First of all, why do we find ourselves in the nearly impossible position of trying to prove the absence of an association between a medication and an important category of adverse events? One reason is that the initial trials of varenicline excluded smokers with psychiatric disorders. And by excluding smokers with psychiatric illnesses,

these trials had no opportunity to evaluate with the aid of a control group the psychiatric adverse event with varenicline in smokers with a comorbid psychiatric illness.

When varenicline came into general use and worrisome events such as aggression, suicidal behavior, or violent behavior were reported, the medication was blamed; action by regulatory agencies was taken; restriction of varenicline by many formularies was made; and there was reluctance on the part of many physicians to prescribe varenicline.

The EAGLES trial now provides the data that we needed then. The EAGLES trial is rightly viewed as a landmark study by clinicians and researchers because of its many firsts. It's the first trial to compare safety and efficacy of all FDA-approved smoking cessation medications in large samples of smokers with and without psychiatric illness. In fact, it's the first smoking cessation trial at all to include a large number of patients with psychiatric illness, and this group consumes the

majority of cigarettes purchased in the United States.

EAGLES in fact enrolled more smokers with major depressive disorder, more smokers with bipolar disorder, more with anxiety disorder and with schizophrenia spectrum disorders than any prior smoking cessation study. So it's thus the first study to allow comparison of both safety and efficacy of all FDA-approved smoking cessation treatments between smokers with these psychiatric disorders.

Because of the broad enrollment criteria, EAGLES results are relevant in a range of clinical settings from primary and specialty medical care to community mental health centers. Those in both cohorts had on average a moderate level of nicotine dependence based on the Fagerstrom score and were thus at risk for nicotine withdrawal symptoms.

Those in the psychiatric cohort had stable psychiatric illness. Half were on psychotropic medication in order to be stable, but many were quite symptomatic at baseline with symptoms such as

delusions, hallucinations, and ongoing depressive symptoms despite best treatment.

The distinction between stable and symptomatic is important. Excluding patients with unstable psychiatric illness is medically prudent for patient safety and scientifically sound to reduce noise in the primary outcome. But including those who are symptomatic despite best treatment makes the sample relevant and useful to clinicians treating a wide range of smokers.

In the EAGLES cohort, half of those with major depressive disorder had a more severe form of the illness with repeated episode called recurrent major depressive episodes. These folks are shown to have an increased risk of neuropsychiatric adverse events during a smoking cessation attempt. One-third of the psychiatric sample had a comorbid psychiatric illness. A fourth had a comorbid substance use disorder that was in remission at the baseline visit, and one-eighth had made a prior suicide attempt in their lifetime. So overall, these are the types of patients that we see in

clinical practice.

Another reason EAGLES is so impactful is that the efficacy conclusions are clear and unambiguous. Varenicline is more effective than bupropion and nicotine replacement therapy, which are each more effective than placebo. This was the case in smokers with and without psychiatric illness, and this is great news. The fact that the EAGLES efficacy findings are consistent with prior findings also raises confidence in the findings.

In addition to confirming our prior understanding of relative efficacy, the EAGLES trial extends what we know about safety by quantifying the rate at which neuropsychiatric adverse events can be expected when smokers attempt to quit smoking on each of the three FDA-approved smoking cessation treatments and placebo, and doing so in the largest sample ever studied.

So the EAGLES results tell us that we can expect about 2 percent of smokers without a mental illness who try to quit to have some kind of neuropsychiatric adverse event during their

cessation attempt regardless of treatment. And they tell us that smokers with a mental illness who try to quit, we can expect 5 to 7 percent to have some kind of neuropsychiatric adverse event, again independent of treatment.

So the results from EAGLES to my eyes show no pattern in the most worrisome of neuropsychiatric adverse events, including hostility, aggression, severe depression, or suicidal ideation or behavior. And as I will show, no psychiatric subgroup appears to be at a particularly increased risk from varenicline compared to placebo.

One of my central concerns as a clinician, teacher, and now as a researcher is that smoking cessation treatments are underutilized by clinicians, underlying our slow progress in reducing smoking rates in the general population, and are virtually no progress in reducing smoking rates among those with psychiatric illness.

Now, clinicians are seeing in a very large randomized controlled trial, in a sample patient

population relative to their practice, that the AE rate is essentially the same with placebo as with active treatments, as shown here for the non-psychiatric cohort, and may make a difference in what they recommend to their patients who smoke.

It may shift their risk-to-benefit assessment and recommendations away from thinking that inaction or delay in prescribing a smoking cessation aid is the safest course of action toward implementing clinical best practices for smoking cessation, an action that dramatically increases the odds of smoking cessation, which is lifesaving.

In this slide, in subjects with psychiatric illness, we see similar findings regarding neuropsychiatric safety and efficacy. These figures from subgroup analyses within the psychiatric cohort show the neuropsychiatric adverse event rates on the top and the end-of-treatment abstinence rates on the bottom, broken out by both treatment assignment and psychiatric diagnosis.

With varenicline, the likelihood of success

with abstinence for smokers with a mental illness is increased by a factor of 2 in those with mood disorders and by over 5 in those with psychotic disorders with no significant increase in neuropsychiatric adverse events rates versus placebo.

You can see here that the neuropsychiatric adverse event rate, on the top row, is actually the same with varenicline as with placebo in the psychotic disorder group and the anxiety disorder group, and not significantly increased in the mood disorder group. This is a very important finding.

Moreover, while varenicline more than triples the odds of abstinence compared to placebo for the whole psychiatric cohort, the odds of abstinence with varenicline versus placebo in those with schizophrenia is over fivefold, and this is because the placebo quit rate of 4 percent is so low.

This 4 percent placebo quit rate is consistent with prior studies. It's what is shown in the Cochrane review of smokers with

schizophrenia. Yet, as is precisely in this subgroup, smokers with psychosis, that varenicline is the most underprescribed and for whom the mortality disparity due to smoking related illness is the greatest. It's now estimated at 28 years compared to the general population.

The efficacy and safety findings from EAGLES are consistent with my experience as a clinician and as a PI of many smoking cessation trials. However, despite the very consistent high quality evidence regarding neuropsychiatric safety of smoking cessation medications, clinicians often attribute any neuropsychiatric adverse events during a cessation attempt to the medication, particularly when using varenicline.

By way of example, in a trial of maintenance varenicline versus placebo, that I was the PI of, published in JAMA for prevention of relapse to smoking, my wonderful mentor, Don Goff, a superb clinician who some of you know, attributed to varenicline a psychotic decompensation in a person with stable, treated, symptomatic schizophrenia.

At his request, we stopped study medication for this person, and some of her symptoms resolved over the coming weeks. However, when we broke the blind, we later found she had been on placebo.

We've had similar situations in prior trials. A woman in a trial of bupropion plus NRT versus placebo plus NRT became manic and psychotic. She broke into the BU law library one night and was found by campus police the next morning and brought to our clinic. She had been reading legal text all night, though she had no connection with the law school, and her clinician attributed the mania to bupropion she could have been on in the study, and asked that she discontinue study medication. We did, and later learned she had been on placebo plus nicotine patch.

I emphasize with prescribers and with patients, who are often wary about using varenicline due to the boxed warning and the negative press about psychiatric risks of the drug, that there are neuropsychiatric adverse events observed in trials as they are observed in

practice. But in trials, we can see that they're not different by treatment. And now with the EAGLES data, we can quantify that risk. With controlled trials and large operational studies, we see that significant neuropsychiatric adverse events have been reported by smokers trying to quit, but the events are independent of treatment.

What do we know about why significant neuropsychiatric adverse events occur in smokers trying to quit? Smoking is an addiction. Thus, it is a psychiatric illness in and of itself with associated increased neuropsychiatric adverse events, notably suicide, even in those not trying to quit.

The neuropsychiatric adverse events seen during smoking cessation attempts, independent of treatment, is consistent with our understanding of the brain changes that occur with chronic nicotine exposure. The act of trying to quit smoking in and of itself is stressful, and it is associated with some psychiatric symptom instability independent of both treatment and abstinence.

The good news is that now in a large meta-analysis recently published, long-term smoking cessation is associated with improved depression, anxiety, stress, and other symptoms as well as self-reported quality of life in addition to living longer.

In my opinion, it's critically important to increase the use of effective smoking cessation treatment in all smokers, particularly those with mental illness. Smokers with serious mental illness are more likely to smoke, more likely to smoke heavily, and to be physiologically dependent on nicotine. They are less likely to be able to quit smoking without a medication cessation aid, and are more likely to relapse after discontinuation of medication cessation aids.

Smokers with mental illness are less likely to receive a pharmacotherapeutic cessation aid from a medical provider. This is universally recommended treatment, a standard of care combined with behavioral treatment, something we aim to improve.

This contributes to the shocking mortality gap, now up to 25 years in those with serious mental illness compared to the general population, largely due to diseases judged by the CDC to be causally related to tobacco smoking. EAGLES demonstrates that the most effective pharmacologic treatment is varenicline by a significant margin in all groups, and the difference is not small.

My group is now involved in three large studies. Two are ones funded by NIMH and one by PCORI, that aim to increase utilization of effective smoking cessation treatments for smokers with mental illness. In these studies, I meet with primary care doctors and psychiatric providers and discuss their attitudes and knowledge about the risks and benefits of smoking cessation treatments and how this impacts their prescribing behavior.

In these interactions I find overwhelmingly that providers overestimate the risk of tobacco dependents' treatment, especially varenicline. Prescriber cite the boxed warning as evidence that varenicline has been proved to cause major

disturbance and psychiatric illness, which we see is not true. And this leads doctors to question whether prescribing varenicline for their patients is safe.

Further, doctors underestimate the benefit of varenicline for quitting. Many express the common societal misconception that if smokers were really motivated, they would quit on their own. In reality, medication significantly increased the odds of quitting. The evidence is consistent in multiple observational and randomized controlled trials. Varenicline is the most effective FDA-approved smoking cessation medications.

These doctors tell us that based on their risk-to-benefit calculation, they often consider not prescribing or delaying prescribing smoking cessation aids to be the more conservative action for many smokers, particularly those with mental illness.

We now have the opportunity to add EAGLES trial results to the labeling for varenicline, and this represents an opportunity to better

communicate the current understanding of the actual risks and benefits of varenicline in light of the large body of high quality new data.

With EAGLES trial corroborating these other studies, it's time to unring the alarm bell on varenicline. It's time to make greater use of the most effective smoking cessation, varenicline, for our patients who try time and again to quit smoking but fail without the optimal treatment that they deserve.

I appreciate your time in considering this important issue, and I'd like to now give the podium to Dr. Jim Rusnak.

Applicant Presentation – James Rusnak

DR. RUSNAK: Thank you very much for your valuable insights, Dr. Evins.

I would now like to present our closing slides, and then we will be pleased to take your questions. The overall summary of efficacy and safety of varenicline from EAGLES is shown on this slide. Efficacy is displayed on the left, NPS safety on the right. The upper panel show results

from the non-psychiatric cohort with the lower panels showing results from the psychiatric cohort.

With respect to efficacy, varenicline has been confirmed to be superior not only to placebo but also to bupropion and nicotine replacement therapy. The superiority of varenicline over other available treatment options is clear in both the non-psychiatric and psychiatric cohorts, and its benefits for smoking cessation are undisputable.

With respect to safety, EAGLES was specifically designed and conducted using a composite NPS adverse event endpoint to evaluate the concerns raised by postmarketing reports regarding the neuropsychiatric safety of varenicline. With respect to safety findings from this study, we have shown the general observations that, first, serious NPS adverse events occur in patients attempting to quit smoking regardless of treatment allocation.

Secondly, serious NPS adverse events that were reported for patients taking placebo in both cohorts were generally consistent with those

reported for varenicline in both this study and in the postmarketing experience. Furthermore, as shown on the right-hand panel of this slide, the incidence of NPS events in the non-psychiatric cohort of the composite endpoint was low overall, and there was a small numerical decrease for varenicline compared to placebo.

In the psychiatric cohort, EAGLES has defined an upper bound for the risk of NPS events, as well as characterized the nature of these adverse events. There was no statistically significant increase in the incidence of serious NPS events in the psychiatric cohort for varenicline versus placebo.

The numerical increase in NPS events observed in this cohort was not driven by events that were serious adverse events, events that were severe intensity, or events that led to treatment discontinuation, or events that led to harm to self or others. The confidence interval for NPS events observed in varenicline treated patients broadly, but not completely, overlapped the confidence

interval of over-the-counter nicotine replacement treatment.

Collectively, the EAGLES safety outcomes showed no increased risk of serious NPS events with varenicline compared to placebo or compared with NRT, regardless of the patient's psychiatric history. These safety data from EAGLES, combined with the efficacy outcomes, significantly increased the understanding of the benefit-risk profile for varenicline.

With the EAGLES data, an update to Chantix labeling is warranted to accurately reflect the benefit and risk profile of this important treatment. Product labeling should accurately reflect the product safety and efficacy profile to allow patients and prescribers to make appropriately informed choices about treatment.

As we have shown today, the totality of scientific evidence from this signal investigation, including meta-analyses of randomized controlled trials, large observational studies, and the outcomes of EAGLES, does not support an increased

risk of serious NPS adverse events with Chantix treatment compared to treatment with placebo or over-the-counter NRT. Of note, serious NPS events occur with NRT, and NRT labeling does not currently include any NPS warnings.

Varenicline is the most efficacious smoking cessation treatment option available. It is an important tool in combating the public health crisis caused by cigarette smoking. The boxed warning in Chantix labeling does not accurately reflect the NPS safety profile of Chantix. Furthermore, the boxed warning has the potential to deter the appropriate use of Chantix.

As such, Pfizer believes the boxed warning should be removed. Pfizer proposed to retain the warning regarding serious NPS events occurring in patients attempting to quit smoking in the warnings and precautions section of Chantix labeling, and to update this warning based on EAGLES. Pfizer believes that such a warning would sufficiently alert prescribers to the possibility that these types of events may occur in smokers attempting to

quit.

Smoking is the leading preventable cause of death and disease in the United States. Varenicline is the most efficacious smoking cessation treatment option available. People need help achieving their goal to quit smoking and to derive the benefits of smoking cessation.

Today is an important day. Today, you will make recommendations on Chantix labeling revisions so that patients and physicians can make appropriately informed choices. I would like to thank the study participants of EAGLES, our study investigators, and we are pleased to receive questions from the advisory panel. Thank you.

Clarifying Questions to Applicant

DR. PARKER: Thank you. We'll now turn to any clarifying questions to the sponsor, to Pfizer. And I'll ask that you place your cards up like this so we'll get you on the list. And Kalyani will make a list here, and we'll call on folks.

Let me remind you that we are 19, and we are going to take a break at 10:15, so if you will

kindly keep your questions specific and to the point. I'd also suggest when possible that you address them to a specific speaker. That's helpful. Please state your name for the record before you speak, and the list has started.

Dr. Narendran, please?

DR. NARENDRAN: Raj Narendran. I had one question. I noticed that a prior treatment of varenicline wasn't an exclusion to get into the study. What was the rationale for that?

DR. RUSNAK: I'd like to ask Dr. Anthenelli to speak to this question, please.

DR. ANTHENELLI: Robert Anthenelli, University of California, San Diego. That was part of an effort to improve the generalizability of the sample. Most smokers in the United States have tried to quit many times, and many of them have tried all varieties of medications. So we did that for that reason.

DR. PARKER: Dr. Higgins?

DR. HIGGINS: Thank you. I have a couple of subgroup analyses questions, which I think would be

illustrative, particularly as I think about the psychiatric population. And I don't know to whom my question should be addressed, so I just state the questions, and perhaps you can field them, Pfizer.

A question about the use of any behavioral interventions in the EAGLES study, I know that bupropion has been shown to be effective in patients with schizophrenia when coupled with CBT, for example. Another question relates to the assessment of the use of any tobacco on medication blood levels. And the third question regarding the comparison between atypical antipsychotic medications and typical antipsychotics because patients on atypicals generally have an easier time quitting.

Were any of these explored?

DR. RUSNAK: I'll ask Dr. Anthenelli to address the first comment regarding behavioral interventions.

DR. ANTHENELLI: Robert Anthenelli again. A three-part question, and I'll do my best.

Part one, which was related to the behavior, you're right. There have been studies done in special populations of individual psych who have schizophrenia or recurrent depression, which have found that more intensive psychotherapies aid smoking cessation better than standard.

In this case, however, we used a standard smoking cessation treatment. It was delivered 10 minutes per session at every clinic visit, and it was based on clearing the air, the standard booklet used by the National Cancer Institute, the agency for health quality research kind of guidelines.

The second part of your question again, if you don't mind?

DR. HIGGINS: Use of tobacco on blood levels for medication.

DR. ANTHENELLI: Correct. Smoking cessation, which cigarette smoke induces the liver to break down certain psychotropic medications, that can actually influence potential side effects during a smoking cessation effort. However, blood

levels of those psychotropic medications were not measured during the EAGLES trial.

Your last one was around the --

DR. HIGGINS: Antipsychotic versus atypical sources, typical antipsychotics.

DR. ANTHENELLI: We carefully recorded, of course, all of the medications that patients were taking in the trial. We did not see an overall effect of medications on treatment. There was actually one analysis that did show that the people taking more medications were likely to have slightly higher more adverse events. We've not done any particular subanalyses to look at atypicals versus first-generation antipsychotics, however.

DR. PARKER: Dr. Roumie?

DR. ROUMIE: Dr. Anthenelli, don't sit down.

Christianne Roumie. Two questions. The first is that there appeared to be the neuropsychiatric inventory was collected multiple times throughout the 12 weeks, and I didn't really see mentioned how you dealt with the multiple

measures in ascertaining the outcome. For example, if someone didn't say a symptom on week 2, but then endorsed weeks 3, 4, and then 5 it went away because of a dose reduction, how is that patient handled?

Then, the second question relates to the blinding procedures. Sometimes clinical trials will ask site investigators at the end what group they thought each patient was assigned to as a way of assessing the effectiveness of the blinding strategies. And I think I just would like to know if you did something like that.

DR. ANTHENELLI: Sure. The first part of your question was related to the ascertainment of neuropsychiatric adverse events. Given that that was the focus, the primary aim of the study, much emphasis is placed on that being standardized. So participants were first asked about a general, non-specific question, "How have you been feeling over the last week?" Any volunteered adverse events would of course had been recorded and embellished as far as their presence or absence.

Then any of those adverse events were of course followed up. So if there was an adverse event that was shown, at each subsequent visit, the rater would actually go back and ask about the event, if it was still active or not. And then in adverse event reporting, we track an adverse event until the adverse event is resolved or not. In some instances, if the adverse event is going, you track it all the way through the trial, and that was of course done.

The same thing done with the NAEI. The NAEI, the Neuropsychiatric Adverse Event Interview, which has 25 items on it, that was asked at all of the clinic visits. At the time the NAEI was done -- in fact, why don't we go ahead, if we could please, and project the neuropsychiatric adverse event interview on the screen so that everybody has a sense of it.

This 25-item questionnaire was asked of the participants at every clinic visit. Regardless of had they reported a spontaneous -- and if an adverse event had been observed -- and this was I

think one of the major innervations of the EAGLES trial, its effort at sensitively assessing neuropsychiatric adverse events.

The trainers, the raters of course have been trained on the instrument. They trained at the investigator meetings. They continued to have buffer training done every six months via webinar training. There were videos used and actor portrayals used to actually rate participants with a variety of psychiatric complaints in the use of the instrument, and this was the instrument that was used throughout the study.

Now, all of this information in addition to the results of, say, the HADS, which is a self-report, Hostile Anxiety Depression Scale, and of course any findings that were obtained in the Columbia-Suicide Severity Rating Scale, that ultimately all funneled down into how the adverse events were determined by the investigator.

Determination of the adverse events first looked at the frequency and duration of the complaints, and then of course there was the

assessment of their severity using a standardized scale that, again, was detailed, and great emphasis in the protocol and in all the training that went along with the study conduct.

Did that answer your question?

(Dr. Roumie nods in the affirmative.)

DR. ANTHENELLI: So if a person answered on the NAEI that they complained of a mood complaint, that would then be written down, and then an effort would be made to determine what was going on.

Let's say that the person had temporary -- was late to the appointment because their bus was late, there was an effort of trying to tease out what was going on in that complaint. And then all of those reports were funneled into the decision by the investigator, what was going on and what type of adverse event that might represent.

DR. PARKER: Dr. Morrato?

DR. MORRATO: Thank you. Elaine Morrato. This question is for Dr. Rusnak. So I agree with Pfizer's statement that investigator verbatim

reporting is really the cornerstone of the ascertainment.

In the FDA's briefing document, they note numerous concerns about trial conduct related to that, citing things around capture of the adverse events, inadequate training of some of the investigators, coding of events, et cetera. In the briefing document, the FDA's Office of Scientific Investigation noted that they were wanting to do site visits to several sites, and at the time of the briefing document, reports hadn't been related.

Obviously, this affects concern under ascertainment of adverse events. So I wanted to better understand Pfizer's position on what was reported in FDA's briefing document, and whether or not you have any sensitivity analyses related to some of these concerns that can help the committee's deliberation.

DR. RUSNAK: Overall, Pfizer took extraordinary measures to collect data in EAGLES. Over 8 and a half million data points were collected, and Pfizer stands behind the data and

the results of EAGLES.

I think there's actually two fundamental core issues that were raised in the briefing document for FDA. I think that one of them is perhaps best illustrated by example. In the FDA briefing document, it was stated that in many cases, no verbatim term for the adverse event was recorded at all, so it is not possible to determine how coding was assigned or how severity was assessed.

I think that there is a fundamental disconnect here because we have 100 percent of all investigator verbatim terms, and that's in fact how all of the adverse events were coded in the trial, the general adverse events, as well as the adverse events that rolled up into the primary NPS composite endpoint.

I have mentioned earlier in today's presentation, we took additional measures to ensure that we collected all of the NPS adverse events that were possible. In the primary endpoint, we had 323 NPS adverse events. About

half of them came from voluntary reporting of adverse events. The other half of them came from the direct solicitation of adverse events through the NAEI usage.

With respect to the sensitivity analyses, one way to interrogate the integrity of the data for the primary endpoint is to conduct a multitude of different sensitivity analyses, looking at confounding factors. Pfizer has conducted these sensitivity analyses, and some of the sensitivity analyses have also been shown in the FDA's briefing document. And both the Pfizer and the analyses that were shown in the FDA's briefing document are very supportive of the primary NPS composite endpoint.

DR. MORRATO: Can you share any of that data, not just your conclusion? So you reference in your briefing document two particular sites that your own audit found were troublesome, and then you excluded them in the analysis.

Did you do any other site interaction testing? This is a global trial. There's also

investigators that were noted in the FDA's document that were receiving Pfizer payments. Those affected certain countries more than others. I'm not saying all investigators are doing that. This is good quality, but what is the analysis where you look at dropping those out?

So if you look at some of them -- I'll just stop with that. Not just your statement; I'd like to see data.

DR. RUSNAK: I'll invite Dr. Gaffney to share that data. Before Dr. Gaffney shares that data, I will note that those two sites that you mentioned were detected by a proactive effort of Pfizer's quality management system. And Pfizer audited those sites and provided full details of that in a prospective fashion. The results were disclosed within the clinical study report. And before the study was unblinded, the statistical analysis plan was updated to include the sensitivity analyses that Dr. Gaffney will describe.

DR. GAFFNEY: Good morning. Mike Gaffney,

statistician at Pfizer. Your question addresses I believe a very important issue that's raised in the FDA briefing document and is part of the EAGLES study. And it has to do both with the variance of the primary composite end point rate among all the sites within the study, which is what you're addressing, and the proper method to analyze that.

FDA in their briefing document has identified unexpected variability in the primary event rate. However, you have to keep in mind that that unexpected variability is being conducted under the assumption that there's a common rate among all of these sites that is operating.

That assumption is not realistic, nor is it necessary for a proper analysis. It's not realistic for two reasons. One, as Dr. Rusnak indicated, we have investigator judgment going on over 140 sites. And the investigator judgment is the primary strength and underpinning of EAGLES, so that's expected to vary over the sites. The second is that the characteristics of the patients can vary over the sites when we're conducting a study

within the U.S. as well as outside the U.S.

We've done analysis that have identified important patient characteristics, which affects the primary endpoint in the study. We'll certainly be happy to share that with you if the committee desires.

But the second point I want to make, though, is that with respect to the methods, Dr. Andraca-Carrera has done an analysis within the briefing document motivated by this excess variability, but remember that excess variability is under the assumption of a common rate.

What I would like to say about that is that that analysis is on the relative scale. As Dr. Anthenelli indicated in designing this study, very little was known about what actually these absolute rates were within these treatment groups, within the populations, so our primary analysis stayed on the absolute scale and the risk difference scale.

In doing that, you have to also recognize that there are 323 events collected in the entire trial. There are 140 sites. So the idea of trying

to look at what's going on differentially among the treatment groups between the sites is futile. What we did was to sum up all of the sites that were in the U.S., and sum up all the sites that were in the non-U.S. regions, and use region as a factor in our analysis. It gave me very interesting results -- which we can also share with you if the committee so desires later -- that is very indicative of more events being reported in the U.S. than outside the U.S.

So I want to leave you with the idea that these analyses that FDA did, that Dr. Andraca-Carrera did, and that we did as the primary analysis are not at odds. They're complementary analysis getting at the same question.

I would invite the committee to look at the forest plot of our analysis, which was presented by Dr. Russ, and compare it directly with the analysis that Dr. Andraca-Carrera did for the FDA. You will see not only the same overall conclusions, but you will see pretty much the same lift of the

confidence intervals on all of the comparisons, albeit, ours is on the risk difference scales; FDA is on the relative scale.

DR. PARKER: So we have 7 more in the queue, and we'll take one brief pointed question before the break and see how time goes. I know it's important to get questions clarified, but unfortunately we have a very long agenda for the day.

So Dr. Marder, if you will pose your last question here before we take a break. Thank you.

DR. MARDER: Yes. I think this should be for Dr. Russ. I noted that the proportion of people with psychotic illnesses was relatively small. I think it was something like 9.5 percent. I'm wondering if you could comment on that.

Does that indicate that the place where recruiting was done, that there weren't more severely ill patients? And can this trial tell us anything about the risk of these adverse events in people who are psychotic, or is the sample size just too small?

DR. RUSNAK: I'd like to invite Dr. Evins to speak to that point, please.

DR. EVINS: Thank you, Steve. It's a great question and sort of near and dear to my heart. The sample size was 9.5 percent for psychotic disorders, but it amounted to about 390 patients. So it actually is the largest trial ever done in people with schizophrenia. It's the first randomized controlled trial data for the efficacy of NRT in people with schizophrenia, and the sample with schizophrenia was I think representative. About 95 percent of people with schizophrenia in the trial were on antipsychotic medication. Many were on two. We enrolled 67 at our site.

So I do think that we can draw conclusions. I would have liked for the sample size to have been larger, but it is the largest ever done in the world.

DR. PARKER: So we'll now take a 15-minute break. Panel members, please remember that there should be no discussion of the meeting topic during the break among ourselves or with any member of the

audience, and we will resume at 10:30. Thank you.

(Whereupon, at 10:19 a.m., a recess was taken.)

DR. PARKER: Thank you, everyone. Since we left seven questions on the table, I'm going to move us right along. We'll go through the FDA presentations, and we will then get directly to questions to the FDA about their presentations, with the hope that then we can return to some of the ones that were left on the table for Pfizer. So that's the order as we move forward.

Thank you for a quick break, and let's proceed now with the FDA presentations.

FDA Presentation - Celia Winchell

DR. WINCHELL: Good morning. I'm Celia Winchell, the team leader for addiction products in the Division of Anesthesia, Analgesia, and Addiction Products. My task is to present the FDA's clinical review of the PMR trial. In my presentation this morning, I hope to provide some insight into the thinking that went into FDA's recommendations for the development of the protocol

for the postmarketing safety outcome trial, and then some observations based on our review of the data from the completed trial.

I'll be enumerating a number of issues identified in our review of the data, but I want to emphasize that this is part of the FDA review process, to explore the data for problems that might limit our ability to rely on the trial to support conclusions. However, as you will see, we were able to conduct a number of sensitivity analyses, and we will present conclusions that we believe the trial can support.

As Dr. Racoosin reminded us, the initial concern about neuropsychiatric adverse events associated with smoking cessation products, varenicline and bupropion, was prompted by spontaneous postmarketing reports involving Chantix. These cases were often quite detailed and specific, and included features strongly pointing to their being drug related, such as temporal relationship to initiating, titrating, or discontinuing the drug; dechallenge and rechallenge

findings; and patients clearly reporting that they had never had an experience like it associated with a quit attempt.

Our review of our own postmarketing adverse event database identified similar cases involving Zyban, but not involving nicotine replacement products. And among the first questions we had internally was how often is this happening? Some of the cases were very serious and concerning, involving suicide, aggressive behavior, or debilitating symptoms. However, given the importance of both drugs, which are both effective at helping people quit smoking, we wanted to be able to compare the risks to the benefits. Without knowing how often these severe and serious events occurred, it was hard to know how to do this.

Additionally, one important unanswered question was whether patients with preexisting psychiatric conditions were at greater risk for adverse events, and additionally whether they had a similar prospect of benefit. Neither drug had been studied for smoking cessation in patients with

preexisting mental health conditions, and it's generally understood that this is a population with a high rate of smoking and a great difficulty in quitting.

Therefore, it seemed important to get at these questions. How often do events of a serious or seven nature occur when people are taking smoking cessation drugs? Is one drug more or less of a concern than another, and do patients with psychiatric conditions have a different likelihood of either harm or benefit?

Because of the seriousness of some of the cases and the potential that harm could be minimized if problems were quickly identified and the drug discontinued, we require that both Chantix and Zyban be labeled to alert patients and prescribers that cases of serious neuropsychiatric events had been reported and to advise that the drugs be discontinued if these events occurred. Meanwhile, we began to work with the sponsors to answer the questions that would allow us to quantitate the risk in a defined population and

compare it to the benefit for patients both with and without preexisting psychiatric conditions.

Next, our deliberations turned to how to answer these questions. This turned out to be very challenging. We knew the study needed to include four arms because we wanted to quantitate the risk of neuropsychiatric events for both Chantix and Zyban, but also to make sure we understood the risk associated with nicotine replacement, and because our adverse event reporting system tends not to capture information for over-the-counter drugs, as well as for prescription drugs. And we wanted to facilitate a benefit versus risk comparison, so a placebo group would allow us to establish an efficacy rate in the trial. And we knew we needed two cohorts, one with a history of psychiatric diagnoses and one without.

Some initial discussions focused on establishing the incidence of some narrowly defined outcome, such as completed suicide, or psychiatric hospitalization. However, trials employing these endpoints would have needed to be even larger than

this one turned out to be, and would have not captured the full picture of the types of events that were reported in postmarketing spontaneous reports.

The adverse events are coded using MedDRA, which has over 20,000 preferred terms and over 70,000 lower-level terms. And even for events that are described in similar language by the patient, it was common to see different MedDRA terms applied in coding, or to see non-specific terms, such as feeling abnormal, applied to situations that involved experiences that had important impacts on patients' functioning.

To identify adverse events with similar concepts, the MSSO, the MedDRA folks, have developed a number of standardized queries, the standardized MedDRA queries, SMQs, that pull together terms associated with a particular syndrome, or a problem, from whichever body system they might be assigned to in MedDRA. Examples include SMQs for neuroleptic malignant syndrome or anaphylaxis.

So there are existing SMQs for depressed mood and suicidality, and for hostility and aggression, but there is no SMQ for a syndrome that encompasses experiences of disturbances in thinking, feeling, and functioning like we were seeing in postmarketing Chantix and Zyban cases, and that led to our novel approach.

To try to define the endpoint of interest, we marshaled observations about the types of experiences that were reported in the postmarketing cases, identified concepts we wanted to capture, and asked the sponsors to develop a tool to prospectively ask patients about their experiences and document them; and to develop a list of MedDRA terms that covered the scope of these experiences, and combine them into a composite. It's essentially the same as a standardized MedDRA query.

We wanted to find a way to capture events involving mood disturbances such as depression or suicidality or mania; events involving hostility and aggression and homicidal ideation; or the

emotional experience that's sometimes described as agitation; events involving perceptual abnormalities or psychotic experiences like delusions and hallucinations, paranoia, psychosis; events of anxiety or panic; and events that defined other descriptions and are characterized as feeling abnormal.

The intent here was to avoid noise by excluding mild events because some emotional and cognitive symptoms, like irritability and impaired concentration, are well recognized symptoms of nicotine withdrawal encountered during smoking cessation, and some symptoms may be expected in patients quitting smoking without pharmacotherapy. The composite outcome focused on adverse events of severe intensity, in some cases moderate intensity, as reflected by the degree of functional impairment experienced by the patient.

So the items that I just listed are five broad concepts broken into 16 narrower terms. These were agreed upon in the protocol, but the choice of the specific MedDRA terms matching to

each of the concepts left to the sponsors to determine was not reported until the statistical analysis plan was described in the interim analysis. And in the end, as you heard this morning, there are 262 different MedDRA terms in this composite.

This is a reminder of the events that were included in the composite endpoint, where you heard that very clearly presented this morning. And we also heard about the instrument that was developed to ensure the events of interest were identified, the neuropsychiatric adverse event interview.

This was intended to be administered by trained interviewers as a semi-structured interview, and any positive responses would be followed up in order to get a full picture of the context of the symptom, co-occurring symptoms, and a rich narrative of the event. There were supposed to be follow-up questions for clarification, frequency, duration, severity, and degree of functional impairment related to the symptom.

Sample follow-up questions were provided in

the training materials, and the interviewer was instructed to probe as needed to assess the subject's experiences and to make an appropriate assessment. And narratives were supposed to be constructed for NPS cases to pull together all the relevant information from reports who include the patient, significant others, healthcare providers, and other sources.

We heard already this morning that information was collected in a variety of ways: routine queries about adverse events, clinical rating scales for Anxiety Depression and Global Functioning, the Columbia-Suicide Severity Rating Scale; and along with the neuropsychiatric adverse event interview, these sources were used to identify symptoms of interest.

Enough detail was supposed to be obtained to understand the impact on the patient. Symptoms not interfering with the subject's usual function were not to be included as cases for the endpoint, and some symptoms were only to be included if a problem interfered significantly with a patient's usual

function.

I should note that the FDA staff and sponsors went into this process realizing that this trial presented a number of challenges. We discussed the need to ensure consistency across raters and across the many different sites, and the sponsors did provide repeated training sessions for the investigators in an attempt to ensure a consistent approach.

The protocol called for full verbatim narratives to be recorded so that enough information would be available to do some adjudication of cases after the fact. Coding was centralized. The interviews were conducted in a local language at each site. However, with sites in 16 different countries, there are inevitably some cross-cultural differences and language differences, as well as differences in how sites familiar with psychiatric patients assess psychiatric symptoms as compared with sites that are not as familiar.

Of course, it was understood that this was a

novel primary endpoint. Ideally, additional work to validate it would have preceded this study, but in the interest of shedding light on this important question, we move forward knowing that there might be some bumps in the road.

So with this understanding of what FDA expected from the trial, I'll move on to what we observed in our review of the data. This is a quick reminder of the basic design of the trial, which was already presented quite clearly to you by Pfizer this morning.

Everyone began on active or placebo tablets during week 1, and then began applying either active or placebo patches during week 2. Visits were initially weekly and then bi-weekly during treatment, and then we had monthly follow-up visits to week 24.

Moving on to study results, as you heard over 8,000 patients were randomized into this study at 140 different centers. The cohorts of patients with psychiatric history -- I might call it the PHx cohort, and patients without psychiatric history, I

might call that the non-PHx cohort -- were roughly equal in size and randomization. Across the treatment arms was 1 to 1 to 1 to 1.

Primary diagnoses of patients in the psychiatric cohort were primarily effective disorders, about 70 percent, followed by anxiety disorders and psychotic disorders. And as you've already seen, specific diagnoses that were eligible for inclusion are listed here.

We've seen this already, the trial disposition. I'll explain that patients stay in the study if they discontinued medication. They could also complete the course of medication but not come to all the follow-up visits. So it was possible to complete the study but not treatment, or to complete treatment but not the study. These are listed separately.

The number of subjects who completed the study, meaning they were followed for the full 24 weeks, was similar in both cohorts and between treatments. The number of subjects discontinuing treatment prematurely were slightly higher in the

psychiatric cohort, and the lowest proportion of treatment completion was 71.4 percent in the subjects randomized to placebo in the psychiatric cohort.

Most frequent reasons for treatment discontinuation were no longer willing and treatment related adverse event. Subjects on placebo were more likely to say they were no longer willing, and review of what they meant by that is displayed by lack of efficacy; they didn't think it was helping. Subjects randomized to placebo were less likely to discontinue for adverse events, especially in the non-psychiatric cohort.

Moving on to efficacy results, which we wanted to see in order to weigh the benefits versus risks, the study demonstrated that all three treatments are effective in patients with and without psychiatric diagnoses. We also explored whether this conclusion held true if patients who had previously had an unsuccessful experience with one of the treatment drugs were excluded, and it did. So I'll repeat, all three treatments are

effective in patients with and without psychiatric history.

The main results of the trial showed that serious and severe clinically significant neuropsychiatric events did happen in this population of 8,000 smokers trying to quit smoking. They occur more frequently in patients with psychiatric conditions than in patients without prior psychiatric diagnoses.

Serious events of a neuropsychiatric nature were reported, about 6 per 1,000, in patients with psychiatric history. Events severe enough to affect functioning but not meeting the regulatory definition of seriousness were reported in about 9 to 12 percent of patients with psychiatric history and about 3 to 4 percent of patients without psychiatric history, including patients on placebo.

These numbers that I'm citing are from an analysis conducted by Dr. Andraca-Carrera that's a bit different from Pfizer's based on sensitivity analyses he conducted to address some concerns that I identified in the review of the safety outcome

trial, and you'll hear about them in a moment.

Before I turn the presentation over to him, let me present to you some of the issues we identified in the review of the data. I'll remind you this is what we do. We dive into the raw data, and we see whether there are problems that prevent us from relying on it to support conclusions.

As you will hear, there are some issues, but in the end, we believed after conducting a variety of sensitivity analyses that we could support certain conclusions based on this trial.

My review observations fall into a few different broad categories. I'll describe some issues that relate to data collection, resulting at times in incomplete or inadequate understanding of events. I'll describe issues related to data coding that created some obstacles to review and limited the extent to which we could place confidence in certain analyses, like the analyses of the subcomponents.

I'll discuss the ways in which the data were reported that didn't meet our expectations. I'll

mention some findings that raised concerns about data reliability and some issues about specific terms that somehow ended up wrongly assigned in the analysis.

The first issue is that, at many sites, it looks from the data as if the NAEI was not used as it was intended. There are items recorded in the adverse event database that just say, "patient answered yes," to a question on the inventory. Some sites simply didn't write down the patient verbatims at all, and the field for that information says, "not captured," or "not recorded," or "N/A," or "missing," similar words. There are a few that's labeled, "event described by reporter."

Sometimes an investigator term would say "anxiety." This actually happened a lot, investigator terms says "anxiety," and the event described by reporter, in that column it says, "as anxiety." In other words, the event was described as anxiety. So there's not anything more than that in that column.

We understand that the sponsor did make an attempt to educate the sites about the importance of capturing the patient verbatim information, but with 139 sites and some of them having upwards of 40 investigative staff at a single site, there seems to have been a problem implementing this.

Another issue with how the data were collected is that sometimes key information seemed not to have been obtained. For example -- I've put one here -- "patient died in a head-on collision," and the report doesn't say who was the driver of the car.

So when an adverse event involves an accident, we're often interested in learning whether the accident could be related to an effect of the drug. In this case, we know that the effects of these study drugs on patients' cognition and perception have been a concern. So a patient being killed in a motor vehicle accident could potentially be a very concerning drug related event if he were at the wheel, but unlikely to be drug related if we were a passenger.

In this case, the information on who was operating the car was not provided, may not have been obtained. As it happens, this patient was on placebo, but it is an illustration of the way key pieces of information were not collected or provided for review.

As I mentioned, in a well-intentioned effort to reduce the noise in the data and to focus only on the types of events that had a significant impact on patients, the types of events that we had seen in postmarketing adverse event reporting, investigator assessments of severity were incorporated into the primary endpoint, and only events with an impact on a patient's functioning were included.

However, in the implementation, a great deal of inconsistency in assessment of severity was obvious. Where patients' verbatims were available, we could see that two patients describing events of similar impact could be coded differently with respect to severity. It seemed as if some investigators found the idea of missing a day of

work significant, while others did not. In another example, a patient with depression required hospitalization, and this event was assessed as mild by the investigator.

So to address this, Dr. Andraca-Carrera performed some additional analyses that incorporated events that were not part of the protocol specified endpoint definition, and you'll see results of his analyses in a moment.

There also seemed to have been some issues with coding of the data. So cases in which the patient verbatim, where available, were identical could have been coded to different MedDRA terms. And in some cases, the patient verbatim said one thing, and the investigator preferred term was something different. This could be significant if the patient in his own words endorsed a symptom that was part of the NPS endpoint, such as anger, but the investigator coded this to a term that is not part of the endpoint, such as irritability, and that did happen.

There also seems to have been considerable

variation in interpretation of the word "agitation." In the NPS endpoint, it's meant to capture a sense of emotional upset, but it appears it was inconsistently applied, and often it's a code for events of motoric restlessness or akithesia.

Patients who reported a variety of symptoms sometimes recorded one term rather than identifying all the symptoms, or just one of their many symptoms, would be considered as significant enough to qualify for the primary endpoint, even though they had a constellation of symptoms. And for that reason, we concluded that any type of analysis of the very subcomponents of the primary endpoint was unlikely to be informative.

In other issues, there were patients for whom an adverse event was recorded where the term selected is a psychiatric diagnosis, not a symptom. The documentation doesn't allow us to determine whether this is a coding error, should the term "depression" have been selected instead of the term "major depression," or genuinely a new diagnosis.

So there are patients in the non-psychiatric cohort who are coded to a psychiatric diagnosis. That would be very significant if it were a new diagnosis, but these were not necessarily flagged as NPS events.

Finally, there were some coding errors that were evident from the review of the way the patient verbatim was translated to investigator's select, and then to a MedDRA term -- I've got some examples up here -- that raised some concern that there might have been some unfamiliarity with psychiatric terminology, or language barriers, or other reasons that might have led to inaccurate coding. There are always errors in coding. These are just some things that caught our attention.

The next issue pertains to how the data was collated and presented to give a full and clear picture of each of the NPS cases. We had stressed from a protocol stage their full narrative, incorporating all sources of data needed to be constructed for each NPS case. And in the end, this was not done as we expected.

Instead, the narratives provided very little other than the MedDRA terms, with start and stop dates for the events, sometimes start and stop dates for concomitant medications, scores on the instruments, but not integrated, that gave the impression of being automatically generated from the case report forms. They provided no context, background, or coherent story, and the patient verbatims, even when collected, were not incorporated into the narratives.

As a result, we were left with narratives that raised more questions than they answered. For example, the narrative of a patient who experienced a skull fracture did not report how the skull fracture occurred. Now, as it turns out, it occurred in the context of an altercation with her boyfriend, which is a very type of event that we're interested in hearing about and understanding how it happened.

This is only one example of many, and we had to ask the sponsor to generate new narratives for the cases that they'd identified as NPS cases, and

even these didn't necessarily present a coherent picture, and they didn't incorporate findings from all the data streams like all the anxiety and depression scales.

Now, earlier today, Pfizer presented an additional analysis of an expanded NPS endpoint that actually did attempt to incorporate findings from all the psychiatric rating scales and adverse event information, everything altogether. And that's very helpful, but that was not included in the study report. In fact, we just got it last week. But this was not a feature of the study report.

Some other issues were identified in our review of the tabulations of protocol violations and the required financial disclosure information and the sponsor's reports of their audits. Some of these issues could impact data reliability, and we did perform sensitivity analyses with and without these sites.

Finally, there are a few issues with the specific terms in the endpoint. For example, for

some reason the term "dysphoria," according to the protocol, was quoted as "aggression." That just seems like an error.

As I mentioned, after identifying these concerns, we then attempted to determine whether any of these issues would preclude our ability to rely on the main conclusions of the study. Based on our concerns about the coding, we concluded that the analyses of the various subcomponents of the primary endpoint were unlikely to be reliable and probably were uninformative, so our analyses focused on the overall incidence of NPS events.

To explore the impact of the concerns identified, Dr. Andraca-Carrera conducted several sensitivity analyses. These included evaluating the impact of the heterogeneous finding across sites, and he'll go into that in some detail.

We also looked, as we always do, at the impact of including or excluding sites for issues of data reliability where identified, and sites where investigators had disclosable financial interest according to our regulations. We also

extended this to sites where personnel were involved in an ongoing way with speakers bureaus for the sponsors. These analyses showed no impact on the overall conclusions, and we won't be presenting those.

Additionally, to address concerns about investigators applying a lower level of severity rating to some events than would seem warranted, or assigning MedDRA terms that took the event out of the NPS endpoint, Dr. Andraca-Carrera evaluated whether widening our net to include some additional adverse events would change the conclusions, and he'll present the findings of those analyses.

I'll come back to my conclusions after we hear from Dr. Andraca-Carrera.

FDA Presentation - Eugenio Andraca-Carrera

DR. ANDRACA-CARRERA: Good morning, everyone. My name is Eugenio Andraca-Carrera, and I'm a reviewer in the Office of Biostatistics at CDER. Today, I will talk about our statistical review of the PMR trial for smoking cessational products.

Dr. Winchell has described the trial design and the primary neuropsychiatric endpoint, so I will start my presentation with a discussion of the statistical methodology of the PMR trial, followed by a discussion of its primary results, as well as sensitivity analyses, and analyses of additional endpoints that we conducted.

As a reminder, the primary objective of this trial was to estimate the risk of neuropsychiatric adverse events associated with bupropion, nicotine replacement therapy, and varenicline in each of the two trial cohorts, patients without a history of psychiatric illness and patients with a history of psychiatric illness.

The statistical analysis plan did not prespecify the statistical hypothesis to be tested, and the trial was not intended to rule out a risk margin of neuropsychiatric events. For this reason, all 95 percent confidence intervals in this presentation are to be considered descriptive and will be presented at their nominal level without multiplicity corrections.

All analyses in this presentation are based on the population of all treated subjects evaluated from the time that they received their first dose of randomized treatment in the trial to the time that they received their last dose, plus a window of 30 days.

Throughout this presentation, I will refer to the primary neuropsychiatric composite endpoint as the NPS endpoint. As you heard earlier, the primary statistical model estimated the risk difference of NPS events for every pairwise treatment comparison in each of the two cohorts in the trial.

Now I will discuss the trial results. I will first present plots of the cumulative event rates of NPS events through time for each of the two cohorts separately. The first plot shows the cumulative NPS event rate for subjects without a history of psychiatric illness at baseline. The 4 colored lines represent each of the 4 treatments.

In this cohort, within the first 7 days after randomization, there was no evidence of a

difference in risk associated with any treatment. There were 3 subjects with an event observed on placebo, 5 each on varenicline and bupropion, and 6 on nicotine replacement therapy within the first week.

Now, let's move to the end of the ascertainment window, which was defined as the end of treatment plus 30 days. Varenicline, which is represented by the purple dashed line, had the fewest observed subjects with an NPS event, with 13 corresponding to a cumulative event rate of 1.3 percent. The other treatment arms, bupropion, NRT, and placebo, were similar to each other in this cohort, with between 22 and 25 subjects with an NPS event each corresponding to cumulative event rates between 2.2 percent and 2.5 percent.

Now, here's a plot for the cumulative NPS event rate in the cohort with a history of psychiatric illness at baseline. The cumulative event rate of events in this cohort was about twice as high as in the cohort without psychiatric history. Within the first 7 days after

randomization, we observed some numerical differences in subjects with a neuropsychiatric event between treatments. In particular, there were 21 subjects randomized to bupropion who experienced an event within the first week after randomization followed by 12 for varenicline and 4 subjects each on NRT and placebo.

At the end of the ascertainment window, we can see that varenicline and bupropion were similar to each other and had cumulative event rates of 6.5 percent on varenicline and 6.7 percent on bupropion. And we can also see that NRT and placebo were similar to each other with numerically lower observed rates of NPS events of 5.2 percent and 4.9 percent, respectively.

This plot shows the estimated risk difference and the corresponding nominal 95 percent confidence interval for each pairwise treatment comparison in each of the two cohorts. The upper panel corresponds to the cohort without psychiatric history at baseline, and the lower panel corresponds to the cohort with psychiatric history

at baseline.

For each pairwise comparison, the treatment arms are labeled by the first letter of their names. V stands for varenicline, B for bupropion, P for placebo, and N for nicotine replacement therapy. And here I have highlighted several pairwise comparisons that I will focus on throughout my presentation.

You can see on the top panel that varenicline had fewer observed events than the other two treatment arms in this cohort, which is the cohort without psychiatric history, therefore, the estimated risk difference in this cohort favors varenicline. In the bottom panel, you can see that varenicline and bupropion were similar to each other and that both of them had more observed events than placebo in the cohort with psychiatric history. So in this cohort, the estimated risk difference between varenicline and placebo and bupropion and placebo show a positive estimated risk difference associated with these two treatments, with confidence intervals that include

zero.

Next, I want to spend some time discussing sensitivity analysis of the neuropsychiatric composite endpoint. As you just heard from Dr. Winchell's presentation, the trial had some issues regarding data collection, data coding, and reliability. During the rest of my presentation, I will discuss sensitivity analysis and secondary analysis, which we conducted to try to look more closely and address these issues.

Also, during our statistical review, we identified one additional issue, which is a possible statistical violation of the assumptions of the primary model regarding site heterogeneity, which I will discuss in the next few slides.

The trial randomized and treated subjects in 139 sites in 16 countries. As a standard part of our statistical review, we conducted a descriptive analysis to better understand the behavior of the primary neuropsychiatric endpoint across different sites. For this purpose, in the next couple slides, I will show you plots of the number of

events plotted against the number of subjects per each site in each of the two cohorts.

This first plot summarizes sites in the cohort without a history of psychiatric illness. Let me try to explain this plot briefly. Each dot in this plot represents one site. There was a total of 117 sites in this cohort. The horizontal axis represents the number of subjects in each site pooled across all treatment arms. The vertical axis represents the number of subjects within each site who experienced a primary neuropsychiatric event.

The pooled event rate in this cohort across all sites was 2.1 percent. So if all sites had the same true risk of events, we would expect for the dots to fall, on average, along the line with an event rate of 2.1 percent with some random variation. The blue shaded area represents -- I think it shows as blue. The blue shaded area represents where 95 percent of the sites are expected to be observed under the assumption of a common true risk of events of 2.1 percent, and the

additional now shaded area corresponds to the 99 percent prediction event.

Given a total of 117 sites in this cohort, we would expect approximately one or two outlier sites to fall outside of the shaded areas, only in this cohort, we observed 4 such sites. We observed slightly more outliers than we would expect, although the assumption of a common true risk of events across sites.

In the cohort with a history of psychiatric illness, we found stronger evidence of site heterogeneity. There was a total of 127 sites in this cohort, and the pooled event of neuropsychiatric events was 5.8 percent, so here, the shaded prediction bands have been adjusted accordingly.

Again, only the assumption of a common true risk of events of 5.8 percent across sites, we would again expect to see one or two outlier sites. But in this cohort, we observed 11, which are represented by the red dots. Also, we would expect to see approximately 45 sites without a single

event, but we observed 60. So there were more sites without a neuropsychiatric event than would be expected by chance under the assumption of common true risk of event. In particular, there were three large sites that did not record any NPS event.

We looked to see if the site heterogeneity could be explained by some covariates that were captured in the trial. For example, could there be a difference between sites in the United States against foreign countries? But what we found was that this site heterogeneity couldn't be explained by country of randomization, some cohorts of patients with psychiatric history, or by other covariates, including randomized treatment.

We also found high heterogeneity across sites in other known adverse events, such as irritability and abnormal dreams. So it remains unclear whether this site heterogeneity could have been caused by differences in patient populations, which is a possibility, or perhaps in differences in how data was collected and recorded in different

sites.

To address the issue of site heterogeneity, we conducted a sensitivity analysis of the neuropsychiatric event using a negative binomial model for the number of subjects who experienced an event within each site. This model was found to fit the data significantly better, and the results are shown here.

The pairwise treatment comparisons in this model are interpreted as rate ratios. They're no longer risk differences. This plot shows rate ratios and corresponding 95 percent confidence intervals for the risk of the neuropsychiatric endpoint. And here, I have highlighted the same pairwise treatment comparisons that I highlighted earlier in this presentation, and what we find is that the results are generally consistent with the primary analysis.

So accounting for additional site heterogeneity, we saw some wider confidence intervals. In the top panel, varenicline had fewer observed events than the other two treatment arms

in the cohort without psychiatric history. Therefore, the estimated rate ratios associated with varenicline in this cohort was less than 1. In the bottom panel, varenicline and bupropion were similar to each other with estimated rate ratios greater than 1 relative to placebo.

Now, I would like to discuss additional analysis of safety endpoints from this trial. In every safety trial, death is always an endpoint of interest. In this trial, there were 9 total deaths across all treatment arms, and we found no evidence of increased risk associated with any treatment. The treatment arm with the highest observed number of deaths was placebo, with 4 total deaths, which included the only completed suicide recorded in this trial, as you heard earlier.

The trial collected planned neuropsychiatric instruments at each in-person visit, and in this presentation, I will briefly discuss the results of the Columbia-Suicide Severity Rating Scale or C-SSRS. This instrument is of a special interest because it tries to measure a very serious event,

but also because the C-SSRS was discussed during the 2014 advisory committee meeting for varenicline.

This table summarizes three components of the C-SSRS by cohort and treatment arm. The three components are suicidal behavior, suicidal ideation, and self-injurious behavior. In the top table, we see that there were fewer instances of suicidal behavior and self-injurious behavior in the cohort without psychiatric history. Suicidal ideation was recorded in fewer than 1 percent of the patients in this cohort, and the risk was comparable across treatment arms.

The bottom table shows that there were also fewer instances of suicidal behavior and self-injurious behavior in the cohort with psychiatric history. Suicidal ideation was recorded in 2.1 percent of the total patients in the cohort with psychiatric history. Varenicline observed 27 subjects, followed by placebo with 25, NRT with 20, and bupropion with 15. So based on these tables alone, we found no evidence of a

difference in the risk of suicidal behavior and ideation associated with any of these treatments.

We also conducted exploratory analysis of other neuropsychiatric endpoints to evaluate whether they were consistent with the primary NPS endpoint discussed earlier, and I will briefly discuss three such endpoints.

The first one will be NPS events that were characterized as severe only. The second will be an NPS-plus composite, which we define as the primary NPS endpoint plus moderate or severe irritability, plus moderate or severe depressed mood disorders. Third will be our corrected NPS event that fixes a mistake, that Dr. Winchell mentioned, where dysphoria was categorized as aggression instead of depression.

Here's a summary of severe-only NPS events by cohort and treatment arm. As a reminder, severe events here are defined as adverse events that interfere significantly with the subject's usual function. Severe NPS events were observed in fewer than 0.5 percent of the patients in the cohort

without psychiatric history, and in 1.4 percent of the patients in the cohort with psychiatric history. Based on this table, there was no observed difference in the risk of severe NPS events between treatment arms in either cohort.

This plot shows the frequency of the estimated NPS endpoint represented by blue circles, compared to the frequency of the NPS-plus endpoint represented by the gray triangles. As a reminder, the NPS-plus endpoint was defined as NPS endpoint plus moderate or severe irritability, plus moderate or severe depressed mood disorders. And what this plot shows is that the NPS-plus endpoint was approximately twice as frequent as the primary NPS endpoint.

Here are the corresponding pairwise risk differences estimated for the NPS-plus endpoint. Again, I have highlighted the same pairwise treatment comparisons that I highlighted earlier, and the results are generally consistent with the previous analysis. In the cohort without a history of psychiatric illness, there were fewer NPS-plus

endpoints observed on varenicline than on the other treatment arms. In the cohort with a history of psychiatric illness, there were more NPS-plus events observed on varenicline than on placebo.

As I mentioned earlier, we also conducted analysis of a corrected NPS event that fixed the misclassification of dysphoria, but what we found is that there were relatively few events of dysphoria in the trial, and therefore the analyses of this event were consistent with the primary analysis, so I will not discuss them further.

Finally, I will conclude this presentation with some overall statistical comments and a brief summary. The review team identified some limitations in the trial, which are listed here. The clinical team identified inconsistencies regarding how the NPS endpoint was reported, collected, and coded. And in this presentation, we showed that the study sites exhibited large heterogeneity in the rate of NPS events under the assumption of a common rate of events.

Our analysis found that the large site

heterogeneity could not be fully explained by covariates captured in the trial. However, sensitivity analysis that allowed for additional site heterogeneity were found to be generally consistent with the primary analysis. We also found that the analysis of different definitions and additional safety endpoints were generally consistent with the primary analysis.

In summary, we found that in the cohort without a history of psychiatric illness at baseline, the trial observed a lower incidence of NPS events among patients on varenicline. In this cohort, the trial observed a low and balanced incidence of severe NPS events and also suicidal ideation and behavior captured in the C-SSRS. In the cohort with a history of psychiatric illness, the trial observed a numerically higher incidence of NPS events on varenicline and bupropion than on placebo, and the incidence of severe NPS events and C-SSRS events was similar in all treatment arms.

Thank you for your attention. That is the end of my presentation, and I will -- Dr. Winchell.

FDA Presentation - Celia Winchell

DR. WINCHELL: Thanks.

So to continue from Dr. Andraca-Carrera's conclusions, across sensitivity analyses, the conclusions about the finding in the study are consistent. In the non-psychiatric cohort, serious or clinically significant neuropsychiatric events, meaning events that had impact on patient functioning, occurred in all treatment groups, but the incidence was similar across treatment arms. In the psychiatric cohort, serious or clinically significant neuropsychiatric adverse events occurred in all treatment groups and were consistently somewhat more frequent in the varenicline and bupropion treatment arms.

The vast majority of events, although having impact on patient functioning, were not of a serious nature, and were usually transient. Serious adverse events in a psychiatric cohort primarily involved psychiatric decompensation, which is an established risk associated with antidepressants such as bupropion in patients with

bipolar disorder.

All three treatment drugs were effective aids to smoking cessation, and the prospective health benefit from abstinence from smoking is substantial. The balance of benefit and risk of smoking cessation products appears to differ based on history of psychiatric illness, but is favorable for both populations.

Next, we'll hear the review of the observational studies.

FDA Presentation – Chih-Ying Pratt

DR. PRATT: Good morning. I'm Natasha Pratt. I'm an acting team leader at the Division of Epidemiology under CDER. About two years ago, I presented a review of observational studies on varenicline's neuropsychiatric risk at our last meeting to discuss Pfizer's request of removing the boxed warning on varenicline. At that time, DEPI's conclusion was all of the available observational studies had limitations that preclude a conclusion of no association of varenicline with neuropsychiatric risk.

We also determined it would be challenging to evaluate such risks using observational data due to the difficulty in capturing all relevant outcomes and correctly classifying varenicline related neuropsychiatric adverse events and the difficulty in avoiding the selection of healthier varenicline users than their comparator. We suggested that the ongoing trial likely offers better insights into this issue. The committee agreed with us and recommended to reassess this issue after trial data was available.

Because our previous review was two years old, and the scope of today's discussion expands to all smoking cessation products, not just varenicline, DEPI updated our literature review from the last AC. I'll first describe how we identified the studies included in our current review.

We conducted a search of the PubMed database and identified 412 English language articles mentioning neuropsychiatric adverse outcomes and the three FDA-approved prescription smoking

cessation products: varenicline, bupropion, and NRT.

Studies were selected for in-depth review if they reported a relative risk of neuropsychiatric events, used adequate design to differentiate a temporary relationship between drug exposure and outcomes, and attempted to account for baseline group differences due to the observational design.

Among a total of eight articles that were eligible for in-depth review, we further excluded two studies because they either used the exact same data or similar data sources as another better designed studies that are also included in the in-depth review. Therefore, the focus of my presentation today is on six observational studies.

I'd like to point out that, first, two of the review articles described studies with FDA involvement, and two members of the DEPI review team are listed as authors on Meyer publication. Because of FDA's participation, DEPI was able to review the protocol and final reports of those studies, which contain more information than the

publication.

Secondly, as shown on this slide, my presentation covered the same studies that were addressed in Dr. Prochaska's presentation earlier today. I want to clarify, only two of the six studies were new studies that were not discussed in the last AC because the Cunningham publication, although it's published after the last AC, it describes the VA study that was already covered in my last presentation.

Next, I will provide an overview of the reviewed observational studies and their findings. The six studies included five retrospective cohort studies and one self-controlled study. We like it because they use real-world data, and they include patients with psych history, which enhanced the generalizability of their findings beyond most clinical trials.

The reviewed studies focused mainly on two types of outcomes. First, neuropsychiatric medical encounters, including hospitalizations, emergency department visits, and outpatient visits; second,

suicide related outcomes such as fatal or non-fatal self-harm identified by mortality data or medical encounter data.

The following slide summarized the main finding of the review studies. It's difficult to see the detail, but our intention is to show the overall trend. I'll start from the orientation of the plot. The findings of the five studies that examine neuropsychiatric medical encounters are presented at the top of the plot against the white background. The dashed lines separate the risk estimates observed from each study. The finding of the three studies that examines suicide related outcomes are at the bottom of the plot against the gray background.

Note that fewer studies estimated bupropion's risk than varenicline's risk. To be specific, only three reported bupropion's risk, and the risk estimates are represented as an open diamond in the plot. Also, most of the studies used NRT as a reference group, except that one study compared varenicline to bupropion, and the

other compared the varenicline exposed period to the unexposed period.

As illustrated in the slides, the reported findings varied considerably. Some reported a positive association between varenicline use and neuropsychiatric adverse events. Others suggested varenicline and bupropion are associated with a lower risk than NRT. But most of the findings did not show a difference in the outcome risk between varenicline versus NRT, varenicline versus bupropion, varenicline exposed versus unexposed time, and bupropion versus NRT.

The hazard ratio is bouncing around 1, and the confidence intervals cross 1. As we heard earlier, the sponsor's interpretation was that the observational studies did not show varenicline has an increased risk. We don't really agree or share the same view because of several study design issues, and I will address them in detail in the following sections.

Our first concern, all studies relied on diagnostic codes to identify neuropsychiatric

events or suicide attempts from medical encountered data, but no chart review was done to confirm those events indeed happened. We have concerns that diagnostic codes might not have well captured the full range of neuropsychiatric events that patients experienced while taking varenicline or bupropion. We also are concerned that medical records may not be the only data source to look for such events because patients experiencing those events might be referred to the legal system rather than the medical system.

We concluded outcome measures likely under-ascertained, and we are uncertain about how many events were missed. We also determined the outcome measures likely misclassified outcomes, and we are not sure if the event observed in those studies fully represents the range of adverse events experienced by the patient while taking smoking cessation products.

The second limitation that we identified, some review studies included data from the time frame after the publicity of varenicline's

neuropsychiatric risk. Because bupropion has also been associated with neuropsychiatric adverse events, we are concerned about differential prescribing or use of smoking cessation products based on a physician or patient's perceived underlying risk of neuropsychiatric outcomes. Specifically, we worried such differential prescribing or use would result in patients with a higher risk of adverse neuropsychiatric outcome being less likely to receive varenicline or bupropion.

Among the review study, we are most concerned about the study by Thomas and Kotz, both of which use UK general practice data, and included data after the UK regulatory agency issued a safety update on varenicline's suicide potential risk. In both studies, the varenicline user and bupropion user were very similar and appeared to have lower baseline neuropsychiatric risk than the comparator NRT user, in that they were less likely to have a history of psychiatric illness and had a lower frequency of previous psychotropic medication use.

Although both studies have tried to account for the baseline differences, we concluded that the trend of a lower neuropsychiatric risk associated with varenicline or bupropion, that were observed in both studies, still carried the bias due to the fact that the varenicline and bupropion user had a lower outcome risk to start from.

In the Molero study we are concerned about the confounding due to nicotine withdrawal symptoms because the study compared outcome risk between varenicline exposed time to unexposed time. The nicotine withdrawal symptom would occur at the same time when patient is exposed to varenicline, but it would not occur if the patient did not try to quit smoking during the unexposed period.

In that case, nicotine withdrawal symptoms would make varenicline exposed time appear to elevate neuropsychiatric risk even if varenicline is in fact risk neutral. It is unclear to us whether the increased neuropsychiatric risk that was observed in the study was due to varenicline use, the choice of comparator, or both.

In a study by Pasternak that compared outcome risk between varenicline users and bupropion users, as I mentioned before, because bupropion has also been associated with neuropsychiatric adverse events, we concluded it would be problematic to interpret the study results. The study found non-significant lower risk associated with varenicline use. However, this finding did not provide reassurance of varenicline's neuropsychiatric safety because the comparator, bupropion, also has been associated with neuropsychiatric adverse events.

Lastly, all the review studies included patients with psychiatric history, but the more relevant question was whether the risk would be different between users with and without psychiatric history. Among the review studies, the impact of psych history was either not examined, as in the Thomas and Kotz study, or cannot be appropriately assessed.

This slide shows the subgroup finding of the Molero study that compared the outcome risk between

varenicline exposed time to unexposed time, similar to our concern on the overall population finding, the main analysis finding. It is unclear whether the observed increased risk was due to varenicline use or the confounding by nicotine withdrawal symptoms.

In the other three studies that conducted a stratified analysis by psychiatric history, they were not able to provide a conclusive finding due to a small sample size or few observed events in the subgroups. But I'd like to point out that consistent with the trend in the PMR trial findings, results of the three studies all indicate that varenicline users with psychiatric history might have a higher neuropsychiatric risk than those without because, first, the majority of the neuropsychiatric events were observed among patients with psychiatric history.

Also, the hazard ratio of neuropsychiatric outcomes were numerically higher among patients with psych history than the overall cohort or patients without psychiatric history. But as

depicted in this slide, those studies were underpowered to confirm the effect modification by psych history.

To sum up our assessment, all studies had a number of study design issues, including outcome misclassification and under-ascertainment, differential prescribing or use due to the perceived baseline psychiatric risk, and confounding by nicotine withdrawal symptoms.

When the potential bias is considered in combination, they restrict our ability to predict the direction of the risk associated with any of the smoking cessation products, besides one study's use of bupropion as reference group to examine varenicline's neuropsychiatric risk was problematic because finding no increased risk did not reassure varenicline's safety given that both products were labeled for neuropsychiatric adverse events. Finally, the inability to assess the risk among those with psychiatric history further restrict the generalizability of the observational study findings.

Because of the limitation, the evidence from the existing observational studies alone is of insufficient quality to confirm or refute an increased neuropsychiatric risk associated with either varenicline or bupropion use. The neuropsychiatric safety of smoking cessation products should be assessed based on the totality of evidence, including to provide a determination of whether or not patients with psychiatric history are at an increased risk for neuropsychiatric adverse events.

This concludes my presentation. Thank you for your attention.

Clarifying Questions to FDA

DR. PARKER: Thank you. So let's turn now first to clarifying questions for the FDA. If you will place your card up again, we'll get your name on the list here. I'll ask that you state your name for the record before you speak, and that you keep the questions brief and specific to the FDA initially. Hopefully, pending time, we will then go back and pick up -- I know we still have seven

people from before that had some questions specifically for the sponsor.

So let's start our list with those who have questions for the FDA. Dr. Narendran?

DR. NARENDRAN: I just have a quick question for the FDA statistical reviewer. It seems like there are 10 to 20 percent of the patients who had already had been on varenicline or bupropion who are entered into the study. You would think that the people who already are willing to go into a study did not have adverse events before. If you exclude them from your analysis, does that change the risk profile or the NPS endpoint?

DR. ANDRACA-CARRERA: This is Eugenio Andraca. Related for efficacy, I don't believe that we have that analysis for safety. I do not know if the sponsor has that. Maybe they can speak to it.

DR. PARKER: Dr. Fiedorowicz?

DR. FIEDOROWICZ: Thanks. My name's Jess Fiedorowicz from the University of Iowa. My question's for Dr. Andraca-Carrera. Slide 29,

which presents a summary of the findings, states that there was, quote, "a higher incidence of NPS events observed on varenicline and bupropion than on placebo." You qualified that statement with a phrase, quote, "numerically," unquote, and I was just wondering how confident are you that these findings are not due to chance.

DR. ANDRACA-CARRERA: This is Eugenio Andraca. The study was not designed to rule out a specific margin. It was designed to be descriptive. We have the estimated parameters with confidence intervals. I think it's up to you and the clinical team to interpret those confidence intervals.

DR. PARKER: Dr. Winterstein?

DR. WINTERSTEIN: This is a question I think for Dr. Winchell, but I imagine that several colleagues from the FDA could chime in, as well as the sponsor. I'm struggling with the endpoint massively. I appreciate the effort that was put into trying to create an endpoint that would be more suitable to capture what had been observed in

the spontaneous reports.

Those of us who are trained to conduct safety studies or review safety studies are alarmed when they see composite endpoints because the big concern then is does that endpoint capture noise. And if we have noise in a safety study, we lose the ability to identify differences.

I'm trying to look at all of these events, those 280 or so various MedDRA terms that were included in this endpoint, and I'm trying to find out what's the noise here and what, and was this study massively underpowered to do anything. There are MedDRA terms that were quoted here that say things like "elevated mood," which I'm not sure that would be a safety endpoint that I would be particularly interested in, even if it were rated as severe, recognizing the fact that it has already been alluded to that the severity rating hadn't been standardized or validated previously, and seemed to be fairly implicit in the judgment as it had been applied throughout the study.

So we have an unvalidated, not particularly

reliable ascertainment system, a variety of events that I don't know what actually captured the drug effect that we are looking at. And I would like to get some input, number one, what was the thinking, and what drug effect would really be important.

Was there an idea to try to remove drug efficacy effects? Because I could see an elevated mood from the efficacy of not having to smoke any longer, which we clearly wouldn't want to have in a safety endpoint. And given that we have 8,000 patients exposed to a trial, what was the power analysis on all of this? I mean, I would imagine that there was some kind of underlying power calculation done that was focusing on some simple size estimate to rule out some increase in safety events, and what were they?

If I could get help with that part. I realize that advisory committee members tend to complain about the results after the fact, and I'm complaining about the results after the fact. I realize that. But I have trouble getting the essence out of this trial that would allow me to

say, yes, there's really no safety problem.

DR. PARKER: Does the agency want to respond? I know the sponsor does.

DR. ANDRACA-CARRERA: This is Eugenio Andraca. I believe that the sponsor presented some slides about power calculations, so maybe they would be the best to address that particular issue. And then if you would like to, we could come back to that, to discussing the endpoint from our perspective.

DR. RUSNAK: Thank you. I'd like to invite Dr. Gaffney to present that information.

DR. GAFFNEY: Thank you. Mike Gaffney, statistics, Pfizer. As Dr. Anthenelli pointed out, this study was not formally designed to address a specific hypothesis. There wasn't sufficient information to estimate a treatment effect or to estimate a non-inferiority margin in this trial. The real focus was on estimating what the rates are and confidence intervals around that rate. However, to address the question, we can in a post hoc way give what the power was in EAGLES

given the observed placebo event rate in a study of 8,000 patients.

Could you put up slide ST-179, please? Thank you. What you see here in the left-hand column are the actual observed placebo non-psychiatric primary endpoint rates. In the non-psychiatric cohort, it was 2.4 percent, in the psychiatric cohort, 4.9 percent, and overall about 3.7 percent.

The next columns give both the risk difference and the relative risk that a study of 8,000 patients, where there would be 1,000 patients per each treatment comparison in the non-psychiatric cohort, 1,000 per treatment group in the psychiatric cohort, and 2,000 patients overall.

So these numbers show that on the risk difference scale, with 80 percent, it would have picked up a difference of 2.32 percent in a non-psychiatric cohort, 3.1 percent in a psychiatric cohort, and overall about 1.9 percent. If you prefer those numbers on the relative scale,

it's just under 2 on the non-psychiatric cohort, a relative risk of 1.6 in the psychiatric cohort, and overall about a relative risk of 1.5.

DR. WINTERSTEIN: So the 8,000 patients was more or less a convenient number? I'm thinking -- there were 8,000 patients who were exposed to this trial to identify -- without an idea of what the incidence of those events would have looked like. Correct?

DR. GAFFNEY: Well, the incidences right here are what was observed, and we presented what was observed. These are what the risk differences -- if there is a true effect on any one of these active treatments versus placebo, of the order that you see here with respect to risk difference or relative risk, this study was sized with enough power, 80 percent power, to detect that.

DR. WINTERSTEIN: Yes, to detect the risk differences that we see here, assuming that all the adverse events that were collected in that composite outcome would actually be relevant. I'm

just surprised that there were no a priori ideas about how many patients were needed to rule out something.

DR. RUSNAK: I think to answer the question, a priori, some estimates were made of what the incidence of the NPS AE events would occur in the non-psychiatric cohort as well as the psychiatric cohort. But the certainty around that wasn't entirely precise. So what the trial did was monitor the overall NPS event rate, and then they had the power to increase the sample size to ensure that we had the appropriate sample size for the study during the course of the trial. And this was done at 50 percent and 75 percent of enrollment.

DR. WINTERSTEIN: Fifty percent difference?

DR. RUSNAK: No. Whenever 50 percent of the subjects were --

DR. WINTERSTEIN: Oh, the interim analysis.

DR. RUSNAK: Interim analysis.

DR. WINTERSTEIN: So what was that a priori idea of a difference that you were trying to shoot for?

DR. RUSNAK: I'll ask Dr. Gaffney to provide the exact details of that.

DR. GAFFNEY: As been stated, there really was very little prior information which to make any estimates, or clinical trials had excluded patients with psychiatric diagnosis. So all we were left to do was to try and recreate what turned out to be the neuropsychiatric adverse event in this trial.

We looked at that over all of our clinical trials, maybe 18 or so. The estimate we got was about 1.75 percent. It was low. We doubled that in the expectation that in getting solicited events rather than just volunteered events, which came normally in clinical trials, that we would have this rate of about 3.5 percent. We doubled that again to estimate that possibly we'd see 7 percent in the psychiatric cohort, which had not been studied.

As you see from the observed events, we got about a 2.6 percent rate within the non-psych cohort, and we had something above 4 percent, 5 percent in the psych cohort. So a little bit

less than what we expected by our assumption of doubling, but certainly more than what we had seen originally in our database.

DR. PARKER: Dr. Hertz, did you have a comment? Oh, sorry.

DR. WINCHELL: I was going to respond to your concern about the incorporation of all of the vendor terms. The broad net for the MedDRA terms was intended to capture events that are sometimes described in terms that are difficult to code. And the intention was that by constructing narratives that told the whole story, what the patient had experienced in their own words and other people's words, and everything together, we'd be able to adjudicate those. And if there was a circumstance where someone had an experience that you wouldn't have considered a concerning clinically significant adverse event, we could exclude those based on review of narratives.

So I understand that not every single term in the list of MedDRA terms is necessarily that specific item would be something you'd be concerned

about, but we also didn't want to lose something by virtue of it having been, for whatever reason, assigned to a term that wasn't on a short list.

DR. PARKER: Dr. Emerson?

DR. EMERSON: Just one real quick follow-up on Dr. Gaffney's presentation. Given that this is really a safety trial and that we're sort of more interested in what we rule out, what was the 97 and a half percent power point for this study? Because that's what would correspond to the 95 percent confidence intervals that are being presented.

DR. RUSNAK: I'd like to ask Dr. Gaffney to respond to this question.

DR. GAFFNEY: If you could call up slide ST-180, please. I'm not sure if this addresses your question directly, Dr. Emerson, but what we also did was to use the observed rates that we saw to look at this study from the perspective of the non-inferiority margin. And I think that's getting at your 97.5 percent confidence interval.

Again, on this slide, you see the placebo rates that were observed in each of the cohorts and

overall, repeated again for EAGLES. And calculating the non-inferiority margin for the non-psychiatric was 2.4, the psychiatric cohort about 1.9, and overall for this study was about 1.6. So the study would have 80 percent power to rule out the 97.5 confidence interval, exceeding those values, which is the definition of the non-inferiority margin.

DR. PARKER: Dr. Conley?

DR. CONLEY: This is primarily to Dr. Winchell, but others can answer with you. Thanks for the presentation. The concern that I have from an industry-wide perspective is though I respect your need to dive into the data and figure out what's going on, sometimes a presentation really seems to lack context. You had mentioned early in your talk that you expect large, multicenter, international studies to have bumps in the road.

Now, at the end of the day, it seems that you're primarily agreeing with what the sponsor has said about both safety and efficacy; at least

that's what I got from the presentation. So if I missed that, please say so. But what I don't know is that we still have raised up -- and that's going to be a discussion issue later on -- quote/unquote, "Should we believe this study?" Are there enough problems here to do it.

I do worry that you all have a bias of never kind of being satisfied when you dig into case reports, and that's because you're looking at safety; I understand that. But there are always going to be some cases you can't ascertain. And what I can't understand, and what I can't put into context, was this some sort of an outlier that there are a lot of problems or not a lot of problems. I think that might be helpful given the questions you're asking.

DR. WINCHELL: I can say that I have reviewed a lot of different NDAs, and this particular one had more barriers to review than typical. The quality of the narratives that were submitted were unusually uninformative. And yes, I found that this was more difficult than typical, if

that answers your question.

DR. PARKER: Dr. Hernandez-Diaz?

DR. HERNANDEZ-DIAZ: Thank you. And actually, you can cross my name from the question to the sponsor because I'm going to ask the same question. What I was going to ask was if we could see the survivor curves, the Kaplan-Meiers that Dr. Andraca-Carrera showed in slide number 8 of his presentation.

If you can put it up; but meanwhile, I agree with the review of the observational studies. I think that some of the limitations that were listed, I would not consider them limitations in itself. I'm referring to the comparison with bupropion. So a comparison with an active treatment, I don't think that's a limitation. We use that in clinical trials all the time. It's just that it's answering a different question.

The reason for that limitation is that we could not assess whether the difference in the risk is to the indication or to the active treatment itself. And I think that in this clinical trial,

because everybody was quitting smoking or trying to, we are now I think left to analyze, if all of the studies are increasing the risk similarly, actually what would have happened to those subjects had there not been exposed to a smoking cessation intervention.

So I was trying to get from this data a sense of what could have been the risk in this population if we didn't try to have them quit smoking. You presented the cumulative risk in your summary, but looking at the graph, it seems to me that there is around 4 percent and 2.5 percent of cases in the first 30 days. Then in the next 30 days after the start to follow-up, there is around -- I was estimating the difference around 2 percent in the bupropion group and 1 percent in the placebo group. Then in the next 30 days, there is around 0.5 percent and 0.5 percent.

So the rate is decreasing over time. I'm wondering if with that data and with perhaps baseline data in similar populations, if we can have an idea of what is the risk of quitting

smoking. Are we talking about 5-, 10-fold increased risk of these events when you start study to quit smoking?

I think this is not going to help further respond to the question about whether one treatment is safer than another, but to inform patients that if you are really trying to quit smoking with a strategy that actually seems to work, this is the risk you are going to have during the first 30 days. And I think that's important for patients and for healthcare providers to keep in mind, to be watching for those initial increases in the risks for all the treatment actually.

I would expect that the better the treatment is -- this is really due to the fact they've withdrawn from smoking. The better the treatment, the more events I would be expecting. So I was wondering if you would agree with that interpretation, that there is an increase during the first 30, 60 days after starting a study.

DR. WINCHELL: I think we're not clear of what specific question you'd like us to address. I

will remind you that quit day was at day 7. So anything before quit day should not necessarily be associated with quitting smoking. So if your question was about the risk of quitting smoking, maybe that will help you.

DR. HERNANDEZ-DIAZ: Well, I was approaching as an intention to treat kind of approach. Like if you start at 7 days, you started the trial, and you see -- my point is that the rates -- you presented cumulative risk after the whole period, after end of treatment and plus 30 days, but the rate, the hazard, is not constant over time; split at the beginning, and then they are pretty flat. And if you look in the non-psychiatric cohort, there is apparently very few cases after 60 days.

So I'm just saying that the interpretation and the study of these groups might help with the recommendations at the end.

DR. ANDRACA-CARRERA: This is Eugenio Andraca. Unfortunately, we didn't compute confidence intervals for the curves, which is possible that it might show some overlap. And

also, the prespecified comparison was only prespecified at the end of treatment plus 30 days.

So I would say that this curve should be informative to give you an idea of what the actual observed pattern of time for these events were. But we didn't prespecify any comparisons at 30 or so days. So that could be sort of a post hoc comparison, and could lead to the wrong conclusions.

DR. PARKER: Dr. Budnitz?

CAPT BUDNITZ: Yes. Maybe we could put up slide 20 from the statistical presentation, if that's okay, because I'm struggling with how the identification of the primary NPS adverse event endpoint actually happened in the EAGLES study. Here we, I guess, have suicide, behavior ideation, and self-injurious behavior events. It's not clear if these were from self-reports or from these instruments, and then a follow-up.

So that's my first question either to FDA or sponsor. How do we distinguish where these events came from?

The first thing I'd like to address is in the briefing booklet, on page 48 -- and I think it was referred to in the presentation -- there are two patients who deliberately took an overdose of the medication. They were not coded as making suicide attempts. These cases were not even selected for preparation of narratives as being a potential interest.

So I'm trying to reconcile these two patients that took intentional overdoses of medication, do they appear in this suicide slide? And if not, then are there other types of adverse events? Or if not that, how do we even have that information?

DR. WINCHELL: It's my understanding -- unless sponsor can confirm -- that when the C-SSRS was administered, patients who endorsed suicidal ideation or reported behavior were then assessed for whether or not that endorsement represented an adverse event. And if the investigator felt that the suicidal ideation reported was not an adverse event, that that was

not included in the adverse event data set, and only adverse events were included in the NPS endpoint.

So I can tell you that the one patient who took an intentional overdose was not coded as a suicide attempt, and we were told that the C-SSRS, he never endorsed suicidality, although we don't really have an explanation of why he took the overdose.

CAPT BUDNITZ: So I'm trying to clarify. So that information about one or two patients that took intentional overdoses of drugs and were not reported as suicides, and were not reported as adverse events, then where did that information come from? I'm still confused.

DR. WINCHELL: So they could be reported as an adverse event. The overdose was reported as an adverse event. The overdose was reported as an adverse event but not a suicidal adverse event, in both cases. One was coded to the term "overdose" but not a suicidal overdose. And one was coded to the term "accidental overdose," although the

verbatim said that the patient took a handful of -- took all of her pills at once. So that didn't sound like an accident. As to how those were or were not handled, I can't say.

DR. PARKER: So maybe direct this directly to the sponsor. If you could answer about those two patients specifically.

DR. RUSNAK: Yes. I'll invite Dr. Cristina Russ.

DR. PARKER: Just those two to start with.

DR. RUSS: Cristina Russ, Pfizer. The accidental overdose with the patient that took 4 bottles of study pills, it was included. It's captured in the Columbia scale, as I mentioned during the presentation, and it was included in the sensitivity analysis as a result of the clinical review. Another overdose with psychotropic medication was coded -- was mapped directly. It's captured by the scale, and it's also included in the primary endpoint of placebo subject.

So those two are -- that's the situation of those two cases.

CAPT BUDNITZ: We'll go into this later, but I guess I'm confused about this is a study trying to identify if an event is associated with an exposure, but the investigator has this -- based on their prior experience can determine if an event, like a clear overdose, is or is not related to the study drug.

It seems like it it's inherently -- it doesn't make sense in the point to have an epidemiologic association because you are using your predetermined assumption about what is a study related event to then be the outcome of whether or not there's epidemiologic association between study related event.

DR. RUSNAK: I'd like to invite Dr. Anthenelli to address that question.

DR. ANTHENELLI: Robert Anthenelli, University of California, San Diego. So investigators were experienced and trained on reporting of adverse events. I'll give you an example, though, of how a positive response and the C-SSRS might not lead to an adverse event report.

And I can give it -- it's actually from case B of the clinical vignette. I know you don't remember my slide show, but that happened to be a 40-year-old woman with bipolar disorder.

This particular patient had chronic suicidal ideation, and in between her manic episode she was low-grade chronically depressed. And as a symptom of that, she was also chronically low-grade suicidal.

We recorded that on, of course, the baseline C-SSRS. So when she came in at week 1 for the evaluation, and she still was reporting that, because that was no exacerbation or change in her preexisting state, that did not get reported as an adverse event at week 1. However, when she came back two weeks late and she had this more major mood change, then that change in the severity or the intensity of the C-SSRS, which was captured in that time, and of course on that neuropsychiatric adverse event, and of course on the HADS, was all captured into that adverse event report.

So there can be some discrepancy in the

C-SSRS finding and an adverse event report.

DR. PARKER: Dr. Rimal?

DR. RIMAL: Rajiv Rimal from George Washington. I have a question for the FDA with regard to something in the briefing document. It mentions wide variations across sites on a variety of measures, including in financial disclosures. And I'm wondering how that variation across sites was taken into account, either in the subanalyses or in the primary analyses.

DR. ANDRACA-CARRERA: This is Eugenio Andraca. The variation across sites in the primary analysis, what we did is we looked for different statistical models that fit the data better. The negative binomial model that I presented was found to fit the data significantly better than the primary model. So we fit that model to account for the additional [indiscernible] heterogeneity, and we presented the results.

In terms of sites that had -- for example, the two sites that were identified previously by the sponsor, we excluded sites that had other

potential problems that were identified either prior to the submission or during the conduct of the study. If we didn't find that the results were significantly different, we did not discuss them. We included some in the background package. We didn't find any major discrepancies in the results.

DR. RIMAL: I guess my follow-up question to that is that my experience is that if there is a problem in the site on one event, it's quite likely there's a whole series of problems in that site. So it may be more cumulative than you're making it out to appear.

DR. ANDRACA-CARRERA: I can only talk in terms of the events that were captured. There is a correlation between the sites that had few events for the primary event, also had few captured abnormal dreams. They had few captured -- it was on irritability I believe. So it's not a perfect match, but there is a correlation that sites that captured few primary events captured few other behavior or psychiatric events.

DR. PARKER: I had a follow-up to that just

regarding the FDA's look at this. I understand that there were about 150 sites -- 139 I think, 16 countries. I wanted to know how many languages the instruments were presented to enrollees in. I know that there were a total of 8,000, but it looks like in the U.S., there are about 4200, 4260 that you presented in your background documents. And I saw that there were over 800 from Bulgaria, the Russian Federation, Slovakia.

Can you give us some idea about what we know about the instruments that were used to garner the data from the neuropsychiatric events, to surveys throughout -- were these instruments that were known to be validated and have good testing characteristics in other languages, or was this the first time they'd been used to capture information on enrollees, almost half the study, or over a third of the study that weren't primary English speakers?

DR. ANDRACA-CARRERA: This is Eugenio Andraca. I can list the countries under the number of events. Perhaps the sponsor might have a better

response about the instruments and how the instrument was collected in different countries.

DR. PARKER: Are you aware of how many languages total?

DR. ANDRACA-CARRERA: If we look on my slide, backup slide 19, statistics backup slide 19, these are all the countries in the trial.

DR. PARKER: No information on the number of languages, how language was --

DR. ANDRACA-CARRERA: I do not. I do not know if some countries had multiple languages or not.

DR. PARKER: Does the sponsor have the answer to that, how many different languages the instruments were available in and used in?

DR. RUSNAK: The instrument was used in two studies prior to EAGLES. One was a major depression study, and the other was a study that was specifically conducted amongst the patient population -- that matched the patient population in EAGLES. We don't have the specific language information now, but we could try to get that

information to you over the break.

DR. PARKER: So that would be the total number of different languages and any testing characteristics about the instruments and data capture using those instruments in other languages.

Dr. Hennesey?

DR. HENNESEY: Thank you. I think I'm addressing this question to either Dr. Winchell or anybody else at FDA who'd like to address it. So my understanding from the statistical review is that serious neuropsychiatric events occurred equally across groups, both in those with baseline psychiatric mental health conditions, and those without.

Dr. Winchell's slide 29 concludes that the use of varenicline is favorable both in patients with and without mental health conditions. If it's true that a boxed warning dissuades people from using a drug -- and we heard at least anecdotal evidence of that today -- if all those things are true, then isn't there a negative -- so isn't the benefit-harm balance of a boxed warning negative in

this context?

DR. HERTZ: Hi. This is Dr. Hertz. That's not a clarifying question. That's a really good discussion question. So I'd like to refer that to a little bit later so we can just keep going with the clarifications.

DR. HENNESEY: Fair enough.

DR. PARKER: That would be called hold that thought.

Dr. Pickar?

DR. PICKAR: Dave Pickar here. I wanted to ask Dr. Eden, just in general, as a psychiatrist who treats seriously mentally ill patients, that group of people are a terrible risk for the hazards of smoking; there's no question. And you started -- and we talked about it that way. We don't have a large number of schizophrenic patients in this package, but we have some. There's no question that adverse events are enhanced in people with a psychiatric illness.

How many people are hospitalized? How many schizophrenics who you gave this drug -- you're

giving a drug that affects the brain to help with withdrawal and to encourage abstinence from smoking. How many were hospitalized? Of course, you have to remove Bulgaria. I'm not familiar with the hospitalization, but I am familiar in the United States.

How many people -- patients, depressed patients, schizophrenic patients, non-history -- how many were in a hospital? The silliness of these reports and the discussion of them is just a little bit much. And you're talking about people dying from cigarettes and so forth. So I'm not pushing it for industry, but I think we got lost somewhere here.

(Applause.)

DR. PICKAR: I mean, really, this has gotten -- but I'd like an answer to a very specific question. Okay? And it's important to me, because when patients with schizophrenia get in trouble, they end up at a hospital. If a depressed patient has a serious relapse, they end up in a hospital or serious care. You're talking black box warning.

We're not talking how was your day today, ma'am?

So what's the answer to that? Can you help me with that? Just in the United States, how many people ended up in a hospital in association with this trial?

DR. EVINS: Can you show MD-76?

Dr. Pickar, I don't believe we know how many in the United States, but we can try to get that for you over the break because that's an excellent question. We do have the number in the psychiatric cohort who have had serious -- severe events. So the number is very low, so 14, 14, 14 and 13. That's in the entire psychiatric cohort, so amongst 4,000, roughly, patients in the study.

If you show the slide from my deck, MD-106 -- and again, we can try to get specific hospitalization numbers for you. You were interested in the -- you mentioned psychiatric disorders. So this was not in the Lancet paper. This is a subanalysis that we looked forward to doing that breaks out by diagnosis those with the most serious illness, psychotic disorders. The

rate of the primary endpoint is quite low.

I don't believe we have the -- okay. So it looks like we have slide S-443, please. So this is hospitalization in the entire psychiatric cohort, so it's not by psychotic disorders, but in the entire psychiatric cohort, we've got 4 people on varenicline, 4 on bupropion, 4 on NRT, and 1 on placebo who were hospitalized. And you can see the neuropsychiatric adverse event that's listed for those, and I'd be happy to try to break those out by subcohort for after the break and discuss it further.

DR. PICKAR: Okay. I certainly appreciate that, and that gives me some picture. There's not particularly a difference among treatment per se.

Question on the schizophrenic persons. Did people have to change medication? You have no data on medication and what somebody was treated with considering they're depressed patients or whatever. I mean, if you're going to do this and understand it -- these are tough questions. This was a colossal, well-done trial. I mean, what a tough

trial. But if you really want to tease it apart, how did it interact with specific medications? Is there any data on that?

DR. EVINS: So I can speak to the fact that investigators were allowed to adjust medications for patient stability. They were treated with the clinical best treatment that had to be stable at the beginning of the trial.

DR. PICKAR: So if someone was experiencing a symptom, their physician, their treating physician could change the medication to attack that.

DR. EVINS: That's right. And if you show MC-29, we can give Dr. Pickar some numbers for that. Those are the numbers of patients in the psychiatric cohort on the bottom -- again, the entire psychiatric cohort -- who required a medication change due to a neuropsychiatric adverse event. And you can see it's about 30 people per arm, anywhere from 25 to 36, who needed a change in medication following a neuropsychiatric adverse event. So again, on the base of a thousand

patients, this is quite low.

DR. PICKAR: I don't mean to be too picky on it, but we're going to have to decide whether it's a black box or not -- that really is the conversation today -- and what constitutes that. So I just had to get a clear picture of that, and we'll discuss it more. But thank you.

DR. EVINS: To me, this is the kind of rate you would see as a base rate --

DR. PICKAR: Yes.

DR. EVINS: -- over the course of 12 to 16 weeks.

DR. PICKAR: If you're treating a significant number of seriously mentally ill patients, you're going to see versions of this all the time. And stable is one thing, but stable doesn't mean the exact same dose every day or every week.

DR. EVINS: Right. And this was oversampled for the more seriously ill patients because while half of the neuropsychiatric cohort were on a psychiatric medication, 95 percent of those with

schizophrenia spectrum disorder were on a medication; 75 percent of those with bipolar disorder were on a medication. So this would oversample for those with more serious illness.

DR. PICKAR: Thank you very much.

DR. PARKER: Dr. Roumie?

DR. ROUMIE: Thanks. Christianne Roumie. So I think one of the comments that have been brought up a number of times is the question of underreporting of events, and Dr. Andraca-Carrera brought out by site the number of sites that was higher than expected that reported zero events. And I was wondering if you have done any sensitivity analysis.

In the psychiatric cohort, it looked like you didn't need but a few more events to tip your confidence interval into exclusion of zero -- I'm sorry, exclusion of 1. So whether or not you did some bootstrapping samples and looked to see whether or not that underreporting -- how many more events would have been needed to tip the findings to positive.

DR. ANDRACA-CARRERA: This is Eugenio Andraca. We did not conduct that analysis.

DR. PARKER: Dr. Morrato?

DR. MORRATO: I had the exact same question as Dr. Roumie. So another way of saying it is how bad would the underreporting had to have been in order for it not to become significant? And that's commonly done in these kinds of studies.

What is the p-value in that? I think it was slide 10, just so that we have an anchoring of the p-values for the V versus P and the B versus P in the psychiatric cohorts.

DR. ANDRACA-CARRERA: This is Eugenio Andraca. We purposely didn't percent p-values because p-values are usually associated with a prespecified hypothesis.

DR. MORRATO: Okay.

DR. ANDRACA-CARRERA: So we think that the trial was designed to be descriptive, and perhaps that's the best way to interpret it, based on the point estimates and confidence intervals.

DR. MORRATO: Then along that line, I know

you used the other negative binomial modeling. So as we consider the data and what might get reported in labeling -- I assume that's going to be one of the questions -- do you feel confident that the primary analysis that's presented is the one that we should be considering?

DR. ANDRACA-CARRERA: This is Eugenio Andraca. We haven't discussed which analysis would be more informative. I can say that from a statistical perspective, the negative binomial model fit the data better, significantly better, than the binomial model.

DR. MORRATO: Okay. Thank you.

DR. PARKER: Dr. Pickar, I think you had a --

DR. PICKAR: There certainly is a hypothesis. Excuse me.

DR. ANDRACA-CARRERA: I'm sorry?

DR. PICKAR: There's a hypothesis here. Am I lost here? The hypothesis of this drug, or this group of drugs or this drug in particular, causes significant adverse events that cause a black box.

I mean, if there's all patients -- some of us who were here on the previous board meeting, who passed it on now, wanted to see this overall data, and there's no question there was a hypothesis.

If I was back functioning as a scientist, the hypothesis -- there have been plenty of studies that were giving the hypothesis that it's going to worsen somebody. And here you have reason to believe it. That is a hypothesis, and I would like to see the p-values.

DR. PARKER: I'm going to take that as a comment.

(Laughter.)

DR. PARKER: Dr. Morgan?

DR. MORGAN: Anybody that wishes can respond to this, but I think it might be in the bailiwick of Dr. Evins or Dr. Prochaska. We heard about some of the limitations from FDA, limitations of the observational studies in that we're not seeing the true frequency of neuropsychiatric events because the medications aren't being described because of the black box warning. And also, I think it's

clear that this is a real health disparity if we have folks with psychiatric disease that aren't getting treatment that can help them quit smoking and save their lives.

Do we have the data from surveys or other studies, or is there speculation, regarding the reluctance to prescribe, or patterns of prescription that have been changed by the black box warning amongst psychiatrists or primary care physicians who treat a lot of people with affect disorders and psychiatric abuse, psychiatric disorders? Thanks.

DR. HERTZ: Just for clarification -- this is Dr. Hertz -- are you asking if there's information about the impact of boxed warnings on medication use in general or specifically here?

DR. MORGAN: Here.

DR. PROCHASKA: I was hearing that as a two-part question, so I'll answer the first with observational, and then --

DR. PARKER: We're going to do this really quickly and very pointedly. Okay?

DR. PROCHASKA: So with the observational data, a couple of those studies were population level, was the entire country of Sweden, the entire country of Denmark. So they do have individuals with mental health concerns. And then the VA data as well had individuals with mental health concerns.

There were differentials, as you saw, at baseline, and that's why they did the propensity score analysis to map -- to measure compounds to have them be equal so that you can get a picture of what's going on in smokers with mental illness. Certainly, there are limitations in the different observational studies, and that's why it's so important to look at the map of them. So they're not just looking at one individually, but each is answering different questions in different ways, and all of them are an enhancement over the case-reporting data that we have.

DR. PARKER: Let's go back. We had several people earlier that had questions --

DR. MORGAN: If Dr. Evins wants to respond

to the second part of the question.

DR. PARKER: Okay.

DR. EVINS: I'll be very quick. Eden Evins from Mass General Hospital. There are reports that people with psychotic disorders particularly are underprescribed varenicline, which is published in the literature. There are surveys that underpin to RO1s to NIMH, and a large pragmatic trial to PCORI, indicating convinced reviewers that psychiatrists underprescribe both bupropion and varenicline to people with serious mental illness, and that this underlies the largest mortality disparity in this country.

DR. PARKER: Okay. Let's go back quickly if we can and try to pick up a few of the folks who had specific questions for clarification to the sponsor. We had a long list before we took the break. So I'll call on these folks, and if you've already had your question answered, that's fine. Otherwise, let's see if we can get these in.

Dr. Winterstein, you had a question.

DR. WINTERSTEIN: That actually got

answered, but I will take the slot real quick. The hospitalization data that was shown, it looked like from the table that this was within the patients who had a neuropsychiatric event reported. It is not hospitalization rate across everyone, correct?

DR. RUSNAK: That's correct.

DR. EVINS: I showed the hospitalization with psychiatric cohort [inaudible - off mic].

DR. PARKER: Sorry. If you don't speak in a microphone, we don't get it. I think the question is if you could maybe get that data for us of all hospitalizations and specified by which cohort. That would be helpful, and maybe you can share that with us when we come back.

Does that answer that? Great.

Dr. Budnitz?

CAPT BUDNITZ: Yes. This is in reference to slide MD-56. This is, again, just trying to get a handle on the NPS adverse event ascertainment. I think it was mentioned that about half of the events were volunteered adverse event reports. I think that was mentioned.

Could you give us, for the second half, just how many were from each of these different methods? I think that will be helpful to see if this solicited reporting, what kind of events those --

DR. PARKER: So for each of these four, to list the end for each of the four that are on the slide.

DR. RUSNAK: Right now, to be specific, the volunteered actually accounted for 46 percent; the solicited, which was the NAEI, was 54 percent. And then amongst the volunteered was also some of the proxy reporting. While this was a novel aspect of EAGLES, it actually represented less than 10 percent of the overall AE reports that came.

CAPT BUDNITZ: And just to follow up, according to this slide, it doesn't say there's any, quote, "deeming" by the investigator for volunteered adverse event reports. Is that correct? And there was this deeming to the adverse events from the other --

DR. RUSNAK: So the investigator was the final arbiter of what gets reported as an adverse

event report. Dr. Anthenelli had already described to you an earlier case by which the patient had some baseline in levels of depression. And it's really an increase of frequency that triggers the difference, not the presence of the symptom itself.

CAPT BUDNITZ: Okay. So just to correct the slide, both the volunteered adverse event reports are deemed to be adverse events by the investigator.

DR. RUSNAK: Correct.

DR. PARKER: And again, the total number of investigators who had the deeming power?

DR. RUSNAK: There's 140 sites in the trial.

DR. PARKER: And do you know the total number of investigators at those sites?

DR. RUSNAK: We could get that information for you at the break.

DR. PARKER: Okay. Thanks.

Dr. Conley? Dr. Hernandez-Diaz?

DR. HERNANDEZ-DIAZ: Yes. I would like the opportunity to -- I'll try a question with Dr. Evins. That was my initial question actually.

If I understand correctly, when we say serious NPS adverse events occur in patients attempting to quit smoking regardless of treatment allocation, do we mean with this that they occur more often in patients attempting to quit smoking than in the baseline population?

For example, in your psychiatric population, do you think that the patients with psychiatric conditions, when they attempt to quit smoking, no matter how, do they have some period of increased risk of these events?

DR. RUSNAK: Dr. Evins?

DR. EVINS: Eden Evins, Mass General Hospital. It's an excellent question. Yes, I think there is a period of perturbation and psychiatric symptoms during a smoking cessation attempt regardless of treatment given. It's generally mild, it's generally transient, and it's generally manageable.

So clinically, when possible -- and it's on a patient-by-patient basis -- I will keep people on their psychiatric medications unless they're having

vomiting with varenicline or they cannot sleep and they're on bupropion, because generally it's manageable, and generally it's due to either the stress of quitting smoking, abstinence, associated withdrawal symptoms, which begin to occur even with smoking reduction, not just abstinence. So, yes.

DR. HERNANDEZ-DIAZ: Thank you.

DR. RUSNAK: May I also invite Dr. Gaffney to address this question?

DR. GAFFNEY: Mike Gaffney, Pfizer statistics. Could we put up slide SAH-1, please? Thank you. In general, we tried to look at patient characteristics in the psychiatric cohort, which were associated with elevated risk of the primary composite endpoint, and we saw quite a few that increased the risk. I want to remind the committee first that we're looking at a cohort that has elevated risk in and of itself. It's higher than those who present in the non-psychiatric cohort.

Within the psychiatric cohort itself, for example, those that have had a history of suicide ideation or behavior are at a 5.8 percent increase

for a positive response. Similarly, alcohol and substance abuse, there's a 3.7 percent increase. Comorbid diagnosis along with their primary psychiatric diagnosis was 3.2 percent.

I won't read through all of them, but you can see that all of them are positive risk features, except for age, which there is a 1 percent decrease per 8.7 years of age. So younger people in the psychiatric group are more susceptible to the NPS AE, and I believe that's correlated with the years smoked because it's saying also that there's a 1 percent decrease per 9.6 years smoked.

Could we go on to slide ST-191? Could you bring that up, please? Thank you. The features, the characteristics that I just showed all could be interrelated themselves. We looked at a multivariate regression to see which of those characteristics present independent addition to the risk of an NPS AE.

We see here the ones that all behave independently, and from these characteristics, you

can almost write who is the subject attempting to quit smoking, who is at most risk. It's actually a young female with a history of alcohol/substance abuse and a history of suicide ideation or behavior, and to increase that a little bit more if their HADS is elevated.

This I think is important from two perspectives. One is the public health finding, which I think is similar to the question you were asking. And secondly, I think it's kind of a validation of the NPS AE endpoint. This endpoint was powerful enough to be able to distinguish these features as being associated with the primary endpoint and increased risk.

It is very important to state that these are true within all of the treatment groups. It's true within placebo, as well as the three active treatment groups. As well as we can tested, there's no significant difference in this association.

So these are the features that EAGLES tells you, along with having a psychiatric diagnosis

which causes these neuropsychiatric events, not treatment -- the NPS AE was not able to pick up a significant treatment effect. It does not mean that it wasn't a powerful tool because we see from these characteristics that it can very well predict who is at risk.

DR. PARKER: Dr. Gerhard?

DR. GERHARD: This question is for -- probably Dr. Evins might be the best person. Just a question that would lead up to the discussion that we're likely to be having on Dr. Hennesey's comment about the risk-benefit.

For somebody like myself who isn't too familiar with the details of smoking cessation and its benefits, could you give an estimate to quantify the benefits or translate the benefit of the difference we see here in successful quit attempts into kind of hard outcomes, cardiovascular events, cancer incidents, mortality rates, just something to kind of give a ballpark of what are these differences that we see between groups mean translated into kind of hard outcomes down the

road.

DR. RUSNAK: I can provide that data to the committee. If you could please show slide PH-58. This slide shows the benefit versus risk treatment with varenicline versus placebo, and benefit was calculated in two ways; first, the benefit to gain one quitter at 12 or 24 weeks, but also the benefit was modeled for the treatment with 12 weeks of varenicline with a sustained smoking cessation at 52 weeks, implying that 52 weeks with a BENESCO model that looks at coronary heart disease, stroke, COPD, and lung cancer -- smoking of course causes a myriad of other illnesses, and this model is limited to only those four benefits.

With respect to smoking cessation in the non-psychiatric and the psychiatric cohort, you would need to treat 4 and 6 patients respectively to gain one quitter at 12 weeks. You would need to treat 7 and 13 patients respectively to gain a quitter at 13 weeks. To prevent one smoking related morbidity over a lifetime, you would have to treat 58 patients, and to prevent one smoking

related death over a lifetime, you would have to treat 93 patients.

With respect to the NPS risk, if you look at the severe intensity only, we were not able to calculate that for the non-psychiatric cohort because the point estimate and the upper bottom of the confidence interval, always below 1, but in the psychiatric cohort, you would have to treat approximately 1,200 subjects to have on severe intensity NPS adverse event.

The overall benefit of the 58 and the 93 is roughly in the ballpark of what people see with statins. This has been calculated with statins. You need to treat for five years with the number needed to treat of 40 to 70 to reduce stroke, MI, or death in that patient population. A similar endpoint for antihypertensive medications, the numbers needed to treat is between 80 and 160, and for aspirin, it's greater than 300.

DR. GERHARD: Thanks, sir. That's very helpful.

DR. PARKER: Last question. Dr. Hennesey?

DR. HENNESEY: Mine got answered. Thank you.

DR. PARKER: Okay. Last question.

DR. EMERSON: This is a question for Dr. Andraca-Carrera. As you've searched through the different models, not like in their primary model, there are several things to change. You change the contrast across groups, the weightings across groups, how you handle that. Which of those things were you most afraid of in that primary analysis?

DR. ANDRACA-CARRERA: We tried to use the same covariates in the model. We looked at Poisson model, zero inflated negative binomial, and binomial. We looked at those models, their AAC and their BAC, and compared them to each other.

DR. EMERSON: But which aspects of the heterogeneity was most fearsome that would cause you to change the endpoints, to change the summary measures?

DR. ANDRACA-CARRERA: We looked at the overall endpoint. So what we did is we first

assumed that the number of events within each side was binomial, which is basically consistent with the primary model, and then we assumed that conditional for all of the other variables, the number of subjects with an event, within a site, follow these distributions. And that's how we calculated the model fit for each of these models. We looked for the primary endpoint, the NPS.

DR. HERTZ: Your question, from a non-statistical perspective, is we don't just look at what we think might cause fear. We look at different sources of unexpected findings or variability, and explore the effects of that on the outcome. So there is no one thing that drives us to do sensitivity analyses.

DR. EMERSON: I understand that. But in switching these models, you're switching from a relative risk -- or you're switching to a relative risk from a risk difference. And how those analyses weight individuals, weight the clinics. And there's really no statistical problem with the primary analysis they did unless you were imagining

that you were fixing perhaps affect modification, unless you thought you were fixing heteroscedasticity. So there are aspects, that sometimes people shift to those models that really didn't matter.

DR. ANDRACA-CARRERA: So we were only part of the interpretation of the risk difference. If the underlying risk is different across all the sites, then you perhaps have a more difficult time interpreting an absolute risk. If the relative risk could be consistent still across sites, it's your interpretation of the parameter.

DR. EMERSON: Oh, it wasn't the underlying risk. You're afraid the risk difference wasn't the same across the sites.

DR. ANDRACA-CARRERA: That's one potential problem, yes.

DR. PARKER: Okay. Let's break now for lunch. We'll reconvene in this room at 1:20, 45 minutes from now. Please take any personal belongings you may want with you at this time. Panel members, please remember that there should be

no discussion of the meeting topic during lunch among ourselves or with any members of the audience. Thank you.

(Whereupon, at 12:38 p.m., a lunch recess was taken.)

A F T E R N O O N S E S S I O N

(1:21 p.m.)

Open Public Hearing

DR. PARKER: Good afternoon, everyone. Thank you.

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

For this reason, FDA encourages you, the open public hearing speaker, at the beginning of your written or oral statement, to advise the committee of any financial relationship that you may have with the sponsor, its product, and, if known, its direct competitors.

For example, this financial information may include the sponsor's payment of your travel, lodging, or other expenses in connection with your attendance at the meeting. Likewise, FDA

encourages you, at the beginning of your statement, to advise the committee if you do not have such financial relationships.

If you choose not to address this issue of financial relationships at the beginning of your statement, it will not preclude you from speaking.

The FDA and this committee place great importance in the open public hearing process. The insights and comments provided can help the agency and this committee in their consideration of the issues before them.

That said, in many instances and for many topics, there will be a variety of opinions. One of our goals today is for this open public hearing to be conducted in a fair and open way, where every participant is listened to carefully, treated with dignity, courtesy, and respect. Therefore, please speak only when recognized by the chairperson. Thank you in advance for your cooperation.

Will speaker number 1 -- you are now at the podium, I see. Will you introduce yourself? State your name and any organization you are representing

for the record.

DR. NIAURA: Good afternoon. My name is Ray Niaura. I'm representing the Society for Research on Nicotine, Tobacco, and I have no financial conflicts of interest to declare.

Thank you for your attention today. I'm here to present a statement from an unconflicted panel of scientists who are members of the Society for Research on Nicotine and Tobacco, known as SRNT. And the panel reviewed the findings from the EAGLES study and the broader evidence base on the efficacy and neuropsychiatric safety of varenicline.

The members were Dr. Steve Bernstein from Yale University; Dr. Matthew Carpenter from the Medical University of South Carolina; Dr. Nancy Rigotti from Harvard University and Mass General Hospital; and, myself, Dr. Ray Niaura from the Truth Initiative and Johns Hopkins University.

The statement was reviewed and approved by the SRNT board.

The main point I'd like to make today has to

do with scientific methods for clinical medical studies and appropriate procedures for assessing the strength of evidence from different kinds of investigations, and this is referred to as the hierarchy of evidence.

Level 1 starts after medication has been approved for marketing, and postmarket data are gathered through a variety of means, and this information may provide a signal regarding possible adverse events. But it is not gathered systematically and via common protocol, so it is prone to error and can be unreliable. However, it must be followed up.

Level 2 data are gathered through observational studies with large and ideally represented population samples over a long period of time. There have so far been several such studies, including analyses of prescription databases in several countries. These studies have demonstrated very low event rates for neuropsychiatric events and no increases with varenicline.

Level 3 and 4 evidence consists of randomized clinical trials and meta-analysis. No conclusive link between varenicline and serious neuropsychiatric events has been found so far.

In an abundance of caution and at the behest of FDA, Pfizer conducted the EAGLES efficacy and safety study with over 8,000 smokers. And this is the top of the evidence hierarchy, because it was specifically designed to look at safety issues.

There were very few significant adverse events overall, which confirmed findings from prior observational studies and clinical trials. Serious event rates were no higher for varenicline compared to another drug, bupropion, the nicotine patch, or even placebo.

In conclusion, appropriate scientific procedures were followed to verify possible evidence for serious neuropsychiatric adverse events that might be caused by varenicline. The highest quality scientific studies did not confirm that there was evidence for serious neuropsychiatric events.

Now, why is all this important? Unless smokers quit, smoking will kill half of them, and smoking is undermanaged and undertreated in medical practice. Varenicline is the most effective medication for smoking cessation, but some doctors and patients are afraid to use it, and this can deprive many smokers of their best chance to quit.

The totality of evidence suggests that varenicline no longer warrants an FDA black box warning, and it should be removed. This, once again, to remind folks, is coming from an unconflicted panel of scientists from SRNT. Thank you very much.

DR. PARKER: Will speaker number 2 step up to the podium and introduce yourself? State your name and any organization that you're representing for the record, please.

DR. ZUCKERMAN: Yes, hi. I'm Dr. Diana Zuckerman. I'm president of the National Center for Health Research. Our center does not take money from pharmaceutical companies. I have no conflicts of interest, except to say we are a

member of the Campaign for Tobacco-Free Kids. I don't know if they have a financial tie to the companies or not.

My training is in psychiatric epidemiology at Yale Medical School, and I'm going to bring that perspective today, but I'll try not to talk about too many numbers.

I wanted to start out by saying there are more than 17,000 serious psychiatric adverse events that have been reported to the FDA pertaining to Chantix, and that's a huge number, 17,000 serious ones.

Just to look at homicidal ideation reports, you can see it's an extremely high number compared to any other psychotropic drugs. These are the ones in second, third, fourth, et cetera, place.

As you have heard, the NAEI is not a validated scale to be used as a checklist, and yet that is how it was used, and there are serious problems of encoding, and that's what I really want to focus on. I want to focus on how hard it is to figure out whether some event is severe or moderate

or mild, and what it really means, and why there is so much difficulty in looking at the differences between agitation and labile mood and anger and depression, and so on.

Here, you have something that's really typical of the patients I have talked to who have had problems with Chantix; just this feeling of being overwhelmed and feeling great fear, but not really knowing why. What do you call that?

This slide is typical of some of the people who have been interviewed, one that I talked to personally, who went to work every day. So he wasn't listed as having been seriously harmed. But he had his own office, he sat on the floor in the corner every day at work, unable to work, feeling like something terrible was going to happen to him, but he just didn't know what it was.

But, fortunately, after a few days of this, he found out that there was a possible link to Chantix, and when he stopped taking it, the symptoms went away.

This is how I feel about my phone half the

time. But not counting that, how do you code this? Apparently, in Bulgaria, this is a normal behavior, but for one of the patients that I talked to, a woman who was a tenured professor at an important college and had a wonderful career, wonderful home life, but when she started taking Chantix, she suddenly just felt really out of control at work, started getting so angry and inappropriate to everyone, dumped her long-time boyfriend, and when he asked her why, she had no idea, and ended up in a psychiatric hospital, without any relationships. And her problem was that this was just before the black box warning went on, so when she went to doctors, nobody knew it might be related.

So there are all these feelings that can be measured in many different ways. Car crashes can be suicide attempts. They can be people out of control. They can be many different things. The key question is, are you sure that a psychiatric event is accurately coded and analyzed.

In conclusion, my concerns are that the issue has to be not benefits versus risks of this

product. Nobody is saying let's take it off the market. What we're saying is that patients and doctors need to have warnings so that when they have bad side effects, they have some idea that it might be related to the drug, so that they can stop taking it and see if that makes a big difference.

It's informed consent, and that is what I think is really essential for all patients. Thank you very much.

DR. PARKER: Thank you. Will speaker number 3 step up to the podium, please? Introduce yourself, state your name, and any organization you're representing for the record. Thank you.

MR. BARS: Good afternoon. My name is Matthew Bars. I'm president of the Association for the Treatment of Tobacco Use and Dependence. I'd like to disclose that I am on the speakers faculty and have consulted with Pfizer. I have no financial interest in the outcome of this meeting.

In addition to ATTUD, which is a global organization of 500 tobacco treatment providers worldwide, I'm also the director of tobacco

treatment for the New York City Fire Department and the Robert Wood Johnson-Barnabas Health-New Jersey City Medical Center I Quit Smoking program. You should try getting that on a business card.

Collectively, as individual clinicians and as the organizations we serve, we have treated hundreds of thousands of tobacco-dependent patients. We believe the EAGLES data strongly supports the removal of the boxed warnings for varenicline and bupropion and respectfully urge this committee's members to so vote.

Whereas others have presented the clinical evidence, my goal here is to share the experience of clinicians who work with tobacco users to become free of this addiction. Our written statement emphasizing the pertinent literature was submitted by ATTUD to this committee under separate cover.

Day in and day out, ATTUD, as clinicians and others, do the very hard work of treating the tobacco dependent. I have worked in the field for over 30 years myself and have personally treated tens of thousands of tobacco-dependent patients.

A case in point I'd like to share today, my patient, Roberta, which is not her real name, is a lovely, 57-year-old African-American woman challenged with schizoaffective disorder. She cannot tell you who the vice president of the United States is, but is aware, in the nonclinical sense, of the boxed warnings of varenicline's neuropsychiatric adverse events.

For example, during intake, while discussing medication options, Roberta commented, quote, "I heard Chantix can make your head explode," end quote. While this statement may seem a bit extreme, many individuals have a faulty or exaggerated perception of the dangers associated with these medications. As committee members may be aware, 44 percent of all cigarettes sold in the United States are purchased by persons with mental illness.

As tobacco treatment providers, we often find ourselves in clinical situations where are patients are more fearful of using FDA-approved medications than they are of continued smoking.

This is not helpful for the patient or clinicians or public health. At present, a very small percentage of tobacco-dependent patients are prescribed and receive these FDA-approved medications.

We believe that neither clinicians treating tobacco dependence nor tobacco users seeking treatment should be discouraged from prescribing or using these medications. The boxed warnings we are discussing currently have just that impact, reducing our capacity to effectively treat the most preventable cause of death and disability.

The EAGLES study's findings should reassure a wary population of smokers and health care providers about the efficacy and safety of bupropion and varenicline. Removal of these warnings will help assure America's 43 million smokers have one less reason to avoid tobacco treatment.

A little talk has been given this morning regarding what should be the true comparator, and that is continued smoking and eventual death.

My colleagues and I are really good at treating the adverse events that are associated with these medications and tobacco withdrawal symptoms. Treating death is way over my pay grade. Thank you.

DR. PARKER: Will speaker number 4 step up to the podium? Introduce yourself, state your name and any organization you're representing for the record, please.

DR. FOX-RAWLINGS: Thank you for the opportunity to speak today. My name is Dr. Stephanie Fox-Rawlings, and I am speaking on behalf of the many members of the Patient, Consumer, and Public Health Coalition.

The coalition includes nonprofit organizations representing millions of patients, consumers, researchers, and doctors united to ensure that medical treatments are safe and effective. The coalition does not have paid staff and does not accept funding from any outside sources, so I have no conflicts of interest.

Pfizer is once again asking the FDA to

remove the black box warning that Chantix is associated with serious adverse events, such as depression, hostility, agitation, suicidal thoughts, attempts, and completion. They want to replace it with a statement that these are associated with quitting smoking. They also want to remove the warning that there may be an increased risk for patients with a psychiatric illness. GlaxoSmithKline wants to remove the REMS requirement for Zyban.

They base these changes on one large, poorly executed clinical study. It is important to point out that these black box warnings were initiated because of the enormous number of extreme, serious psychiatric adverse events, including suicide, aggressive behavior associated with smoking cessation products.

Research has also confirmed that some patients have extreme psychiatric responses that can be deadly to themselves and others.

The purpose of these warnings is to let patients know that if they seem to be having

uncontrollable feelings when on these drugs, that there's a good chance that getting off of the drug will help solve the problem almost immediately.

Pfizer's study concludes that Chantix does not have these risks, but the FDA reviewers have clearly shown that are extensive problems with how the data was collected and analyzed.

First, the study measured psychiatric problems with the NAEI. This is not a validated test, so it is only supposed to be used to start the conversation about psychiatric symptoms. Instead, it was used as an unvalidated checklist, which contributed to inaccurate data. For example, it did not identify cases of suicidal behaviors that were identified by validated scales.

Second, when patients reported psychiatric problems, those problems were not coded consistently. The FDA pointed out that the staff doing the interviews and coding were not always trained mental health professionals, and they didn't seem to understand some of the categories they were coding.

Even worse, their very subjective measures of the severity were sometimes completely incorrect, such as a patient who became severely depressed being coded as having a mild problem from taking Chantix.

Third, since 70 percent had tried to quit smoking previously using one of these drugs, the study was biased toward people that previously tolerated the drug. This would drastically underestimate the percentage having serious adverse reactions.

In addition, anyone with suicidal thoughts or behaviors in the past year or anyone with self-injurious behaviors were excluded. While these patients should not be treated with a drug that would make these worse, this could also bias the results to make the drugs seem safer than they really were.

In summary, patients deserve access to smoking cessation treatments, but they also deserve warnings about the risks. There remains considerable credible evidence that some patients

are severely harmed by Chantix and Zyban, and those patients' lives depend on warnings about these risks so they will recognize the sudden suicidal, paranoid, or violent thoughts as side effects of the drugs.

Thank you for your time and consideration of our views.

DR. PARKER: Will speaker number 5 step up to the podium? Please introduce yourself, state your name, and any organization you're representing for the record, please.

MR. MOORE: My name is Thomas Moore. I'm senior scientist for the nonprofit Institute for Safe Medication Practices, and I have no financial interests to declare and was not supported by anyone in making this presentation.

I think we have a barius [indiscernible] proceeding, not intentionally, but to assess a drug adverse event really requires us to think about five lines of scientific evidence. And today we spend about 80 percent of the time on one line of scientific evidence, about 20 percent on the second

one, which was inconclusive.

So I would like to use the time that I have to look at the evidence you are not seeing and summarizes it very briefly.

These are the three lines of scientific evidence for which we have multiple publications and multiple people, different countries, and we've all reached very similar conclusions. The most important one we really haven't heard about is are serious psychiatric adverse events and particular bizarre or aberrant behaviors, are they plausible given how this drug works. And the answer to that is clearly it falls somewhere between plausible and probable.

This is an alpha-4 beta-2 nicotinic acid receptor, partial agonist-antagonist, which causes the release of dopamine. We know quite a lot about dopamine, and we know that this drug is active in dopamine pathways, because we see nausea and we see abnormal sleep patterns, which clearly are mediated in this pathway.

Let's move on to the second part, which is

case reports, including the narratives, many of which were flawed in this study, really form the core of how we decide whether a drug was really causing the effective.

We have elaborate protocols, which are widely used, and so we have many, many convincing case reports in patients who had no previous history, who had symptoms before the smoking date cessation, whose problems resolved when they stopped the drug, and we have a smaller number of re-challenge cases where they clearly reappeared when the drug was restarted.

The other part about these case reports to remember is this was not just done by ISMP. There are three FDA pharmacovigilance reports with striking case studies that struck them as credible and important, as well as a peer-reviewed ISMP paper in medical literature.

This is just to give you the flavor of what one looks like and how complex they might be to code. "I was completely out of control. I woke my boyfriend up in the middle of the night and started

physically beating him."

The problem with case reports and the limitation is they tell you if some cases are happening, but they really tell you very little about how many. We have statistical studies that were completed by the FDA, by ISMP, and by the French, and all of us found many more than expected cases.

Here is just one little example. What do we see here when we're looking at suicidal and homicidal thoughts? What we see is Chantix was three times more than any other drug.

Now, I'd like to ask another question. How many of these drugs on that list you see right there have greater person-years of exposure? And the answer is all of them. How many of them had suicide behavioral warnings? The answer is also all of them.

I have run out of time, so I will have to leave that slide for you to consider. But this trial, as we have heard, has many, many defects. Thank you for your consideration.

DR. PARKER: Will speaker number 6 step to the podium? Please introduce yourself, state your name, and any organization you're representing for the record.

DR. ALMASHAT: My name is Sammy Almashat. I'm a physician and researcher with Public Citizen. I have no financial conflicts of interest, but Public Citizen was a cosignatory to a petition to the FDA in 2014 for a stronger boxed warning on Chantix.

First of all, I want to reiterate that Public Citizen is in favor of keeping Chantix on the market. We think it is a good drug. We think it should be used in patients. We are simply in favor of retaining a warning to those patients in case they do experience an adverse event that, which I will go into in my talk, was not adequately assessed in this randomized trial.

It's important to remember that the boxed warning was placed on Chantix in 2009 due to a deluge of postmarketing adverse event reports of suicidality and neuropsychiatric events. Up to

15,000 serious psychiatric events have been reported so far.

Now, the EAGLES trial was powered to detect an absolute difference in event rates between Chantix and placebo between 26 and 52 events per 1,000 patients. Now, these seem to be very high estimates of the absolute risk difference between Chantix and placebo.

A back-of-the-envelope calculation shows that if we assumed a 10 percent reporting rate of voluntary adverse event reports to the AERS database over the last 10 years, that would represent approximately 10 per 1,000 neuropsychiatric events that have been reported to the agency over the past 10 years.

This is roughly the same order of magnitude of excess risk on which the FDA based its suicidality warning on antiepileptic drugs, which was approximately 20 per 1,000; so, far off from the up to 50 per 1,000 in psychiatric subjects that this trial was powered to detect.

The other problem with the trial was the

issue of the inconsistency of data reporting. I won't go into the details, but pages 46 to 49 of the briefing packet detail the FDA's serious concern with how the adverse events were collected and classified.

These inconsistencies led the FDA reviewers to conclude that, quote, "the exact incidence of neuropsychiatric adverse events of significance and perhaps their scope was not accurately captured by the study."

In a trial that found a numerically and almost statistically significantly increased risk of neuropsychiatric events in psychiatric patients between varenicline and placebo patients, even a few events either way that were not adequately captured or were missed during the data collection process could have tilted this toward a significant finding. And it is important to ask ourselves what then would our conclusion be about removing a boxed warning in the face of a significant finding of increased risk with Chantix relative to placebo.

It is also important to remember that this

trial would be the sole basis by which you would be voting to remove a boxed warning on a drug. Two years ago, you voted 17-1 to retain the boxed warning in the face of all of the evidence, including the adverse event reports, including the observational studies that were conducted up to that time. And as the FDA noted, there is very limited precedent for removing a boxed warning, so we argue that the threshold for evidence to do so should be very high.

Again, we think that Chantix is an important drug, we do, and we just think that even if it is a rare adverse event, it is a life-threatening adverse event that was not adequately assessed in the study, and that patients should be warned about the event, should they experience it, so that appropriate action can be taken.

DR. PARKER: Speaker number 7, if you'll step up to the podium and introduce yourself. State your name and any organization you're representing for the record. Thank you.

MS. SOUTHARD: Good afternoon. My name is

Carol Southard. I am based at Northwestern Medicine in Chicago. I am a tobacco treatment specialist. I have been in the field for over 30 years. I feel very old when I say that. I have seen over 3,000 clients.

I am speaking to you not as a researcher, but as a clinician who has been in the field, who reads the literature religiously. And I am here to advocate changing the label, because what I think is getting missed in this discussion is the fact that tobacco users are not being treated in terms of what is recommended by our clinical practice guideline, which was last updated in 2008.

But back then, which was based on over 7,000 clinical trials and surveys, as you well know, it was stated that all tobacco users should be offered some form of cessation pharmacotherapy. That has not happened. There has been data showing that less than 8 percent of tobacco users are given any kind of pharmacotherapy. And there was a study that just came out that said in some states, only 1 percent of Medicare and Medicaid patients were

being offered a cessation pharmacotherapy.

My concern is because of these alarming labels that are on these products, that they are not being offered by providers, and many, many tobacco users, as you have heard, are afraid to use them. And I think it is because of false information that is out there.

There are still, as has been said, over 40 million Americans who are still using tobacco. It is not that we don't know how to help smokers quit, it's that it is not being done, and that's the disconnect that troubles me the most.

More than 95 percent of smokers try to quit without any kind of treatment, even though all the evidence says, even brief counseling, plus use of cessation medication significantly increases success rates.

I have the luxury of over an hour with my clients. I am very assertive about use of cessation medication. In fact, I have greater than 56 percent success rate a year, which is phenomenal. I should be rich and famous.

But what's really important is the majority of my clients try a cessation medication. Most of them are not comfortable with cessation medications, because, I think, of the false information that's out there, that's been made front page news, and that's not in the evidence.

I'm not cavalier about medications. I don't want any of you to think that. But I'm a huge advocate of use of because I've seen firsthand what's in the literature. Using medication increases success rates. And even if there was a risk with these medications, every medication has a risk, I understand that, the benefits of quitting far outweigh any risk that could occur.

I have had over 900 clients on Chantix. One report of a man who did have a history of psychiatric comorbidities reported increased incidence of -- he was hearing voices, mild schizophrenia. I had one very young man who felt that he was feeling suicidal. Of course, I took them off Chantix immediately.

But what I want you to hear is the majority

of my clients have no untoward effects from Chantix, and that is, frankly, true with the majority of my clients who use medications, which, as I said, is most of them.

I'm not going to go over the EAGLES trial, because that has certainly been done, but I do want to reiterate that these medications are safe, they are effective. I have had firsthand experience with them. And despite the knowledge of the tobacco risks, we haven't achieved the goal of making tobacco use a rare occurrence in this country.

I hope that this committee will decide to change the labeling, take the black box label off so that providers will be much more comfortable in use of, and the tobacco users will be much more comfortable in using them, as well. Thank you for your attention.

DR. PARKER: Speaker 8, if you'll introduce yourself. State your name and any organization you're representing for the record, please.

DR. BERGER: My name is Dr. Tom Berger. I'm

executive director of the Veterans' Health Council for Vietnam Veterans of America. I have no financial interests in the outcome of this meeting.

One of the major issues we have learned in the years following the war in Vietnam is that combat exposure to veterans gives the high risks that can affect their health throughout the rest of their lives.

In our war, for example -- I'm speaking about Vietnam, but it's the rest of your war, too, who were around at the time -- PTSD is a condition which, for many of us, has impacted our lives long past the end of our military service.

Then there's smoking. It's well documented that individuals coping with mental health issues are two to three times more likely to smoke. Similarly, some groups of veterans have higher rates of tobacco use, including those with psychiatric disorders, such as depression or PTSD.

As a matter of fact, among veterans, mental illness and smoking are tightly linked, with PTSD being a known risk factor that increases the

likelihood of smoking. And in case you didn't know, currently, 60 percent of Vietnam veterans with PTSD smoke.

In addition, the CDC reports the following data from the years 2007 to 2010, that male veterans aged 25 to 64 years old were more likely to be current smokers than non-veterans.

The fact of the matter is that in order to effectively treat the total health of veterans with mental illness and reduce smoking rates in all our veterans populations, treatment plans must combine specific smoking cessation initiatives, including pharmacotherapies and mental health programs.

The integration of smoking cessation activities, programs, into mental health programs is critical to addressing the compounded mental and physical health issues of Vietnam veterans, in particular, especially those suffering from mental health illnesses. The unique needs of veteran smokers living with mental illness must be met to help them quit smoking and share in the positive results of decreased tobacco usage.

Now, while the VA currently has programs in place that try to lessen the toll of tobacco related consequences on veterans, especially veterans with mental illness, we need renewed emphasis and commitment to this issue.

There are things, activities, in which all of us must focus on access to treatment and resources by the VA, by yourselves, that is, members of the FDA, and the CDC to ensure that vets receive quality health care, quality regarding access to smoking cessation treatments to make certain no veteran is left behind. Thank you very much.

DR. PARKER: Will speaker number 9 introduce yourself?

(No response.)

DR. PARKER: We'll move to speaker number 10. If you will, step up and introduce yourself. State your name and any organization you're representing for the record. Thank you.

MR. COUNTS: My name is Nathaniel Counts, director of policy at Mental Health America. And

our disclosure, I think about five years ago, we had funding from Pfizer, and we might take it again in the future, but presently have no financial interests, especially in the outcome of this day.

Mental Health America was founded in 1907 by an individual with lived experience, and since then, we have grown to over 200 affiliates nationwide and a growing number of associate members. So we have a lot of interest and experience in making sure that people with mental health conditions have the best chance of a happy life in the community.

We're really here today to thank you for careful consideration of this issue, the considering of revising the black box in light of the published study in the Lancet, and mostly just highlight the opportunity presented by all of this.

Our stance is we need all the tools and options that can be safely made available to individuals with mental health conditions.

Without going through the study, mostly just to talk about the surgeon general finding in 2014

that 5.6 million people between the ages of zero to 17 -- and since that was 2014, now it would be 2 to 19 -- will die prematurely because of smoking related causes. And if people with mental health conditions, according to SAMHSA, smoke 40 percent of the cigarettes, that means a disproportionate share of those individuals will be people with mental health conditions.

Just highlighting, given the fact that people who have the option to quit before the age of 40 have a 90 percent reduced likelihood of mortality from smoking related causes, there is a chance to, if additional options are made available, prevent at least some of those 5.6 million deaths.

We thank you for your time and very careful consideration of the issue.

DR. PARKER: Thank you. Speaker number 11, if you will introduce yourself, state your name and any organization you're representing for the record.

DR. SACHS: Good afternoon. I'm Dr. David

Peter Sachs, a pulmonary medicine and clinical care medicine physician and specialist for over 35 years in this field. And because of the toll I saw early on in my pulmonary medicine training at Stanford, I decided I needed to become more actively involved in development of the treatments to help people stop smoking, because we can't treat lung cancer very well even today, let alone back in the '70s when I was a pulmonary fellow. We cannot treat COPD very well today, let alone back 30, 40 years ago.

I also am the chair of the American College of Chest Physicians' tobacco dependence treatment committee, and we produced the 2010 tobacco dependence treatment toolkit approved by the board of regents, with external review.

I am also a member of the American Thoracic Society, the largest pulmonary medical organization in the world; and I serve on the tobacco action committee.

Our committee, independent of me, prepared this letter for you, and I hope you've had a chance

to review it and read it because this is official American Thoracic Society policy, and that is that the black box warning should be removed, because it deters both physicians and patients from using effective medications, specifically, both varenicline and bupropion.

I have no conflicts of interest to declare. I have flown here from California on my own dime. Since 1985, I have conducted over 30 tobacco dependence treatment trials, and I have also personally, in my pulmonary medical practice, as part of my routine pulmonary medical care, treated over 7500 tobacco-dependent patients one-on-one.

Now, I mention this because I have spent my life, my career, over the last 35-40 years, treating the downstream consequences of tobacco dependence. When I sit down, three minutes from when I began, three Americans will have died from tobacco dependence and the myriad diseases that it causes.

By the end of this meeting today, from the time it began, 500 Americans will have died from

tobacco dependence and the diseases it causes. Twenty-four hours from the time this meeting began, over 1,200 Americans will have died from the multitude of tobacco dependence diseases.

These causes of death include, but are not limited to, lung cancer, heart attack, and stroke. Tobacco dependence causes 18 percent of all deaths in the United States and 10 percent of all hospital costs in the United States. This need not be.

Tobacco dependence is treatable, like any other serious chronic medical disease, which tobacco dependence is. It is not a habit. It is a CNS-based disease.

I have a handout for the FDA committee, which is outside. I was going to present to you a short case summary of an attorney I treated with varenicline, but there's no time.

He needed actually 18 months of varenicline treatment, and in the first 6 weeks, he needed a dose as high as 5 milligrams per day in order to suppress all nicotine withdrawal symptoms. When he tried to taper too soon, he relapsed.

I would urge you, please, remove the black box warning, because, remember, cigarette smoking kills. Varenicline, bupropion, and nicotine patch don't. Thank you for your attention.

DR. PARKER: Speaker number 12, if you'll introduce yourself. State your name and any organization you're representing for the record, please.

DR. KERKVLIET: Hi. Thank you for the opportunity to speak today. My name is Gary Kerkvliet, and I am here on behalf of myself, although I have spoken on behalf of Chantix as a useful drug by Pfizer.

I'm here because I'm in the trenches, and I'd reiterate what other speakers have said about the difficulty of treating the tobacco user. I come from a slightly interesting point of view, because although I have never smoked, I'm a physician who can prescribe the medication, and I have also suffered from major depressive disorder. And I realize that in this population, you have to be very careful about any medications that you use.

We know that there's, as many speakers have mentioned, a prevalence of smoking in people with psychiatric disorders, major depressive disorders, and I'd like to reiterate that I think the black box warning is important in pointing out those things that one should consider, although the article in discussion today certainly shows that maybe it's not as bad as we think it is.

My concern is that as before the black box warning and certainly after the black box warning, I was reticent to use the medication because actually I have heard from -- the patients have already heard. They don't want to take Chantix possibly because they've heard things about it. We've seen a number of studies that show that there's not a major difference. And I think we need to remember, too, that nicotine withdrawal is going to be giving some of the symptoms; obviously, not all.

I previously had been of the mind-set that if a patient had major depressive disorder, perhaps it was okay to let them keep smoking until we got

them through the difficult part of their depression. But, in fact, varenicline can be used very safely. I have seen that a number of times.

Again, I would just like to say that if the black box warning is removed, I think we will see the use of it increase. As with any patient, you're going to discuss side effects, possible side effects with them, and I think as long as that's monitored well by the physician, that it's a safe medication to use. Thank you very much.

DR. PARKER: Speaker number 13, if you'll step up and introduce yourself. State your name and the organization you're representing for the record, please.

MS. FODERINGHAM: Good afternoon. My name is Shelina Foderingham. I'm with the National Council for Behavioral Health, and I have no financial conflict of interest to declare.

The National Council for Behavioral Health appreciates the opportunity to provide commentary on the labeling of prescription drugs that treat tobacco addiction. As an association representing

more than 3,000 community-based behavioral health organizations who serve 10 million patients annually, the National Council strongly supports evidence-based approaches to eliminating tobacco consumption by people living with mental health and substance use disorders.

We agree with the growing body of research, which includes that pharmacological interventions paired with behavioral health services are efficacious and improve the likelihood of long-term tobacco abstinence.

The National Council has long advocated for policies that maximize access to effective behavioral health, pharmacological and medication-assisted treatment interventions. The National Council is also doing work to support tobacco cessation in states, tribes, and provider organizations across the country.

To this end, we support the removal of the FDA's black box warning label on varenicline, as it serves as an unwarranted barrier to treatment. The National Council's position on this topic is

informed by robust evidence indicating varenicline's effectiveness and the fact that people living with mental health and substance use disorders are more likely to consume tobacco.

I'd like you to consider the following: people living with mental health and substance use disorders often experience shorter than average life spans. These disparate outcomes in mortality are exacerbated by tobacco consumption. People living with mental health and substance use disorders are also more likely to consume tobacco, as you heard from previous presenters.

While people living with mental health conditions represent nearly a quarter of the overall adult population, they consume nearly 40 percent of all cigarettes.

Adverse neuropsychiatric effects from the use of varenicline are very rare. Peer-reviewed analyses indicate that when a mental health disorder is already present, varenicline has not been shown to exacerbate neuropsychiatric symptoms.

Also, when varenicline is used as

prescribed, there is no evidence or little evidence of increased risk of suicide, attempted suicide, suicidal ideation, depression, or death. In fact, studies show that pairing varenicline with behavioral health and/or other pharmacological interventions can reduce the likelihood of already rare adverse neuropsychiatric reactions among people that have a history of attempted suicide, suicidal ideation, or depression.

Clinicians, as you heard, are using varenicline with great success and few to no adverse reactions. As just one example, as you heard earlier, a tobacco cessation specialist within Northwestern Medicine in Chicago saw an 87 percent success rate 12 months post-treatment in those clients who utilized pharmacotherapy for at least three months compared to 56 percent among patients overall, with very few complaints of adverse reactions.

This example echoes many we have received from our members from across the country. More of these examples can be found in the addendum to our

written testimony.

Eliminating tobacco consumption among behavioral health clients, staff, and practice settings requires the sensible deployment of all effective tools. Accordingly, the National Council urges the removal of the FDA's black box warning for varenicline. I appreciate your time and consideration. Thank you.

DR. PARKER: Thank you. Speaker number 14, if you'll step forward and introduce yourself, state your name and the organization you're representing. Thank you.

MR. MYERS: Thank you. My name is Matthew Myers. I'm the president of the Campaign for Tobacco-Free Kids, both this nation's and the globe's largest advocacy organization devoted exclusively to reducing tobacco use.

I was about to say that I have no conflicts, but I realized as I was coming, both GlaxoSmithKline and Pfizer have made contributions to our annual fundraising gala. It amounts to less than one-quarter of 1 percent of our annual

funding, and they had nothing to do with our presence here today.

I would like to explicitly talk about what has implicitly been discussed, and that is the challenging job that FDA has to put in context the review of tobacco cessation products. And I think that is what -- and our organization believes that's what truly been missing from the review of tobacco cessation products over the years.

You've heard many spokesmen already talk about the health effects of tobacco. Despite all of the progress we have already made in the United States, current estimates are that we still have close to 480,000 Americans dying from tobacco use.

Tobacco use isn't a behavior. It is an illness, tobacco addiction. If I replace the term "tobacco addiction" with "lung cancer" and told you that one out of two long-term users would die, that close to half a million Americans every year would die, that today over 1,000 Americans would die, and that in the last 30 years, we have exactly three new drugs that have been approved, that virtually

no true innovation, while each of these drugs has been shown to improve the likelihood that an individual would smoke. But we still have success rates far below what are necessary to treat literally what is an epidemic.

So in many critical respects, you are here today to take a very narrow, very focused look at a specific study that looked at risks, not benefits, of a particular set of drugs, when the real question, I believe, needs to be from FDA, which is how do you use your authority to ensure that you're fostering a discussion about how do we produce the most effective drugs, deliver them to the widest population with the least harm, but the greatest public health.

The real measure ought to be, how do we use the power of the Food and Drug Administration to reduce the number of Americans every year who die from tobacco use?

This has become even more important, for two reasons. In 2009, Congress gave the Food and Drug Administration authority over all tobacco products.

And in 2010, the courts defined nicotine derived from tobacco as a tobacco product.

What that means is that while, before 2009, the FDA was able to carefully control the delivery of nicotine, since 2010, we have had a situation where nicotine is being delivered widely to consumers of all ages in completely uncontrolled doses, often discouraging Americans from using the most effective products, often resulting in Americans who want to quit to use products that are not effective.

What I would urge your advisory committees to do is begin the real conversation that I think is necessary, and that is how do both CDER and the Center for Tobacco Products combine their authority to maximize the discussion about how we promote the creation, development, and marketing of the most effective products designed to reduce the number of Americans who die from tobacco use to the greatest degree possible. Thank you.

DR. PARKER: Speaker number 15, if you'll step up and introduce yourself, state your name and

any organization you're representing for the record, please.

DR. SPERLING: Good afternoon. My name is Andrew Sperling. I'm with the National Alliance on Mental Illness. I'm here in place of our medical director, Dr. Ken Duckworth, who could not be here today. NAMI has no financial stake in the outcome of this meeting, and NAMI paid for me to be here. Noone paid for me to be here or cover my expenses.

NAMI is the nation's largest organization representing people living with serious mental illness and their families. We have over 1,000 organizations all across the country and advocate for people living with disorders such as schizophrenia, bipolar disorder, and major depression.

You've heard some numbers here today at this hearing about early mortality and mental illness, and they are fairly shocking statistics. You hear different numbers largely because that denominator is sometimes different, the comparator group and the general population we're comparing it to.

You've heard numbers about 18 years of lower life expectancy, 20 years, 24 years of lower life expectancy.

The easy takeaway, the easy measures, just to remember, that if you're an adult living with schizophrenia or bipolar disorder in America, your life expectancy hovers just below an adult in Bangladesh.

This early mortality is largely not due to the underlying psychiatric illness. It's due to lots of comorbid chronic medical conditions, most of which are linked to high rates of tobacco consumption.

People with mental illness not only smoke in higher volumes, they smoke differently. We don't actually know and have the cause of this yet, but we believe that nicotine can actually, very temporarily and on a short-term basis, relieve the symptoms of paranoid delusions or auditory hallucinations. So they smoke more and they smoke differently.

When they're smoking, they draw more

heavily, and they smoke much, much higher volumes than the general smoking population does. And this is a major contributor to comorbid chronic illnesses and early mortality. It is a public health crisis that we're just now coming to grips with, and it is a major public health crisis that NAMI is very, very concerned about.

People with mental illness face bigger challenges in quitting smoking than the general smoking population does, and their relapse rates, even when they've been able to quit on a temporary basis, are much, much higher than any other measured population.

So the single biggest thing we can do to improve the public health of people living with mental illness in this country is to address tobacco consumption and tobacco related illnesses.

That is why NAMI believes that the FDA needs to ensure broad access to the full range of smoking cessation therapies to help people with mental illness quit, including addressing and removing this current black box warning that is keeping

tobacco cessation from getting to people who need it most. Thank you very much.

DR. PARKER: Speaker 16, if you'll step up and introduce yourself. State your name and any organization you're representing for the record. Thank you.

MS. WITCZAK: Good afternoon. My name is Kim Witczak, and I came here on my own. Thank you for the opportunity to address this committee.

As you heard earlier this morning, I was recused from serving on today's advisory committee as consumer rep because of a lawsuit against Pfizer for an unrelated drug that was resolved almost 10 years ago.

Since I had spent the time preparing and studying the briefing documents, I felt it was important that I was here and represented the consumer perspective.

There is no doubt that cigarette smoking is a huge contributing factor to premature deaths in this country and, in fact, around the world. And I fully support the need for having treatment

options, including the drugs that we are discussing today, available for smokers to help them quit.

With this being said, there are a couple of things that caught my eye about this large safety study that we should consider before we remove the black box warning.

As we heard earlier, the FDA found problems with the study accurately identifying the psychiatric events. For example, I'm personally concerned that agitation and anger cases were coded as irritability, which could be seen as a result of just quitting smoking.

I also, like you, wanted to know more about the two intentional overdose cases that were not coded as suicide attempts. These are just a few examples with coding after reviewing the specific patient individual cases. But I also wonder how the payments that the investigators and sites received from Pfizer, that may or may not have influenced any of the results.

The bottom line is that I think there's too much room for subjectivity or incomplete

information around the narrative and coding of the events.

I know there was more information in our FDA briefing packets about the adverse events that were reported through MedWatch, and we heard a little bit about earlier, but I also would love to have heard about the 2700 victims from the lawsuits that weren't able to present.

Here is the current medication guide that we're looking at today, and I actually think this is a really strong medication guide. It lets people know that some people have had serious side effects while using Chantix to help them quit smoking. Some people had these symptoms when they began taking Chantix, and others developed them at several weeks of treatment or after stopping.

Stop taking Chantix and call your doctor right away if you or your family or caregiver notice agitation, hostility, depression, or changes in your behavior, or thinking that's not typical, things like thoughts of suicide, depression, anxiety, panic attack, agitation, restlessness,

aggressive behavior. You can read the warnings up there.

This is really meant to be a conversation with our doctor, and we're potentially taking that away.

Now, let's look at the proposal that's on the table. But at the end of the day, the real question, in my mind, given all the earlier comments from the FDA and all the anecdotal reports that the FDA has received over the years, is can we, in good conscience, sleep comfortably tonight knowing that it's truly safe enough to eliminate and delete an important patient protection.

There are real-world consequences to this decision, and I'd like to thank you for being willing to listen, and I look forward to the discussion.

DR. PARKER: The open public hearing portion of this meeting has now concluded, and we will no longer take comments from the audience. The committee will now turn its attention to address the task at hand, the careful consideration of the

data before the committee, as well as the public comments.

We'll turn now and ask Dr. Racoosin to provide us with the charge to the committee. Thank you.

Charge to the Committee – Judith Racoosin

DR. RACOOSIN: Good afternoon. Today you have heard presentations from industry and the FDA about the safety outcomes trial that FDA required after the emergence of the concerns about the risk of neuropsychiatric adverse events with smoking cessation drugs Chantix and Zyban, as well as discussions of the observational studies that have been published on this topic. You have also heard from members of the public who have traveled here to share their thoughts about this topic.

Now, we turn to all of you for an in-depth discussion of the questions that we'd like you to consider. I'm going to run through the questions as a group, and then you'll consider them one by one.

First, discuss the strengths and weaknesses

of the completed randomized controlled trial with regard to the study design, including the novel primary endpoint.

Two, discuss the potential impact of the variability in data collection, adverse event coding, and case definition on the primary endpoint. Because of this variability, discuss which analysis and results -- and by this, I mean the sensitivity analyses -- and results that most appropriately describe the effect of the smoking cessation therapies on neuropsychiatric events.

Three, discuss how you weigh the evidence contributed by the observational studies when evaluating the risk of serious neuropsychiatric adverse events in patients taking smoking cessation therapies.

Four, based on the results of the clinical trial and observational studies, discuss the impact of psychiatric history on the occurrence of neuropsychiatric adverse events during smoking cessation therapy.

Five, and this is a voting question, based

on the data presented on the risk of serious neuropsychiatric adverse events with smoking cessation products, what would you recommend: A, remove the boxed warning statements regarding the risk of serious neuropsychiatric adverse events; B, modify the language in the boxed warning; or, C, keep the current boxed warning?

Six is related to question 5, which is explain the rationale for your answer to question 5, and discuss any additional labeling actions you think the agency should take regarding the risk of serious neuropsychiatric adverse events with smoking cessation products.

With regard to this last question, I want to comment that some of the discussion today has had a certain flavor of being an all or nothing component about how FDA might handle labeling of this safety issue, and I think it is important for you to discuss any level of or part of labeling that you think would be applicable, including this morning I talked about the boxed warning, the warnings and precautions section. The medication guide is

another important part of labeling that relates to this issue. Really, whatever section that you think might be in play, we would be interested in hearing your thoughts. Thank you.

Questions to the Committee and Discussion

DR. PARKER: Thank you.

We're 19, aren't we? I'd like to defer to as many people as would want to comment on each of these to be allowed to. So what I thought we would do is I really would like for each member to have an opportunity to comment on all these topics.

In order to maximize that opportunity for input, I thought what we could do is we'll take each question in order, and we'll start with any clarifications that anyone has regarding the content of that discussion point, so that we're on point with our comments about each one.

So if you have clarifying questions related to the topic, we'll start with that, so that we're clear what we're discussing. Then we'll go around, and I'll ask that everyone have an opportunity to make a succinct comment, and if you have no

specific comment or its repetitive of what has already been said, you can simply say "no comment" or "I agree with" whatever has been said.

At the end of each discussion point, I will summarize for the record what I've heard, and then we'll move on to the next one. Game plan.

Let's put up the topic number 1 for discussion, discuss the strengths and weaknesses of the completed randomized controlled trial with regard to the study design, including the novel primary endpoint.

Let me ask if there is anyone who has any specific clarification that they would like about that topic, that they would like to pose to the agency in order to make sure we're clear what we're actually discussing and what content will be most useful. I see Dr. Roumie has her hand up.

DR. ROUMIE: I'm just wondering whether or not the sponsor got any information on the total number of hospitalizations in each of the four arms, which are typical SAEs in clinical trials, hospitalization, any cause.

DR. RUSNAK: Yes, we do. We do have some follow-up information on that, as well as the other items that we were looking for during the break.

If I could have the slides for the hospitalizations projected, please. Maybe while that's coming up, I'll tackle one of the other components. For the NAEI, it was used in 15 languages, and it was linguistically validated.

Moving over to the hospitalizations, the subjects with NPS AEs leading to hospitalizations in the United States by treatment group and the overall study population are on this slide for the psychiatric cohort.

As you can see, the events were relatively infrequent, 4, 3 and 3 amongst the active treatment groups and one for placebo, and then they are further subdivided by mood disorders, anxiety disorders, and psychotic disorders.

DR. PARKER: Dr. Morgan, I believe.

DR. MORGAN: [Inaudible]. This has to do with me forgetting to put the mic on -- the conclusion set, in one of the FDA talks, I think a

summary of the study. And I may be mischaracterizing this, but I think I have it right.

It was stated that serious or significant neuropsychiatric AEs occurred in all treatment groups, both within the non-psychiatric history and the psychiatric history cohort. Treatment groups were used.

Were you including the placebo group in that characterization, no difference between placebo group, as well as the three treatment conditions?

DR. WINCHELL: Yes. The treatment groups include the placebo groups.

DR. RUSNAK: No difference in adverse events between the medication groups and placebo.

DR. WINCHELL: I'm not sure exactly which slide you're quoting, but, yes, our conclusion was that in patients without psychiatric history, serious and clinically significant, which is not always serious, regulatorily speaking, events occurred at similar rates across all treatment groups.

DR. RUSNAK: That was my interpretation. I wanted to make sure. Thank you.

DR. PARKER: Dr. Morrato?

DR. MORRATO: According to the study design, certain specialties or trainings were the ones that were collecting the adverse events, and there was mention that every six months, I believe, they would get retrained.

Do we have any data on the adherence with that, either qualification of being able to get the events and then the training throughout the study?

DR. RUSNAK: If I could have Dr. McRae address that question, please.

DR. McRAE: Good afternoon. Tom McRae, clinical sciences, Pfizer. The question was specifically in regard to the NAEI; is that correct? Individuals who conducted the NAEI were qualified by virtue of training to do so, and then they were required to have refresher training every six months during the course of the study.

In most cases, these were sub-investigators at the study sites. So they had professional

qualifications of some sort, but they did not have to be trained mental health professionals.

DR. PARKER: If I could just redirect this slightly. What I would like for us to do is focus on any clarification to the agency regarding the topics of discussion.

They have presented us with five different discussion topics, and I'm going to go around and give everyone an opportunity, if they would like, to provide a comment regarding each of these. But before I began that, what I wanted to say, do you have any question about what it is you're responding to as you read the topic right here for discussion under number 1?

So as I read this, we're focusing on study design. We're going to go to conduct under number 2. So as we provide our discussion and input for number 1, we're to be giving our input on the study design itself, including specifically the novel primary endpoint.

As you think about the study design and specifically about the novel primary endpoint, if

you would like to comment, as you seen the design, on strengths and weaknesses to begin the discussion -- I'm sorry. So anything that we need to clarify, what that's about.

Dr. Rimal, would you like to be our first commenter?

DR. RIMAL: [Inaudible - off mic].

DR. PARKER: Okay. I thought what we'd do is we'd just go around the table and offer everyone an opportunity to offer any comments that they have. That way, anyone who has a comment is given an opportunity, and like I said, if someone has already offered one and you're in alignment, you can just say that.

Dr. Conley, if you'd like to begin.

DR. CONLEY: Sure. Rob Conley. I'd like to comment on the design in that I do understand that it is a challenge in our field to be able to measure literally new things, and yet we don't have an intimate amount of fullness at times.

What I see from the design that I think is important to understand is I think the agency and

the sponsor did their best to design an outcome study and conducted that study. We'll talk about the method of conducting it later.

But at this time, one of the concerns I have for the committee to consider, even in critiquing it, is that the fullness of time isn't there forever. We talk about a not validated outcome measure. I think you also have to think about what is a reasonable outcome measure and is the NAEI a reasonable outcome measures.

Validation takes a long time, and in this population, I think that is just not there. So I think the agency and sponsors did the best job they could, from what I can see, to develop the study, which is my comment on the design alone.

DR. PARKER: Okay.

DR. EMERSON: Scott Emerson. I felt that the design, such as it was, was generally fairly good. In terms of the composite endpoint, and this is probably just a statement in retrospect and as the conversation has proceeded, it seems that the composite endpoint is driven a lot by components

that weren't as much of interest from the anecdotal reports in the adverse events.

The idea that the suicidal ideation was not a very big part of the composite endpoint at the end, and agitation, I guess, was a fairly common aspect, but not in terms of the severity. That is my only fear in this, I think being driven by retrospect rather than beforehand.

DR. MARDER: My concern regarding strengths and weaknesses of the data is that the key question of whether or not people with psychotic illnesses, individuals, where there is a plausible mechanism by which varenicline could make them worse, that they seem to be underrepresented. And it seems to me like it's not that strong a database for looking at these kind of relatively rare events.

DR. PARKER: Dr. Morgan, please state your name for the record.

DR. MARDER: That was Dr. Marder.

DR. MORGAN: And this is Glen Morgan. Given the purposes of the study, I thought the design was reasonably strong.

DR. MORRATO: Elaine Morrato. I also agree that the study design was strong. I like the margin of error that was aiming for the 1 to 3 percent range.

I like the -- it hasn't been mentioned -- the independent data monitoring committee and a real attention to trying to ascertain the adverse events in terms of the solicitation breadth, probing and the training of the investigators collecting.

DR. PERKINS: Ken Perkins. I really have very little to add. I also thought it was generally a strong design, and the sample size is pretty substantial given that there has been nothing in the literature to date.

Although some of the psychiatric issues or the population included might have been less than desired, I still think that it was substantial to identify whether or not there really was a significant risk, as it was designed.

DR. HERNANDEZ-DIAZ: Sonia Hernandez-Dias. I also think that the process worked, that we had

some adverse events reporting cases. And we know that they can provide signals, but they have limitations; like, for example, that we cannot know whether it is a medication or the reason why a medication is being used. And we have, also, the original studies that are challenged by confounding and other important biases in this specific case.

We needed a randomized clinical trial, and I think it's the best evidence we have, with some limitations. For example, we won't have the power to look at things like suicidal attempts. It is also not a real-world situation. We have more counseling and more probably supervision of the patients.

They both had some limitations, as we have been discussing, but both the company and the FDA have done sensitivity analysis and have beat the horse to death, and the results seem to be very robust no matter what you do, and the conclusions and this analysis. I think that is a good thing.

I think we can learn some lessons from the whole experience. The design of the study focused

on efficacy and collected the adverse events reports rather than, I think, focusing on the safety of the main outcome in the sense that we typically select hard outcomes for efficacy when we want to study it. And I think if we were to go back, probably we would have selected a higher outcome, like hospitalization, as has been proposed. That would, of course, required larger sample sizes and not little sample sizes.

Finally, I think the multicenter design of the study is always good, and we look for generalizability, but sometimes it has been generalizable globally and competes with internal validity and the difficulties of maintaining the standards across many centers around the world. But overall, I think it was a very helpful study that we needed to have, and now we have the evidence.

MS. HIGGINS: Jennifer Higgins. I concur. I think it was a very strong study design. I had some trouble with the NAEI assessment composite tool. I feel like it could have been tested a

little bit more for validity purposes. And I think there are some other methodological flaws or challenges, which I'll get into later.

MS. GILLESPIE: Terry Gillespie. No comment.

CAPT BUDNITZ: Dan Budnitz. Some of the obvious strengths of the randomized trial are the blinding and randomization to address the bias and channeling of the observational studies.

In terms of weakness, as other folks said, this non-validated outcome for NPS adverse events, because it is non-validated and does not seem to be focused on the particularly unusual adverse events that might be of concern, I think it's challenging to use and interpret in a single study.

There was potentially an opportunity to compare some of the findings or these events identified by the instrument to events identified other ways. That was not done, and it appears that opportunity may be lost for how data was collected in the design.

Then, finally, it does not appear to be

powered with the appropriate sample size to detect rare adverse events, like suicide attempts.

DR. WINTERSTEIN: Almut Winterstein. I agree with the previous speakers about the strength of the study. It is an important study, there is no doubt. It's large, and the randomization and blinding were certainly good.

I think what is important to recognize is that the study was not powered to rule out whether varenicline can increase the risk for suicide, suicidal ideations, psychosis, aggression, what have you.

What it can rule out is that there is not more than a 50 percent increase in the risk of those ascertained -- conglomerate of ascertained adverse events. And what exactly that is, is obviously a little bit difficult to interpret given the problems in the ascertainment and the definitions that we have discussed.

I'll stop here, because the rest is interpretation.

DR. GERHARD: Tobias Gerhard. I agree with

many of the comments that have already been made. One point in terms of strength that I want to point out is that this really shows the wisdom of FDA, and I guess also the previous advisory committee, in making sure -- requesting the RCT to be done, and then not making a decision based on observational studies when thinking about looking at trying to rule out a risk of neuropsychiatric and adverse effects with observational studies, which is very problematic, versus a clinical trial.

We see exactly what we would have expected, that we have much higher rates of the adverse events, even given all the issues with the outcome measures in the trial that are in the group with psychiatric history , about 5 percent, while there were between 0.5 and 1 percent in the observational study.

Again, I think there is a lot to be learned from this; that for some outcomes, some methods were better than others. And particularly, you are showing some of the limitations of observational studies to rule out concerns.

I think there is a lot to be learned regarding the outcome measures. If there would be another trial of a different product, trying to evaluate neuropsychiatric adverse effects, this particular instrument would be used probably.

One thing to make sure is to get much more detail on the individual reports, individual vignettes, making sure that that is collected. A lot of that came out in the comments of FDA. Again, I want to leave it with that for now.

DR. PARKER: Ruth Parker. I would agree with the comments that the design overall was quite good. I see both sides of multicenter. I think multicenter is incredibly important, multinational in so many different places and languages.

I'm not sure what it means to be linguistically validated in 15 languages, and how robust that really is in terms of its performance. And I think trying to hold quality control over that many sites and that many investigators is a huge challenge, and the results are only as good as the data.

I agree with the comment about the non-validated outcome measure for the neuropsychiatric symptoms, and also about the study being underpowered for less common outcomes related to imported events, like suicide, suicide attempts.

DR. NARENDRAN: Raj Narendran. I'm still trying to -- I just could never really get a good answer on -- if your whole idea is to look at the safety of these compounds, you already are enrolling. At least 15 to 20 percent of the people have been exposed already to these drugs. It seems like they should have been excluded, and probably in terms of safety, you'd probably want to remove them and see if there are any differences. That I think is a weakness.

I also have concerns about the NAEI and the way it was used. It didn't seem like it was used in the spirit of how it was supposed to be used, and the narratives are missing, which raises some concerns, as well. But overall, I think the study otherwise was well conceptualized, but I do have concerns with how well it was executed.

MR. PICKAR: I agree with much of the discussion and Dr. Parker's comments included. I think it was a heck of a study to do. It's far from perfect, but I'm just glad I didn't have to have the charge to have to carry that one out.

DR. FIEDOROWICZ: I won't re-echo all the comments, but Robert Conley had mentioned that this is a difficult task to validate an outcome when it's such a complicated outcome to capture and there is limited time to do it.

In spite of that, though, there is I think great risk that we all sort of highlighted of under-ascertainment of outcome. Even related to Pfizer's initial estimates, they observed less outcomes than were expected. And I would share the concern that that may bias the results to the null.

I am somewhat reassured by the fact that we also have the Columbia and the HADS so that we can measure anxiety and depression and suicide risk by those self-report measures that don't t rely on the investigator to identify events, especially in this case, where we have reason to be concerned whether

that was identified.

It may have also been nice to perhaps have anger and sleep measures, since that highlighted some of the concerns from before. But I think that the results from both the NAIE and the Columbia and the HADS were fairly consistent with each other.

DR. ROUMIE: Christianne Roumie. I really agree with a lot of what's been said. I felt like there were -- that efficacy part of the trial design was very strong and used objective measures of quit rates.

I feel like the safety part took some -- there was a lot more looseness to that, and a lot was left to the site investigator rather than a central adjudication process, which could have been used if case reports had been collected and often is used in large clinical trials.

DR. RIMAL: I have two thoughts. First, I think there is a mismatch between the objective of the overall effort, on the one hand, with what the objective of the study actually was, on the other hand.

I think the overall effort, the question is should the warning be removed. The overall study design asked the question, is one of these drugs better than placebo. Those two things, to me, are not the same thing.

Yes, they do work, the three drugs do work in increasing quitting rates, but this very expensive trial, in my opinion, does not address the key question, which is does the presence of the particular warning label reduce the use of the particular pharmacological treatment. That to me is the critical question, and this trial comes nowhere close to answering that question.

The second issue I think is very much in line with what Dr. Parker said, which is I wonder what does a warning label like this mean in the U.S. conceptually, and how is that different in a culture where such warning labels perhaps are not as prevalent, like Bulgaria. So the linguistic equivalence, I am very doubtful about that.

MR. HENNESSEY: Sean Hennessey. I think the strengths of the study were the randomization of

the three treatment groups with blinding, the two equally sized strata, those with and without mental health conditions.

I think that the outcome in the NPS represented the best thinking at the time, both on the part of the sponsor and the agency. And basically, the pilot study of the NPS was an 8,000-patient multinational randomized trial, and not surprisingly, we've learned a lot about the outcome as a result of that pilot study.

DR. PARKER: We forgot Dr. Besco on the phone. Dr. Besco?

DR. BESCO: Thank you. I also agree with the comments made about the strength of the study, especially when you compare it to available published articles. And I also agree with some of the earlier comments made about applying some of the learnings just from this experience to study other medications where we have observational reports received by FDA, that have received concerns

But like others, I do have some concerns

about the powering of the study and the validity of the measurement tool, and the interrater agreement between the investigators.

DR. PARKER: The daunting task begins. Let me see if I can summarize what I believe we have said here in response to the discussion, and then the agency can let me know if we have adequately addressed what you'd like for us to have commented on here in this discussion.

Regarding the strengths and weaknesses of the design, overall fairly good to quite reasonable, to good and strong, definitely a nod to going beyond case reports and the observational studies and the input that was available under those prior to the postmarketing research having been done; the need and what we learned from doing randomized controlled trials providing us with now what appears to be the best evidence that we have about it; the importance and the robustness of the sensitivity analyses that have been done.

The strengths and the concerns related to the multicenter design, that obviously being a

strength on many hands, but also raising some concerns, on another hand, about quality control and conduction across multiple sites with many investigators; linguistic issues related to the instruments, et cetera.

Randomization and blinding, again, being strengths related to the design; that the efficacy components, in particular, being quite strong. Then regarding the weaknesses, comments about the underrepresentation of the population that has the most severe psychiatric illness; that the notion regarding the influence -- the key question, which relates to the black box warning and its influence and what happens when it is there and when it isn't there being a central, I would say, concern. Do we know the answer to that based on what we have from the study that has been done?

Linguistic equivalence across many languages being a potential weakness in design; lack of power to look at suicidal events; that safety indeed probably does deserve a harder outcome than that which was garnered here, requiring a larger sample

size. Several comments related to there not being enough power to look at rare events, like suicide, suicide ideation, and that perhaps safety deserves, indeed, more robust outcomes in the study design.

Concern that 15 to 20 percent of the enrollees were exposed to the drugs that were actually under investigation for safety in the study; also, a comment about the large reliance on many different site investigators and how hard it is to control that many people that are actually providing data on what you're seeing and finding.

Specifically, regarding the question about the novel primary endpoint, there were a couple of comments about validity of measures, specifically, the NAEI, with a desire to have had more validity testing and understanding about the robustness of that as a tool.

It was designed to have narrative as a complementary component, and maybe in design it was a good idea. We can talk later about how the conduct reflected the intended design.

Non-validated outcome measures for the

neuropsychiatric symptoms, that being a concern about that outcome measure. However, it was also noted that the outcome for the neuropsychiatric symptoms reflected the best thinking at the time. And looking back, there's this retrospectively-driven way to look at that endpoint, and that's what we're left with at this point.

I'm sure I have missed some things. I hope I didn't add too many things that no one thought of or said.

Let me ask the agency before we turn to the next point of discussion, if we have adequately given you the kind of input you wanted regarding that topic.

DR. HERTZ: Yes. Thanks. I think that was what we were after.

DR. PARKER: Good job, team.

Let's turn to topic number 2. Discuss the potential impact of the variability in data collection, adverse event coding, and case definition on the primary endpoint. Because of this variability, discuss which analysis, or

analyses, and results most appropriately describe the effect of the smoking cessation therapies on neuropsychiatric events.

Let me just ask if there are clarifications for that topic before we, again, go around and offer points for discussion here. What questions dose anyone have about what we're being asked to discuss here?

(No response.)

DR. PARKER: It must be crystal clear. Why don't we start on this side? Dr. Hennessey?

DR. HENNESSEY: So we're going to do that?

DR. PARKER: You didn't know how lucky you guys were going to be. Here we go.

DR. HENNESEY: Sean Hennessey. I'm going to answer the second part of the question about which analyses we should pay most attention to. I think that some of the expanded outcome definitions that were seen in the sensitivity analyses should be those that we pay attention to. Even in those, I didn't see much in the way of cause for concern for the safety of varenicline with regard to serious

adverse events, serious neuropsychiatric adverse events in people with mental health conditions.

The impact of variability in data collection, adverse event coding, and case definition, there's a lot that's been said, and I'm not sure all of it has stuck in my brain. I'm going to pass on that.

DR. RIMAL: I think my concern is with patients with prior mental health --

DR. PARKER: I'm sorry. Your name at the beginning.

DR. RIMAL: -- Rajiv Rimal; thank you -- those with prior mental health issues. I think this study was very underpowered to detect differences in that group.

DR. ROUMIE: I would agree, and I think -- this is Christianne Roumie. I would add that I think some of the sensitivity analyses that were done, which exclude certain sites, don't actually get at the fundamental underlying issue, which is there may have been a systematic underreporting of events.

I never heard an answer to my question about blinding because even though there was randomization and the randomization seemed to work, there did seem to be an underreporting of events that seemed more lopsided in certain exposures.

So I'm not 100 percent confident on the accuracy of the blinding and would have liked to see some data that showed that investigators truly -- it was like a coin flip.

Ruth, can the sponsor respond to me?

DR. PARKER: Yes. I will quit curiously writing. Yes.

DR. RUSNAK: The direct answer to your question is, no, we did not ask investigators to guess at the treatment allocation. However, some data that we did collect within the study population, particularly if you just look at the overall population in panel 3, gives some data regarding the blinding.

Again, this a triple-dummy, blinded in both oral agents, as well as a patch. And if investigators had any inclination as to what

treatment allocation group they would be assigned to, you would likely see some substantial variation in the differences between the all-cause treatment and the treatment related adverse event reporting. And we didn't see that in the overall population or in the psychiatric cohort or the non-psychiatric cohorts. It was pretty flat between each one of these treatment groups, indicating that the blind was strong.

The same data also addresses the difficulty of ascribing causality in the postmarket reports.

DR. ROUMIE: Okay. I think, again, that while there were a lot of very suitable hard endpoints that were collected in the efficacy realm, there was more variation and variability for the neuropsychiatric outcomes, which is really what we're here to talk about.

I think we could have seen some other sensitivity analysis regarding how many more events would need to have occurred to tip results one way or another. But if you're telling me that it would have had to occur in 20 percent more varenicline

versus placebo patients to make the results positive, that gives me a little more confidence that there is less events noted in the varenicline group, and that truly this is by chance. There is truly no difference between the two.

DR. FIEDOROWICZ: This is Jess Fiedorowicz. As previously mentioned, related to concerns about ascertainment, I would weigh heavily the Columbia and HADS data when I review this data.

As far as the primary outcome, which was previously defined, it seems the negative binomial model of Dr. Andraca-Carrera perhaps best captures some of the heterogeneity that may be related to variability in ascertainment. And I share people's concerns about the internal validity of any specific sub-items or specific measures, such as irritability, given issues related to classification.

DR. PICKAR: First of all, the impact of variability in data is always to work against the statistical --

DR. PARKER: Dr. Pickar, can you please

state your name?

DR. PICKAR: Dr. Pickar. Dave Pickar. I think that's it. Generally, variability always works against your finding of statistical significance. It's our enemy and noise, as Dr. Winterstein appropriately said.

But what do I take away on this? Slide MD-65, can I do that? I'll tell you exactly what it is, for me. Sorry, I'm very personal about this.

DR. RUSNAK: Would you please project MD-65?

DR. PICKAR: In this, what you're looking at on the left is overall non-psychiatric and psychiatric. In the overall patient group, there is no difference between the incidence of these adverse events, between these different treatments, including placebo. And you don't talk statistics, but, in fact, if you go the risk difference slide, you'll see that there is absolutely no statistical difference.

Now, if you go the non-psychiatric patients, the middle group, a fascinating thing here. There

is a statistical difference there, whether they talked about it or not. For whatever reason, Chantix was less provocative of adverse events in the non-psychiatric population, and in the psychiatric population, much higher overall.

So the take-home message, overall, not much in the non-psychiatric -- your Chantix may be doing something that is beneficial, probably reducing withdrawal in the placebo, compared to the placebo group, and the psychiatric patients are particularly at risk on this.

If you go to just the last one, MD-67, when you look at this with just risk analysis --

DR. RUSNAK: MD-67, please.

DR. PICKAR: There you go. I know we're not talking statistics, but if you did, I believe -- and please, statisticians here, correct me -- the signal on the very top would be statistically significant.

On the other one, nothing else is, including, interestingly enough, the Chantix or the Wellbutrin compared to placebo, probably related to

the variability issue, that this is noisy data, so it is tough to get a statistical significance. Numerically, it's worse. But the one statistical significance is, for whatever reason, in the non-psychiatric people, the Chantix people experienced less adverse events.

To me, that's the take-home message, and the noisy study works against statistical findings, but it was a tough one to do.

DR. PARKER: If we can just put the discussion question back up just to keep us reminding ourselves, that we really also want to make sure that we focus and put our opinions around the conduct of the study here.

Dr. Besco, we have come to you.

DR. BESCO: Actually, my comments have already been expressed, so you can pass on any comments from me.

DR. NARENDRAN: Raj Narendran. I do have concerns about the way the data was collected and the variability that comes from it. I think the sensitivity analysis clarified somewhat, but I'm

still concerned that if you look at the data that's being reported to the FDA and being collected in your database, the study didn't really capture that very well.

So for some reason, I have less faith in the primary endpoint that was derived from this trial based on numerous other problems that were found in the review.

DR. PARKER: Ruth Parker. I, too, have some concerns about the data being quite noisy, data collection, some coding variability.

Specifically, I didn't hear it come up, and it struck me. I don't really know what it means, but I know in the FDA background, on page 45, there was a comment, "Office of Scientific investigation inspections of several of these sites had been requested, as well as inspection of other sites, in which similar issues were listed among the protocol violations, and the results of these inspections are pending at this time."

I don't know if there is a comment from the FDA regarding that. That seemed to be noteworthy,

to me. At least it made it into the briefing document. So when I see something like that, I scratch my head kind of hard, something that is getting investigated.

DR. HERTZ: We often will do some inspections, so part of that is the regular review in an application. We did a few more in this case to explore further some possible issues. We have some preliminary reports back, but we don't have any of the final reports from the inspectors yet. It's still an ongoing process.

We haven't been alerted to anything that we needed to describe at this point, so we are just going to have to wait for that final report to come in. But so far at least, nothing preliminary has identified additional problems.

DR. PARKER: I would just underscore the other concerns that have been raised about the sample size and going halfway, and then figuring out how many are we doing and why. I got a little lost in that.

There is just something about the idea that

this happened in 19 countries, and you're talking about whether or not you feel abnormal, and there are so few people -- I mean, if I ask in this room how many people feel abnormal today, I bet there'd be some people who do.

I don't know. There is just something about the validity of the measures and their meaning, and how I interpret that, and whether or not we're capturing all we can. I can't exactly articulate it better than that.

DR. GERHARD: Tobias Gerhard, Rutgers. Regarding the first question, the potential impact of the variability and the coding issues, case definition and so on, I don't think that's a major concern. I think that's somewhat supported by the fact that these different sensitivity analyses basically show pretty much the same thing.

What could be a potentially big issue, as pointed out by Dr. Roumie, is if there is really substantial underreporting across the board, or even worse, if it's a differential.

I would think that although that's a bit of

a concern overall, that at least for some of the more severe events, like hospitalizations, that seems unlikely to have occurred, and I think that's somewhat reassuring here.

Moving on to the question of which analysis is the most appropriate, again, I would less focus on which of the analyses, because they actually say pretty much similar things, from my perspective, but I think what they all say in the group with psychiatric history is not that there is no risk, from my perspective.

I think the question of statistical significance here, in the context of a study that wasn't powered to show this difference, is really about estimation. And the best estimate here of the data is that there is a small difference between both Chantix and bupropion of about 1.5 events, plus/minus a little bit, in these events; not, however, in serious events.

I think that provides us the information to quantify, with a confidence interval, the potential risk difference, and allows us to put that in light

of the benefit and make a decision about risk-benefit, which then we can use to make our decision for the voting question and the consequences for the black box.

DR. WINTERSTEIN: My most compelling slide would be MD-79, if you could bring that up, please.

DR. RUSNAK: Just to clarify the slide that you are requesting, 7-9?.

DR. WINTERSTEIN: 7-9.

DR. RUSNAK: MD-79, please.

DR. WINTERSTEIN: Just to be a pest and talk about the noise, the FDA opened this meeting quoting two cases. The cases were aggression and suicide. When we are looking at the composite outcome, this is mainly driven by agitation. And I don't know whether this is agitation because of drug exposure or whether this is agitation because of nicotine withdrawal, what that is, and that is my noise.

This first agitation bar, if we think about the number of events that we have that are compiled with this in relationship to everything else, this

is what drives the analysis, unfortunately, here.

If I interpret this correctly, aggression, we see we are comparing 3 cases -- or 2 cases to 2 cases to 1 case, and we have already seen the subanalysis for suicidal ideation. To me, that explains the noise in the analysis that makes it too hard to interpret this endpoint.

DR. PARKER: Please state your name for the record.

DR. WINTERSTEIN: Almut Winterstein.

CAPT BUDNITZ: Dan Budnitz. I'm not certain which analysis is most appropriate, because I think the variability for me is what others have expressed concern on, is an underreporting, or potential underreporting, of adverse events brought about by this investigator deeming what is an adverse event.

Maybe this is more particular to the outcome of it being a vague neuropsychiatric outcome as opposed to a very hard outcome of a cholesterol level or something, where there is no deeming involved, or an adverse event that is able to be

diagnosed by CT scan for stroke or something.

But when you don't know what you're missing, like those cases that we heard about, overdoses that are not deemed to be an adverse event, there is no statistical approach that can address that.

MS. GILLESPIE: I'm Terry Gillespie. I agree with most of the people here. It seems to me, after listening to everyone, that the coding was based on individual interpretation rather than clinical hard data.

MS. HIGGINS: Jennifer Higgins. With respect to the data collection, I just really can't get past the heterogeneity and the coding issues.

Regarding safety, I don't think I ever -- I think that David asked for total number of hospitalizations for both populations, both cohorts, and I didn't see that. I don't know if it's possible to ask for that now. We saw the psychiatric cohort only.

DR. RUSNAK: I'm sorry. I didn't catch all of that question. Could you just recap it?

MS. HIGGINS: I'm just seeking a total

number of hospitalizations for both cohorts, not just the psychiatric cohort, which is all we've seen thus far.

DR. RUSNAK: Do we have that data? Yes. Would you please project that?

The slide shows subjects with any adverse events leading to hospitalization in the U.S.

Do we have a total slide?

We apparently don't have the non-site data readily available.

DR. WINTERSTEIN: Are these the reported adverse events that led to hospitalization, or all these all hospitalizations that were then attributed to an adverse event?

Were these all hospitalizations with a principal diagnosis of some type of mental disorder or symptom, related symptom, or are these the reported ones that were collected that actually led to hospitalization?

DR. RUSNAK: Could you re-project the slide, please? This is subjects with any adverse event leading to any hospitalization.

DR. WINTERSTEIN: My question is where did you get the adverse event from. What came first? Did you collect all hospitalizations, and then look at which ones were attributed to an adverse event, or if there's obviously adverse events that were collected, we already discussed ascertainment method, and then you followed-up to see whether these led to hospitalizations?

DR. RUSNAK: I believe that this is the data that comes out of the SAE reporting, and hospitalization is an SAE.

DR. WINTERSTEIN: That would mean this is not all hospitalizations.

DR. RUSNAK: Actually, I'll ask Dr. Russ to clarify this, please.

DR. RUSS: These are the hospitalizations for any type of adverse event. So this would be part of the serious adverse event definition that leads to hospitalization.

When such adverse events are reported, the hospitalization is part of what the investigator would indicate. But these are non-psychiatric and

psychiatric adverse events. This would be any hospitalization.

DR. WINTERSTEIN: That was your AE reporting. That is not the ascertainment. This is just the AE reporting that would be part of any RCT.

DR. RUSS: There are very, very strict rules for serious adverse events. When adverse events lead to hospitalizations, they are very carefully monitored, and that's part of that.

DR. RUSNAK: We also have a slide similar for ex-U.S. Would the panel like to see that data?

DR. PARKER: Let's go ahead.

DR. HERNANDEZ-DIAZ: Sonia Hernandez-Diaz. Regarding the case definition, I think that if we wanted to go after hospitalizations or aggression specifically, by putting many things together, then we could be missing the answer, because we will dissolve the case in the broader case definition.

However, if we had agreed with the case definition that was used for the data collection, then you got in the coding on the classification.

As has been said, we could be missing some, but I would be more worried about the specificity than about the sensitivity in the sense that we'd have more biased results.

I think we can assume that the missed classification would be non-differential across treatments in this randomized setting. But in order to increase the specificity, the company and the FDA run several sensitivity analyses to focus on those outcomes, with probably higher positive rate, and by doing that, results did not change.

In my conclusion regarding which results most appropriately describe the effect, I would take the primary outcome because that was the primary analysis, and then consider all the subanalyses and sensitivity analyses, and I think all of those thoughtful analyses did not move the conclusions.

I would say that I would take the whole analysis overall in order to interpret what are the best results, taking the primary plus, all the sensitivity analyses conducted. And again, the

conclusions didn't change after all that.

DR. PERKINS: Ken Perkins. I was just going to conclude the same thing in terms of the lack of bias by treatment condition in whatever variability there was. So I didn't have much to add at this point.

DR. MORRATO: Elaine Morrato. I'm going to comment on the impact of the variability from a slightly different angle, more on the impact of decision-making.

Given I think the precedent-setting nature of this discussion of a study, removing a boxed warning, for me, the impact of the variability effects my confidence in making the decision to remove completely or modify, especially in light of the thousands of case report findings that we've heard and so forth.

For me, I'm most concerned, I think, with not the overall group, but those with a psychiatric history, because the point estimates are trending higher and bordering on statistical significance.

So I know we got to see some sensitivity

analyses, but I would have preferred to have seen perhaps a little bit more systematic or robust. I felt like there is still some analyses in progress. It was mentioned that the sponsor had updated theirs a week ago.

We heard several ideas from committee members in terms of analyses of, as Dr. Narendran mentioned, removing those that were on the drug before, not looking at the sensitivity analysis of how much under-ascertainment would need to be occurring in the psychiatric cohort to see significance.

In the FDA's document also, they talked about there hadn't seemed to be a full synthesis of the different safety data sources in terms of the endpoints. So it leaves me, at least, a little at pause as to has that all really been completely wrapped together and adjudicated and looked at as robustly as it could be.

DR. MORGAN: Glen Morgan. I have nothing further to add.

DR. MARDER: This is Steve Marder. My only

concern is that I would have hoped that the primary endpoint would have defined the kind of moderate adverse events that one would expect in this population. And what keeps gnawing at me is why is the incidence of these events so low in the placebo group? I thought that people who were withdrawing from cigarettes felt crappy and they would have more -- so I agree with everything else that has been said, but just sort of a gnawing concern that the primary endpoint may not have been very sensitive.

DR. EMERSON: Scott Emerson. The potential impact of variability in data collection, adverse event coding, and case definition on the primary endpoint. Potential impact is huge. That's really what it comes down to.

Again, as I say, I think as it was designed, I think the plan was good. There are some problems about how it turned out that I have some problems with.

First off, I'll mention that I agree with Dr. Parker on this idea of dealing across countries

in psychiatric disease, different cultures, is very hard. My whopping eight weeks of psychiatry that I've had in my life, one of my professors remarked on the fact that diagnosing mania; what was mania.

He said in his native India, in post-colonial times, mania would be people -- their concept or delusions of grandeur was carrying an umbrella and speaking English. In Seattle, where I have lived now for quite a number of years, believe me, the people carrying umbrellas and speaking English are depressed.

I worry about how all of these things translate, and perhaps we're seeing some of that in the variable rates across the centers. Again, I focused, unlike Dr. Marder, on what was the placebo rate, and the sponsor spoke to, in their design, saying, we had no idea, so we made it up. So I'll account for that a little bit.

What I was going to say when I first read the things, of saying, well, you were expecting 7 percent and you got 5 percent. So is that a problem due to the underreporting? And realize

that if you underreport enough, the two groups just look identical at zero and zero.

So while I was all in favor of their primary analysis start out, this is what I would do. I would do a risk difference analysis. I'd like to judge the public health impact, but all of those zero-zero centers with a risk difference, we're treating that data as it's very real. And when you switch instead to a risk-ratio analysis, those sort of become non-informative; that is, the zero-zero, and you're not as much pretending that you have a lot of precision with that.

So when I was asking the FDA about why you were doing this, one of the reasons I would have gone to an analysis, in my exploratory analysis, for the risk ratio was in fact to say, well, if you've got some centers that are just reporting nothing -- because if their biases are that this is all okay -- in this study, unlike an efficacy study, you can make anything look safe if you say nobody has a problem, then we've got to worry about what happens with that.

If it's just individual centers, then you have to say, well, we can contaminate a study with non-response, we can contaminate a study with people where we can't possibly show the difference on either arm, and the risk ratio will stay fairly constant.

If the contamination is pure, that you won't have any response on either arm, in which case, that drives me a little bit more towards thinking about the risk ratio rather than the risk difference. As much as the sponsor said that they of course were not going to test any hypotheses, a number of them got up and interpreted lack of statistical significance as absolute equality.

So I'll channel Tom Fleming in saying, "Absence of evidence is not evidence of absence," and we need to be very careful in saying we know that it's equal because it's not statistically significant.

We don't know any such thing. And we should really, in a safety endpoint, be focusing a little bit more on the upper bound of the confidence

interval and what do we see. And that is somewhere a relative risk, I'm going to claim, of 1.8 to 2, using the negative binomial data that was precise or faking analyses using the risk difference.

We are still in that same target, with the point estimate at being somewhere in the 1.35 to 1.4 increased risk; not statistically significant, I'll grant you, but it's just saying that's what our estimate is. And if we're worried about safety, we don't live and die by the .025 one-sided significance.

So the impact can be large here in terms of what we're looking at. And I'm having a slight bent toward looking at the analysis that they've have done, relying on my belief that I didn't see anything big that made me fear the randomization was not good, that the blinding was not good; the missing data aspect that we're missing data on about 20 percent of the subjects, of which I believe it was around 10 percent of those subjects. It was missing data, that they dropped off the study during the treatment phase, not just the

treatment.

The bias, the large tails of this, the low numbers of events in a few sites was countered with extremely high events in the other site, and that's also the way the to lie with statistics, is just say everybody has an event. And if everybody has an event, then it also looks equal.

Well, it went up to 15 percent rates in some of those sites, and I don't really know what that would do. But recognize that if you want to mess up a relative risk, then throw in bias of noise on that. So it just makes it much more difficult to determine what happened.

So as I look at all of these analyses, living with the primary endpoint as it was originally defined, clearly, this study was never designed to really have high power to detect the most severe neuropsychiatric adverse events, but was instead trying to cast a slightly wider net on what some less severe events might have.

Using that endpoint, I'm struck by that 1.4 relative risk as an estimate that is with a

confidence interval that goes on up towards 2, and that is bothersome given the plausibility in that patient population of exacerbating an underlying condition.

Then it comes down to what do you believe the baseline rate is, and we'll have to discuss is a black box warning worth a relative risk of 2 not being ruled out.

DR. PARKER: And we're coming to that.

DR. EMERSON: Exactly. So it's just this concept of how to look at those events. But in the most severe events, we don't have enough power to really assess what those rates would be, and I'll comment more as we come to the observational data with respect to those.

DR. PARKER: Dr. Conley?

DR. CONLEY: Yes, thanks. I agree with a number of the comments that have happened here today. First was the study, Dr. Rimal, is it really designed to take off that black box warning as opposed to detecting what is a change in agitation.

I do think that the concern about can we really get the real number of hospitalizations actually is important. Why that is, is partly addressing this question of the -- in the data variability from the sites, a concern I have is that the FDA's presentation seemed pretty informal.

I recognize you felt like you had a lot of problems, but it would have been nice to understand the precision in some ways of those problems. How many cases were you having a hard time ascertaining, not so much are there just examples of there are a few cases where we can't figure out what's going on. I figure that's always true, but it was hard for me to still gauge exactly how bad this is or not or how accurate stuff is.

That said, I think what was done statistically made a lot of sense to me, a lot of the sensitivity analyses coming out more or less the same. I did have a bit of concern that it seemed at the end of the day, you said, well, there's a little more of a rate in the psychiatric group, which I think I can understand why, but you

underemphasize that you never did lose the lower rate in the non-psychiatric group.

So I do wonder why you weighted one more than the other, since they had about the same number in it.

It is helpful that the Columbia and the HADS were the same. I think that's important in understanding that probably you're not seeing anything here. But I do hear your last comment, that if you get a bunch of zeroes, it's always a worry.

So to me, the things that are hard to not detect, like hospitalizations and things like that, that being flat would be reassuring to me that there isn't a real difference, understanding that there is potentially problems with the sites in the study.

DR. PARKER: Dr. Hertz?

DR. HERTZ: I can't help myself, but feel compelled to comment on one thing you said. I don't really think that it's completely the responsibility of the agency to detect all of the

problems with the company's lack of data capture.

The fact that we have identified a problem is, I think, the relevant point here. I know how much work has gone into a literal page-by-page review of -- I'm not looking for a response. This is not a discussion. I'm just addressing the fact that our expectation is that a company will identify problems and bring that forward.

We attempt to anticipate problems and specifically request in advance certain things, which we did not get with this. This is not our first time asking this company for informative narratives.

This expectation of if two sites are found not to be providing the type of study conduct, that the inspection would go on further to look to see how broad it is, but we do our best to fill in when we have questions like this.

So we have done a substantial page-by-page analysis, and that's why these issues have been brought to light here. But it would be nice if not just presenting the best possible analysis, but we

could get to some of the weeds from the company itself, who clearly must have access to all of the potential issues not only that we identified, but presumably their own analyses detected at some point.

DR. PARKER: Let me offer a summary here, if I can. The variability of data collection -- this relates broadly to the conduct, and the first, design, this being a focus on the conduct.

The variability of data collection, the coding of adverse events; impact of variability, and that impact really is the issue becoming that of the black box warning removal, which is what we're moving toward; the potential impact of this variability in data collection; and the coding of adverse events.

Its potential impact is huge. There were comments about heterogeneity; comments about language, culture; validity of measures; variability across sites, quality control there; noise in the data, much of that relating to this notion of agitation and whether that's drug

exposure or nicotine withdrawal.

The noise itself is driving an endpoint and makes it incredibly hard to interpret the endpoint. Precision around outcomes or events like hospital admission, there being a lot of zeroes and hard to know what to do when there are a lot of zeroes; the NAEI; the lack of informative narratives, which were to be a part of an established protocol for the study. It was not to be used as a checklist, but anytime there was affirmative response, it was to be a narrative, and those narratives were not provided, which was a part of the intended study design.

Potential for misclassification; bias; again, sample size came out again. Could there be a systematic underreporting across of adverse events? If we don't know what's really missing, you can't really address it statistically, being a comment.

How accurate the blinding really was; the primary endpoint; why is the incidence so low in the placebo group, and a question about the face

validity of that; there not being more happening in the placebo group, who is going through nicotine withdrawal.

Which analyses? In general, the primary outcome first, the sensitivity analyses that were done did appear to line up with the results there and have similar findings.

I hope I didn't miss anything major in that. Can we do one more before we move? Is that okay? Everybody on board? Here we go. Let's do number 3. I'm sorry.

DR. WINCHELL: If I might, I think we can circle back to this when we talk about people's recommendations for labeling, but if people felt that one or another of the sensitivity analyses is a better representation of the overall findings, for expressing the results of the study, that's something we'd like to hear.

DR. PARKER: Let me remind us of that when we get to that, if we've lost that thought chain.

Let's move on to topic number 3. Discuss how you weigh the evidence -- but this relates to

the observational studies -- when evaluating the risk of serious neuropsychiatric adverse events in patients taking smoking cessation products.

This should focus specifically on the observational studies, and how you take in and use that evidence when you are evaluating the serious neuropsychiatric adverse events.

Any clarifications needed to that or can we start? All good?

(No response.)

DR. PARKER: So focused on the observational studies and impact. Remember, focus comments, all 19 of us, and if you don't have anything to add, it's fine to just say so. Thank you. We'll start with you, Dr. Conley.

DR. CONLEY: Rob Conley. I appreciated the analysis of the observational studies. To me, it does suggest sort of the same problem we were talking about in the trial, is that at one level, it was hard to really have correct ascertainment of the cases because of the level of data you have in observational studies, and that's an obvious thing,

important, though.

The thing that I was sort of missing in it is whether or not this can be supportive evidence. Here, there were some studies positive, some studies negative, most hovering around the middle. What I didn't come away with a clear understanding of is can I trust that to say I heard that, but at the end of the day, we'd like to have this study get done.

Well, it's done. So can we have a help with these observational studies? I still just have a question mark in my mind about that.

DR. EMERSON: This is Scott Emerson, and I don't weigh them much at all. The issue is that the biases that have been pointed out, particularly the time frame where there's going to be a channeling bias, the ascertainment bias.

The other aspect is just the definition of the endpoints that makes it very hard to compare, and particularly some of the studies that were showing a hazard ratio of .5 with a, roughly -- with a very, very narrow bound, if

you're looking at it and saying -- if we had seen that same data on an endpoint that we thought we could compare with the clinical trial, I would be holding this up and saying, yeah, this proves that the observational study is just completely worthless.

So I wasn't certain how much I should drop back because of that, where there's too many cases in which the observational studies show results counter to the clinical trials that were later done, and I just think that I trust the clinical trial here far more.

DR. MARDER: Steve Marder. I agree with Dr. Emerson. I have nothing to add.

DR. MORGAN: Glen Morgan. I'm kind of in the middle. I feel that the trial that was conducted and the clinical trials generally are stronger indications of the signal that we're looking for, but that doesn't mean that an observational study or a case report is without utility. I consider it all data that we should attend to.

DR. MORRATO: Elaine Morrato. I participated in the 2014 review in which we looked at the observational data. And at that time, I agreed with the committee members and the FDA's decision that observational studies weren't adequate to address the safety question, and the data was required.

Now, I believe some of the observational data was incorporated into labeling, to some degree. As we think about label changes and if we decide to added in the trial data, and if some of the trial data is inconsistent with now a wide variety of observational data, we may need to go back and relook at, well, which observational data do you now include.

While it is in sort of the matrix of all available data, I don't think we want labeling that has some data inconsistent with the overall warning message. So we can discuss that when we discuss if the trial data goes in and how and so forth.

DR. PERKINS: Ken Perkins. I don't have anything to add.

DR. HERNANDEZ-DIAZ: Sonia Hernandez-Diaz. I think that the observational studies are sometimes as good as clinical trials and can help in many occasions. They have larger sample sizes for some events and longer follow-up sometimes, and they represent the real-world evidence.

However, in this case, when the outcome that we are going after is some psychological, psychiatric, or even like a feeling kind of outcome, that of course is not going to be capturing claims databases, and that together with the room for confounding of -- if it is smoking cessation itself, what is affecting the outcome, et cetera, and that makes observational studies potentially biased because of this confounding. And actually, only those with active treatment would be close to an answer.

When I support observational studies, in this particular case, I think the best evidence is coming from the clinical trial.

MS. HIGGINS: Jennifer Higgins. I have nothing further to add.

MS. GILLESPIE: Terry Gillespie. I have nothing further to add.

CAPT. BUDNITZ: Dan Budnitz. No additional comments.

DR. WINTERSTEIN: Almut Winterstein. I agree with what Dr. Hernandez-Diaz just said.

DR. GERHARD: Tobias Gerhard. You'd be hard-pressed to find a bigger proponent of observational research, but in this case, I also think it's an application where the observational studies, as demonstrated or as shown in the thoughtful presentation by Dr. Pratt, really has severe limitations and really don't contribute much to what we have from the trial here.

DR. PARKER: Nothing to add. Ruth Parker.

DR. NARENDRAN: Raj Narendran. Nothing to add.

DR. PICKAR: Dave Pickar. Nothing to add.

DR. FIEDOROWICZ: Jess Fiedorowicz. Nothing to add.

DR. ROUMIE: Christianne Roumie. I have nothing to add.

DR. RIMAL: Rajiv Rimal. I also have nothing to add, except to say that I think it's a sequencing, that if that had come -- given that we now have an RCT, I don't think it adds much.

DR. HENNESSEY: Sean Hennessey. I'll just note that this is a difficult outcome to study using health care data, and I don't believe that the existing studies add much to our understanding.

DR. PARKER: Dr. Besco?

DR. BESCO: I agree also with the previous comments.

DR. PARKER: Let me give a quick summary here for the observational studies. Good job on that, team, by the way. That was really well done. Well done.

Though observational studies can be even as good as clinical trials and do have utility, especially given their size, there was a lot of discussion about the outcome related to psych outcomes not being well captured in claims data; comments about biases specifically related to channeling time frame in the observational studies;

definition of endpoints; more trust, therefore, in the clinical trial; the observational studies alone not being enough to adequately address safety; yet, the sequencing of the trials as they happen with the observationals preceding the clinical trials was not a bad idea.

Another comment that relates to how labeling changes, especially as we move toward addressing whether there will be any or need to be any, that they should reflect beyond the observational studies, since current labeling up to this point did not have the availability of the results from the clinical trial that we've discussed today.

How about we take a really only 10-minute break, and then we come back, and we're going to power through the rest of this. Thank you.

(Whereupon, at 3:44 p.m., a recess was taken.)

DR. PARKER: As we're all taking our seats, let me just let everyone know, we've got a couple more points specifically for discussion, and we do have our voting question. We will not be taking

time to go back into data and pull slides. We're going to stick very specifically to the task at hand and offer our advice on these discussion points rather than dig back down into the data points.

Okay. Topic number 4. Based on the results of the clinical trial and observational studies, discuss the impact of psychiatric history on the occurrence of neuropsychiatric adverse events during smoking cessation therapy. So the focus here is the impact of psychiatric history on the occurrence of the neuropsychiatric adverse events with smoking cessation therapy.

So true to our form before, we'll go around, and if you will offer your comments related to that. We'll start with you, Dr. Hennesey. Thank you.

DR. HENNESEY: Sean Hennesey. The absolute frequency of neuropsychiatric adverse events is not surprisingly higher in people with mental health conditions than people without mental health conditions. And in addition, the frequency of all

neuropsychiatric adverse events appears to be higher in the varenicline group compared with the placebo group. But when you look specifically at serious events, they do not appear to be more common in the varenicline group than the placebo group.

I hope that answers the question.

DR. RIMAL: Rajiv Rimal. We just, for lack of a better term, trashed observational studies. I think that given the study design, we don't have the requisite data to answer that question because, in my mind, to answer that question, we would need to randomly assign people who have history to these different arms and sufficiently powered to detect those two kinds of differences. And I don't know the answer to that question.

DR. ROUMIE: Christianne Roumie. I think, based on the data we've seen, there appears to be numerically more events among patients with a past psychiatric history. And I will leave it at that.

DR. FIEDOROWICZ: Jess Fiedorowicz. Persons with psychiatric disorders are a potentially

vulnerable at-risk population that is known to have a high prevalence of smoking in excess and dramatic burden of related morbidity and mortality, and to be undertreated with smoking cessation therapies.

Subsequently, we must be cautious and avoid overinterpreting the numerically higher incidence of neuropsychiatric effects observed on varenicline and bupropion than on placebo. We should require compelling evidence to separate this subgroup from the general population in terms of risk-benefit analyses.

With regard to the cohort with the psychiatric history, there was no evidence of a cohort by treatment interaction on outcome. Within the cohort with the psychiatric history, the numerically higher incidence of NPS events in the cohort of persons with psychiatric disorders was not demonstrated to be a non-chance finding. That is all I have to say.

DR. PICKAR: Dave Pickar. I think Jess has gotten that right, but I do come away pretty clearly in my mind that psychiatric populations are

particularly vulnerable to behavioral side effects as they are with virtually all drugs that interact with the CNS.

Fortunately, Chantix is a prescription drug, and doctors administer it and manage patients, and that's what people do when you manage psychiatric patients. So I think that part of the story is the clearest thing.

DR. PARKER: Okay. Let's get Dr. Besco.

DR. BESCO: Dr. Besco here. I also have some difficulty answering this one to speak on the underrepresentation of the patients with known psychiatric issues and the random controlled trial, and also based on issues with the pairing of the study to detect the rare events.

DR. NARENDRAN: Raj Narendran, and I agree with the previous speaker.

DR. PARKER: Ruth Parker. I agree as well.

DR. GERHARD: Tobias Gerhard. I agree with Dr. Hennesey's comments. I think there is some evidence. The study wasn't formally set up to detect an interaction and test the difference

between the groups. But I think the evidence -- also as shown on FDA's slide 29; I think that's the second presentation -- is that there is a higher incidence of events in the varenicline group, in the group with psychiatric history, but not in the group without. I think that difference is meaningful for the label.

DR. WINTERSTEIN: Almut Winterstein. Yes, if the confidence intervals of the presumably protective effect in the non-psychiatric group and the non-significant potentially increased risk in the psychiatric group, if those confidence intervals were compared, they probably would just barely touch. So there may actually be some statistical proof at least for an interaction or for a modifying effect.

But this said, again, given that the composite endpoint is composed of so many different components, for the non-psychiatric group, the types of outcomes that are reported, just looking at those distributions, are also a little bit different. So in the non-psychiatric group, that

is primarily agitation. In the non-psychiatric [unclear] group, agitation is still the leader, but then aggression, panic, mania, depression, anxiety have a much larger contribution as well.

So it's hard to compare those results because even though we're thinking we're looking at the same outcome, we actually don't. So I do think it makes sense to at least consider a warning that would say that this might be that the psychiatric population might be a more vulnerable population with respect to side effects.

CAPT BUDNITZ: Dan Budnitz, no additional comments.

MS. GILLESPE: Terry Gillespe, no additional comments.

MS. HIGGINS: Jennifer Higgins. I believe the data really show a propensity for increased neuropsychiatric issues among this population, psychiatric population cohort.

DR. HERNANDEZ-DIAZ: I think there were more events -- with psychiatric history.

DR. PERKINS: Ken Perkins. It does seem to

be a trend, but I was convinced that it was not statistically significant. So that's my comment.

DR. MORRATO: Elaine Morrato. I would just add that I found it interesting in the FDA slide 7 and 8, I believe, when it was the Kaplan-Meier curve looking at the cumulative incidence of events, that in the psychiatric -- those were the psychiatric history, while it leveled off, was much steeper, continuing out through the whole 120-140 days, whereas those without the psychiatric history kind of leveled off much sooner, closer to 40 days.

So we didn't really discuss the time clustering of events in the trial and whether or not it was similar or not to case reports, which seemed to be around 14 days, but that may be something worth looking into.

DR. MORGAN: Glen Morgan, no further comment.

DR. MARDER: Steve Marder. As I look at it, I see a higher risk of these adverse events in the psychiatric group. The thing is, most psychiatrists don't see the world of psychiatric

patients as a single group; they see them as people with individual illnesses. And it would be useful in the future to take that data -- you know, that comes more risk related to what illness they have.

DR. EMERSON: Scott Emerson. I'll just note that when I am studying drugs that I think might be renotoxic, I naturally separate the population into those who are already renal compromised and those who aren't. So in that same token, I would do this here.

I would not personally recommend to anybody that they power a study to detect that interaction, possible interaction, statistically. That takes a sample size 4 times greater than a single study, whereas I can answer in each subgroup separately, maintaining my type 1 error with about 2.3 times as many subjects. So answering it in each subgroup is important. I think it's not significant trends towards acting differently in those subjects with prior history.

DR. CONLEY: Rob Conley. I would agree with what I've heard from the group that I did certainly

see a higher rate of psychiatric events in those who had prior psychiatric illness, but I really didn't see compelling evidence that there was a difference in the non-psychiatric versus psychiatric ones. I mean, it does look like there might be a trend there -- I'd agree with my earlier commenters -- but I wasn't sure that that was really enough of a separation to know what that meant.

DR. PARKER: So a summary here; a trend to an increased risk for adverse events among those with neuropsychiatric diagnoses, that it would also be helpful to have data more specific to which psychiatric diagnosis as we think through this a little more.

Also, other comments that we don't actually have enough data to answer this completely, noting that there are more events among patients with a past prior psych history, and also comments that indeed those with a psychiatric history are more likely to be smokers, more likely to have neuropsychiatric symptoms and other comorbidities

as well.

One other comment related to the underrepresentation of patients with neuropsychiatric symptoms overall, and another comment that I believe probably again relates to the agitation noise -- I will call it -- warning for the psychiatric population, that they may indeed be more vulnerable to side effects, and that being something for consideration as we move toward thinking specifically about the labels and the content that is on the labels.

Let me just ask the agency, is this adequate for what you were looking for?

(Dr. Hertz nods in the affirmative.)

DR. PARKER: Great. Okay.

So we will move now to question 5, which is the only voting question that will be before us today. We have question 5, and then in question 6, we'll be looking for the rationale for the answer that is provided in the voting question number 5. And it will be at that time that we'll look at any additional labeling actions that we would advise

the agency considers, they think about this broadly. But we'll first go to the voting question here.

It looks like we've got some clarification about what all this is about. That doesn't surprise me. So let's start with that before we move to the specifics about the voting.

Yes, Tobias?

DR. GERHARD: Not so much about what it is about, but more about the process. So if I understand correctly, we'll vote on this, and then justify, go around, which is question 6. So there's just one round. That's what I would propose, otherwise, there will be a lot of duplication.

DR. PARKER: That would be fine because -- that's fine. The specifics around the vote relate to the black box. And as you know, labeling is more than just the black box. It was my understanding that the agency was asking us to look at more than just the black box. If all you want us to specify is this about the black box

warning, that certainly would relate to the justification for the vote in number 5.

DR. GERHARD: So are you planning to do one more round of comments or two rounds?

DR. PARKER: Let me ask the agency what they're looking for in their input, but I have a feeling they want to know more than what we're going to get at just with the black box. Let's just ask them, but I hear what you're saying.

DR. GERHARD: I agree, but I think it all fits in one justification. Going around once and concluding everything that one has to say about the removal of the black box and suggestions for the label might be most efficient.

DR. HERTZ: This is Sharon Hertz. That's certainly an acceptable option, as you go around and have people say their vote for the record and their reason for the vote. If you want to ask for any additional labeling comments in that same round, that certainly is an acceptable option.

DR. PARKER: A couple of other clarifications.

That's fine. So let's plan that what we'll be doing then is we'll vote, and then we'll go around once, and we will explain why we voted the way we did, the rationale for that; and at that time also address any additional labeling action so that we go around once. I'm all about efficiency. But it looks like there are some questions related to that, so let's make sure everybody's clear because this will be the last time we all go around.

Yes, Dr. Winterstein?

DR. WINTERSTEIN: I actually have a question to the FDA, and it's a history question. There have certainly been occasions that I remember where spontaneous reports were sufficient evidence to create a black box warning in the past. Usually then, there's also some biological pathway, and the pharmacology makes sense, so there's plausibility in some way or fashion.

From what I understand, the decision to put a black box warning in place was based on this spontaneous report, right? There was no other

evidence at that point. I'm just curious. If you just could recall the evidence that was available then for us, that actually triggered the black box warning at that point, that would be helpful to me.

DR. RACOOSIN: Judy Racoosin. It was the review of spontaneous reports that had been submitted to FDA, to the FDA adverse event reporting system, as well as reports that had been submitted to the sponsors, and that were then submitted to FDA. So yes, based on case reports.

DR. PARKER: Dr. Budnitz?

CAPT BUDNITZ: This is a question to the agency as well. We saw some of the ordinary situations leading to black box warnings, but we didn't hear any precedents or considerations for removing or changed boxed warnings. Can the agency add any insights in regard to that?

DR. RACOOSIN: So there are relatively few examples, and none of them are particularly contributory to where we are, what we're considering today. The one that might be the most relevant but is not terribly relevant is

ambrisentan, which is the second in a class of treatments for pulmonary artery hypertension. The first drug that was approved in that group was bosentan, and there is a signal for hepatotoxicity and teratogenicity.

Bosentan had a boxed warning for both of those, and then when ambrisentan was approved, it also got a boxed warning for both hepatotoxicity and teratogenicity. The data supporting the hepatotoxicity boxed warning was not terribly robust at the time, but then there was the consideration that, well, maybe this is a class effect, and so it originally got that boxed warning and approval.

Subsequently, over time, the sponsor collected information through a variety of data streams that did not bear out the hepatotoxicity risk with ambrisentan, so eventually the boxed warning for hepatotoxicity was removed. So it was not based on -- it was based on a number of streams of postmarketing data that had come through that the sponsor collected.

The other examples I don't think are particularly relevant. There was one where inhaled corticosteroids had a boxed warning related to concerns about patients being treated or having to be -- when inhaled corticosteroids were first approved, there was a boxed warning about the risk of adrenal insufficiency. There were concerns about that. Eventually, after 20 years, that box got removed because the practice of medicine, people understood how to use inhaled corticosteroids. I think that's about it.

CAPT BUDNITZ: So this could be potentially precedent-setting if this box was removed beyond this particular case.

DR. HERTZ: Yes. So I think that if an analogous situation were to occur where case reports and spontaneous reporting supported a box, and a subsequent study suggested the signal might not support it, people would refer to this. But I don't know that -- I mean, it's all going to be the devil in the individual details of the strength of the risk in the studies.

DR. PARKER: If I could just also ask the agency to clarify. I know in our comments, as we go around, we'll be mentioning, discussing any additional labeling actions. And you mentioned previously this relates to the box, the warning, precautions, the med guides, potentially REMS, because I assume this is across all of the smoking cessation products. It's not just the black box warning here.

So maybe it's important to make sure everybody understands, as we provide comments, we're also providing comments not just about the black box warning for the varenicline, but I assume we would be looking at these other potential components of the warnings, precautions, med guides, REMS, for other products, including the other ones that have been discussed today and whether or not there's input regarding that.

DR. HERTZ: Yes, that would be true for the bupropion for the smoking cessation indication. But the over-the-counter products don't carry those types of warnings, so it's a different

consideration there. But certainly for varenicline and bupropion, yes, all of the traditional prescription labeling options, the warning, any other labeling within the full prescribing information, the medication guide.

DR. PARKER: Was the request also from GSK about the REMS or about any particular components that you want addressed here?

DR. HERTZ: Well, the REMS issue in this case, this was just the MedGuide REMS. So I don't know that there's a lot of the REMS discussion per se. I think it might be most helpful if the discussion was more about the medication guide, and then we'll take are of that connection with the existing REMS.

DR. PARKER: Let's go to the phone. Dr. Besco?

DR. BESCO: I'm so sorry. I forgot to unmute myself. Kelly Besco. Actually, my question was about the precedent of removing a black box warning. So that's already been answered, so I don't need to ask a question.

DR. PARKER: Dr. Morgan?

DR. MORGAN: It was on an earlier slide. The criteria for having a black box is that there's a high risk of a serious adverse event. Is that correct? Or is it when there is a unknown yet very serious event? I'm trying to get a sense of what are the criteria to decide. It's a little difficult to say black box/no black box, or change the black box.

DR. RACOOSIN: Slide 21.

DR. MORGAN: They've got it up now. Great. Thank you.

DR. PARKER: Dr. Morrato?

DR. MORRATO: I thought it would be helpful maybe if you can share with us a little of the thought process that FDA uses. So when it's something that is so serious in proportion, how do you -- and maybe this is what you want us to comment on. But in general, how do you approach that? Can it be so serious in terms of magnitude, severity, life-threatening; so serious in terms of frequency versus rare; so serious relative to the

benefit, so we really should be looking at serious in a benefit-risk way?

DR. HERTZ: I would say that it's not so much an absolute frequency. It's more the severity aspect of it and the importance of making it prominent in the risk-benefit considerations for the product where the risk might outweigh the benefit, or the risk can be mitigated in a certain way. So that's kind of what this is trying to get at.

I do want to remind everyone that the presence of a box or the removal of the box doesn't negate that there are other warnings in labels. So section 5, which is our standard warnings and precautions, would also still have information. Decisions about the box would not change the decision of having a warning necessarily, unless we actually had no reason to consider any longer that signal.

DR. MORRATO: So does the box make a difference on the ordering of warning information in a med guide? That's what's being proposed, is

that whole section that starts the current patient med guide around psychiatric be deleted.

DR. HERTZ: The ordering in section 5 is reflected. So even if there's not a box -- so typically if something rises to a box, it will be higher in section 5, but in the absence of a box, the ordering in section 5 is meant to reflect some degree of concern.

DR. PARKER: Okay. We're going to be using an electronic voting system for the meeting. Once we begin the vote, the buttons will start flashing, and will continue to flash even after you've entered your vote. Please press the button firmly that corresponds to your vote. If you're unsure of your vote or you wish to change your vote, you may press --

DR. FIEDOROWICZ: I had a question.

DR. PARKER: Oh, I'm sorry. I do apologize. I didn't see it. Yes?

DR. FIEDOROWICZ: I just wanted to clarify, we mentioned a precedent, but is there any reason that we wouldn't apply the same criteria for the

black box, to overturn the black box as to initiating it? This is a question for the FDA.

DR. HERTZ: No. We showed the criteria for a box because if you feel that the data support any of those criteria, please feel free to let us know that.

DR. PARKER: Let me check. Any other questions here, before we do the vote?

(No response.)

DR. PARKER: We will be using an electronic voting system, as I said. Once we begin the vote, the buttons will start flashing and will continue to flash even after you've entered your vote. Please press the button firmly that corresponds to your vote. If you are unsure of your vote or you wish to change your vote, you may press the corresponding button until the vote is closed.

After everyone has completed their vote, the vote will be locked in. The vote will then be displayed on the screen. The DFO will read the vote from the screen into the record. Next, we will go around the room, and everyone who voted

will state their name and their vote. We will also go ahead and offer comments related to number 6 at that time. We'll continue in the same manner until we have completed.

Based on the data presented on the risk of serious neuropsychiatric adverse event with smoking cessation products, what would you recommend? A, remove the boxed warning statements regarding risk of serious neuropsychiatric adverse events; B, modify the language in the boxed warning; or C, keep the current boxed warning.

If you will enter your vote.

(Vote taken.)

MS. BHATT: The voting results, A is 10; B is 4; C is 5; and there is zero no voting.

DR. PARKER: Great. We're going to go around as suggested, and we'll start on this side with Dr. Emerson and go around if you will. State your name and your vote. And also in addition to your rationale there, discuss any additional labeling actions that you feel the agency should take regarding the risk of serious neuropsychiatric

adverse events with smoking cessation products.

DR. EMERSON: This is Scott Emerson. I voted to remove the box. It was a hard decision between that and modifying the wording. But I decided the way I'd modify the wording would be watered down enough that you'd wonder why the box was there.

I personally believe that the evidence from the clinical trial is certainly suggestive enough, that on a safety endpoint, you would want to have strong warnings that there's a suggestion that patients with prior psychiatric conditions should be watched carefully while they're on the drug and be sure to withdraw it otherwise.

But I wasn't convinced that based on the data that we had -- again, relative risk of approximately 1.4 on something that might be up to a 7 percent baseline rate but was only observed at a 5 percent baseline rate, that a lot of that was driven by things that were unpleasant certainly, but not necessarily life-threatening, any aspect of that, that that rose quite to the level of a boxed

warning.

But I would be very much against anything that watered down that concept that there was a warning, and that the recommendations were in the psychiatric cohort in particular, but with some notice even in the non-psychiatric cohort that these adverse events could occur; and that certainly for the most severe things that led to the anecdotal reports and the postmarketing surveillance, we had nowhere near the sample size that would have picked that up.

So we can't really claim that we're doing that. It's just that the underlying risk and casting the slightly wider net didn't raise any strong suspicions to the level of boxed warning.

DR. MARDER: Steve Marder. I gave a lot of thought to this, and I wavered. Ideally, most physicians would prescribe varenicline and bupropion, and they would also watch their patients more carefully, but I didn't have that choice.

I really thought that increasing the amount of prescribing was important. The fact that over

the years, with additional data that's come up, that there hasn't been anything that's come up that's sort of reinforced the initial anecdotal spontaneous report, made me think that the signal is just not strong enough to justify a black box.

DR. MORGAN: Glen Morgan. I voted in favor of dropping the black box for three principle reasons: the effectiveness of the medication; the general population, but also specifically for the psychiatric population that are especially vulnerable because of their high rates of smoking and their higher rates of failing to quit smoking when they sought to cease. And the third reason is the outcomes of the study that was presented overall.

In terms of where we go from here with guidelines and warnings, I would start with a description of potential adverse events and serious adverse events and say perhaps with an adverse event, contact your practitioner. With a serious adverse event, perhaps stop the medication and contact your practitioner immediately.

DR. MORRATO: Elaine Morrato. I voted C, and partly in principle I think in terms of the precedent-setting nature. Not that the trial couldn't change the boxed warning in my mind, but I just felt that some more time is needed to really complete a more robust sensitivity analysis in light of some of the shortcomings that were identified in the data collection and ascertainment.

Not that the sensitivity conducted looks promising, but I just had this sense that the sponsor is turning a new analyses a week ago. FDA I would imagine would like to have more time to complete their analyses. In the briefing document, there was mention that hundreds of narratives had to be requested -- FDA could only do a sampling -- to look at the adjudication, et cetera. So I feel that because of the precedent-setting nature of taking a box off, I would feel more comfortable if that had been given the time to be more robust.

In my mind, it's very similar to the Avandia

situation, where it was around the REMS and loosening the REMS, and the RECORD study initial analysis raised questions around the conduct. In that case, an independent adjudication occurred. And once the committee saw that full report, it was more comfortable in voting in terms of lessening the REMS. Assuming that turns out, and it confirms the initial sensitivity analysis, I would support what others are saying in terms of removing the box and making it a warning.

This is the time where I know we're supposed to talk about warning, and I know Dr. Perkins will maybe add, but I was struck by -- I don't think I would take out as much of the warning information as what was proposed in the redacted labeling. I think the sponsor was taking out all mention of alcohol. And we haven't talked that, but I seem to remember data that was suggesting that alcohol use was one of the determinants of likelihood of adverse events. So I think some of those data or information shouldn't be removed completely when this warning is discussed.

As I said earlier, I think careful attention if you're going to add in the trial data. If there is sensitivity analyses that might be illuminating, some of that may need to be into the label, depending, but I would recommend I think removing the observational data or putting it into context.

The reason why I feel strongly on the robustness of the sensitivity analysis is because I worry; the unintended consequence of the message of we removed a warning, and the message meaning, oh, now it's safe, and then kind of the pendulum swings the other way, and people assume everything is safe. I think in this case, it requires careful messaging that while a box is maybe being removed, it doesn't mean it's being removed at all in total from labeling as a warning.

DR. PERKINS: Ken Perkins. I voted to remove the box. Clearly, there's a great deal -- and the reasons are very similar to what Dr. Morgan said, the clear efficacy of the drug for help to quit smoking, especially the neuropsychiatric population is less likely to get

it with the warning as it is, the lack of clear evidence of an increased risk overall.

In terms of the labeling, there already is quite a bit of information about symptoms to look for and what to do if they occur, including added symptoms. And I'm just going by what was provided here as the suggested changes for, which appear quite substantial in terms of adding the EAGLES trial data as well, that those who prescribe this drug and those who are doing the prescribing will have clear information about what to watch for anyway.

So we're not abandoning that possibility, so that I think the information provided will be improved in some respects and more clearly indicating how people should proceed if they have these serious neuropsychiatric symptoms.

DR. HERNANDEZ-DIAZ: Sonia Hernandez-Diaz. I voted for removing the box, although I like very much Dr. Morrato's qualifications of saying something like pendulum, the feeling from FDA that they have all the analysis they need to have. But

assuming the extra information on the sensitivity analysis support the conclusions, then I will vote to remove the box.

It was a hard decision to me because I think it's important to highlight a warning that this is especially important to prescribers. So I hope that still in the label it's very clear that, yes, there is an important increase of very important outcomes right after starting the interventions to quit smoking, however, we'll have to put a box to all the interventions if we want to put a box.

So I think that that's important to make it very clear in the label that risk is particularly high in some groups with psychiatric disorders, and probably also had a high risk before the intervention, of course.

So I would not remove the wording of the causal effect on the outcome from the interventions since it may be an effect that is mediated through quitting smoking or anything associated with it. But there's still an effect. According to the clinical trial in the first days, months, right

after the intervention, there is an increased risk of these events. And I think it's important to highlight for prescribers and patients that they have to be watching out and careful about them during these interventions.

MS. HIGGINS: Jennifer Higgins. I voted to modify the language in the boxed warning. And I appreciate the effort that was involved in conducting the EAGLES trial, but I still worry about the psychiatric population in particular, with whom I work daily, and I did mention that in some of my comments earlier. I think we also need additional research to remove the boxed warning.

With respect to language changes, I'm looking at the warning right now, and I think something along the lines of, although risks of neuropsychiatric symptoms may be present for all, they're potentially enhanced risks for the psychiatric population.

MS. GILLESPE: Hi. Terry Gillespe. I voted to keep the box for the main reason that there are a lot of people out there who don't admit that they

have problems. And if it's on there, and they do have problems with this medication, they go to the doctor, and the doctor says, oh, yeah, that's part of the risk because of this reason.

I think that this drug should be given out in combination with psychotherapy or some type of psychiatric care because of the reasons that people hide things that may cause effects, different effects, of this medication.

CAPT BUDNITZ: Dan Budnitz. I voted B, to modify the language of the boxed warning by adding a description of the EAGLES study, but also including caveats and limitations of the study, many of which we discussed today.

I was conflicted in this decision. I work at CDC, but I don't represent the agency here. But I certainly want to promote the availability of smoking cessation therapies because, obviously, decreasing smoking is a national public health priority, and we should try to reduce barriers to treatment options.

If removing a box would do this, that's

good, but on the other hand, I don't want to necessarily set this precedent because removing a box based on this single study, where it's an unvalidated outcome measure that's really been used, and there's concern about complete ascertainment of this outcome of adverse events, I don't think is a good precedent to set for removing a box.

I would add that the rationale for keeping the box that I used was having a box to highlight especially important information for prescribers. And I think that especially important information is these suicide attempts and suicidality, which was an outcome that, I think we agree, the study was not powered to address. So if that's the rationale for the boxed warning, we don't have the data to remove it.

So in conclusion, I'd hope that modifying the language would still allow prescribers to have that information, but also improve access for folks as well for treatment.

DR. WINTERSTEIN: Almut Winterstein. I

voted A, and I could easily have voted B. And I was conflicted in what I was supposed to do, and here is the reasoning for that.

I do think that the evidence for the boxed warning has actually not changed much to the time when it was implemented. We do have a clinical trial that was fabulous to put together, and 8,000 patients is an amazing accomplishment. And I think we all agree that the claims data doesn't have the ability to really address this problem or measure the outcome effectively or adequately. But this trial unfortunately didn't do this either, so now we are left with a trial that cannot tell us whether there is a causal association between drug exposure and more severe events such as suicidal aggression. I think these are more of the things that we would care about other than besides the agitation noise, as Dr. Parker coined it now.

So from that perspective, I don't think that the evidence has changed so much. But then thinking about what evidence was there when the black box warning was put in place, if that was

really only based on spontaneous reports, and we're looking at spontaneous reports that capture symptomatology that goes hand in hand with the indication for the treatment, then it becomes very difficult to think about causality and what really -- this is different from hepatotoxicity for an antihypertensive, where clearly the indication has nothing to do with the hepatotoxicity. But here we have all kinds of alternative explanations, so it makes it much harder, which basically means that we really don't have causality here.

Then looking at the criteria for a black box warning, the evidence that we have available to support causality didn't really be strong enough, to me, to warrant the black box warning. So it's just kind of a logical step in this decision of whether this warrants a black box warning or not.

I am concerned, though, that the decision to remove the black box warning will be misinterpreted by consumers and clinicians as that there is no problem. I think we need to be very cautious about making clear that there is not a clinical trial

that is a de-warning [ph]. There is a clinical trial that hasn't raised additional questions or concern, but it also hasn't produced a de-warning of what was the original trigger for the black box warning, which in my understanding was aggression and suicide. So I don't think that this question has been answered.

So whatever labeling decision was made of whether this was still included in a black box warning or not, I would advocate for if the clinical trial was presented, that it was made very clear that this composite outcome that we're dealing with needs to be interpreted with all the restrictions and limitations that we have discussed here.

DR. GERHARD: Tobias Gerhard, Rutgers. I voted A, to remove the boxed warning. Although I agree with almost everything that Dr. Winterstein just said, I actually think that the trial added additional evidence. And I think what it did is basically to allow us to quantify -- even with all the limitations that were discussed about the

endpoints, it allowed us to quantify the risk compared to what was known at the time the box was put in place based on these spontaneous reports.

With that quantification of the concern, even with wide confidence intervals, I think it makes it pretty clear that in this specific drug indication, the benefits outweigh the risks pretty clearly. And that I think puts it in -- makes this a warning and not a black box warning. Here, the benefit outweighs the risk and not -- it doesn't raise to that level which would really change that risk-benefit for a lot of patients.

So that being said, however, I think the warnings should be in the label. I actually would argue that this trial raised our confidence in that these concerns in the population with a history of psychiatric illness are real. So I don't think that this issue of statistical significance and whether the confidence interval crosses one here as particularly meaningful, this wasn't a trial that was powered that way. The interpretation of the findings that's most compatible with the data is

that there is an increased risk in this group, and we know have some point estimate and confidence limits for it, and I think that should be very clearly communicated.

Probably we won't be able to avoid that impression that the removal of black box is evidence for safety or and endorsement of the safety of the drug, but I think the agency should do whatever it can to prevent that perception.

Two additional comments, briefly. I think there are some real concerns about, in part, somewhat poor conduct of parts of the trial, and that is I think something that the agency needs to be very careful with and the advisory committee needs to be very cognizant of in these trials that try to -- in safety trials where there is a motivation to basically conduct a sloppy trial because it will make it less likely to find results. I don't think that this was the case here, but it's just something to be very cognizant of.

The last comment is that I don't think that

this should set a precedent in the sense that whenever there's a safety study done for an existing black box, and this safety study doesn't show a statistically significant finding, then that black box should be removed. I think that would be a very incorrect reading of this discussion and what has been shown. I think it's rather an issue of evaluating the risk-benefit ratio in each context. So that is I think an important distinction.

DR. PARKER: Ruth Parker. I voted B, which is a different mode than my colleague to my right, but for the same reasons. I had concern about the trial conduct that we discussed at length and its potential huge impact. And that really influenced my vote more than anything; a great concern about life and narratives and just a lot of unknown about the conduct of the trial.

I think we all have a very high bar for safety, and I do believe that removing the black box warning will indeed be read as safety. As mentioned, the agency should do what it can to

prevent a perception that its removal means it is safe. I don't know how you do that. I don't see that as a doable task, and that's what led me to vote the other way.

I think that removing the black box warning, it's going to be interpreted that way. You see it. It's easier to find. The other stuff is harder to read and harder to get to. So I have concern about that. And the reason is really because my read of what we heard today is that I don't know. And if I don't know, then it's harder with safety to make that leap that we just talked about, for me.

I do think that it should be changed to reflect the trial findings, the higher frequency of adverse events among those with a prior psychiatric history, which has been mentioned by others. Thank you.

DR. NARENDRAN: Raj Narendran. I voted to keep the black box warning. I just didn't think the trial was conducted with the same elegance and cleanliness that typically is done for an NDA for efficacy. I don't really feel fully reassured that

this trial -- which was a very complicated thing to do, so I do appreciate them for doing it. But I just don't think it really captured the essence of what is being reported by the general public as adverse events.

So I wasn't fully reassured, and I think removing the black box does send a wrong message that, okay, this drug is now safe. Let's just prescribe it. Although you could potentially add it in the warnings and precautions section, let's face it, how many people are really going to read that as, "Oh, no. It's not in the black box; it's all of a sudden in the warnings and precautions."

So I think it has a potential to cause more problems, and I think keeping it in there, this information for the prescribers and the public, in my opinion has very little harm contrary to what several people voiced.

DR. PARKER: Dr. Besco?

DR. BESCO: Kelly Besco. I voted to retain the black box warning. What it came down to for me is the impact of the variability that was found in

the control trial, which really influenced my comfort in downgrading a warning for these products, especially when you consider the severity and harm associated with these events.

I also agree with much of what others have said about the potential precedent of this decision, especially the need about careful messaging. Again, I am very fearful that the public will assume that, since we've removed -- if the black box is removed, that these products do not present any safety.

DR. PICKAR: Dave Pickar. I voted to do away with the black box with no ambivalence whatsoever. The risk-benefit ratio is as clear as anything I've seen. If you work with these patients and you see what their life span is due to smoking, it's extraordinary. I have never, ever, ever seen a schizophrenic patient on this drug. I don't know if I've ever seen a psychiatric patient on the drug.

By the standards of what you have to put up with in terms of behavioral problems with these

patients, this is mild. You're seeing these as very serious. There's a certain naïveté about serious mental illness and what's involved in managing those folks. So for you, okay, they are adverse events, but not warranting of a black box, particularly when the benefit to these patients could be substantial in the most fundamental thing, which is being alive. So that's why I say with no ambivalence.

DR. FIEDOROWICZ: This is Jess Fiedorowicz. I voted A, to drop the black box warning. I felt compelled to update the prescribing information based on the aggregate of current evidence, which included more than just the EAGLES trial, but cumulative studies since the warning was placed.

As far as regarded changes in the labeling, I have some concerns about the underreporting of incidents in the trial, and there are a lot of proposed updates that list the frequencies of events. And I think that we need to take that into consideration when we're updating that and whether those are the best and most accurate estimates of

risk, because I certainly share the concerns about people underestimating risk, and I think that could promulgate that concern.

I am certainly open to suggesting and would indeed recommend close monitoring for those with psychiatric disorders. They are at higher risk of these complications. As Scott Emerson mentioned, the folks with a specific illness are more likely to have problems like that, and we saw in the placebo groups, the placebo group for those with psychiatric disorders, higher rates of these events.

I think we should be very cautious, however, about any insinuation that folks with psychiatric disorders are less likely to respond to varenicline or more likely to have adverse effects. While the point estimates were in that direction, I don't think that was convincingly demonstrated, and I think we run the risk of further disenfranchising this potentially vulnerable population, and doing so without evidence.

DR. ROUMIE: Christianne Roumie. I voted B,

which was to modify the warning, and for many of the reasons brought on by my colleagues who both voted to remove the warning and to keep the warning. So we basically all use the same reason and fall in all three patterns. But my primary reason was listed in the boxed warning as to highlight a warning that is especially important to the prescriber.

As a clinician, I always have this discussion with my patients before I prescribe it. And I use it, and it's a great drug, and it's very efficacious. But I do think that there should be a conversation and not this kind of carte blanche, it's fine because there's no black box warning. And I think that it should be a thoughtful conversation between the clinician and the patient, and that should be highlighted, especially in patients where certain side effects may be more likely to occur.

DR. RIMAL: Rajiv Rimal. I voted for C, to keep the box. I think the EAGLES study was -- for all the problems we've identified, it's still a

very important study that has very important implications. If for nothing else, it certainly demonstrates the efficacy of this approach for both cohorts.

So I think in that sense, it's a very important study. I am just not convinced that the study addresses the question that we want to see addressed. In my mind, the big question here was not is this drug efficacious. The big question was, for a certain class of patients, does it introduce harm? And for that question, I'm not convinced that we have, as a result of this study, a definitive answer to that. So in the absence of that kind of evidence, I decided to go with the status quo.

There were two other considerations that I think many of my colleagues have talked about. One is the message that it sends when we remove the label, that people are going to construe that as, "Oh, so now the FDA has removed the label -- the warning." Sorry.

But there's another nuance to that, which is

that because this study was done in multiple countries, in many other countries that don't have the same safeguards as we do in the U.S., the perception is going to be, "Guess what? The U.S. FDA has now removed the warning, and therefore" -- right? I mean, I think that sense of complacency can be further exaggerated in many other settings.

Then lastly, as someone who studies doctor-patient communication, I think having the warning has created opportunities and instances for discussions between -- I would guess. I don't know the data on that. But I would guess that just having that warning has provided a venue for discussion between patients and their physicians who are prescribing these drugs. And now removing that is yet another instance where I think, in a round-about way, we're sending the message that perhaps those kind of discussions are not as important, inadvertently, but I think we are sending that message.

DR. HENNESEY: Sean Hennesey. I voted for

removal of the boxed warnings. I did so without any ambiguity. I think it's the right thing to do from a public health perspective. I think that the trial did show evidence in people with a prior history of mental health conditions that varenicline is associated with an elevated risk of neuropsychiatric events, but not serious neuropsychiatric events.

It was underpowered for neuropsychiatric events, but the strongest evidence, the only evidence we have that is associated with serious neuropsychiatric events, are case reports, some of which are very convincing. But there are also very convincing case reports in people with placebo. So I think that reduces the convincingness of the spontaneous reports.

If varenicline does cause serious neuropsychiatric, it does so in a relatively small proportion of people. Smoking causes severe adverse events in a very high proportion of people. I think there's very good evidence that varenicline is under-used, and that continuation -- so people

are concerned about the message that taking away the boxed warning has. Us revisiting and continuing the boxed warning would also send a message. It would say that there's a continuing serious safety problem with the drug, and it would continue to promote under-use of an effective smoking cessation therapy, particularly in the group that most needs it, those with serious mental health conditions. So that was the reason for my vote.

DR. PARKER: So before we adjourn, let me ask if there are any last comments from the FDA?

DR. HERTZ: What I'd like to say is, as always, it's really interesting to hear everyone's comments, and it's incredibly helpful to hear how you think about the issues that we bring before you and how you think about the data. And as much as the actual vote is considered, the comments surrounding the questions, the discussion, is something that we will work with quite a bit.

So thank you all for your time. We really greatly appreciate it.

Adjournment

DR. PARKER: Okay. We will adjourn the meeting. Panel members, please leave your name badge here on the table so that they may be recycled. Please also take all your personal belongings with you as you leave. The room is cleaned at the end of the meeting day. Meeting materials left on the table will be disposed of, and thank you very much.

(Whereupon, at 4:56 p.m., the meeting was adjourned.)