FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

ENDOCRINOLOGIC AND METABOLIC DRUGS ADVISORY COMMITTEE (EMDAC)

Wednesday, October 24, 2018

8:30 a.m. to 3:47 p.m.

Day 1

FDA White Oak Campus

Building 31, the Great Room

10903 New Hampshire Avenue

Silver Spring, Maryland

Meeting Roster

DESIGNATED FEDERAL OFFICER (Non-Voting)

LaToya Bonner, PharmD, NCPS

Division of Advisory Committee and Consultant Management

Office of Executive Programs, CDER, FDA

ENDOCRINOLOGIC AND METABOLIC DRUGS ADVISORY COMMITTEE MEMBERS (Voting)

Michael Blaha, MD, MPH

Associate Professor, Cardiology and Epidemiology

Director of Clinical Research

Johns Hopkins Ciccarone Center for the Prevention of Heart Disease

Baltimore, Maryland

Daniel Budnitz, MD, MPH

CAPT, US Public Health Service

Director, Medication Safety Program

Division of Healthcare Quality Promotion

Centers for Disease Control and Prevention

Atlanta, Georgia

Kenneth D. Burman, MD

Chief, Endocrine Section

Medstar Washington Hospital Center

Professor, Department of Medicine

Georgetown University

Washington, District of Columbia

James de Lemos, MD

Professor of Medicine

Sweetheart Ball/Kern Wildenthal, MD, PhD

Distinguished Chair in Cardiology

UT Southwestern Medical Center

Dallas, Texas

Susan S. Ellenberg, PhD

Professor of Biostatistics

Professor of Medical Ethics and

Health Policy (secondary)

Department of Biostatistics, Epidemiology and

Informatics

Perelman School of Medicine

University of Pennsylvania

Philadelphia, Pennsylvania

Cecilia C. Low Wang, MD

Associate Professor of Medicine

Director, Glucose Management Team

University of Colorado Hospital

Department of Medicine, Division of Endocrinology,

Metabolism and Diabetes

University of Colorado Anschutz Medical Campus

School of Medicine

Lead Clinician-Scientist, CPC Clinical Research

Aurora, Colorado

Anna McCollister-Slipp

(Consumer Representative)

Founder, VitalCrowd

Washington, District of Columbia

Peter W. F. Wilson, MD

(Chairperson)

Director, Epidemiology and Genomic Medicine

Atlanta Veterans Administration Medical Center

Professor of Medicine and Public Health

Emory University

Emory Clinical Cardiovascular Research Institute

Atlanta, Georgia

Susan Z. Yanovski, MD

Co-Director, Office of Obesity Research

Senior Scientific Advisor for Clinical Obesity Research

National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)

National Institutes of Health (NIH)

Bethesda, Maryland

TEMPORARY MEMBERS (Voting)

Brendan M. Everett, MD, MPH

Assistant Professor of Medicine

Harvard Medical School

Director, General Cardiology Inpatient Service

Brigham and Women's Hospital

Boston, Massachusetts

Judith Fradkin, MD

Director, Division of Diabetes, Endocrinology, and Metabolic Diseases

NIDDK, NIH

Bethesda, Maryland

George Grunberger, MD, FACP, FACE
Chairman, Grunberger Diabetes Institute
Bloomfield Hills, Michigan
Clinical Professor, Internal Medicine and
Molecular Medicine & Genetics
Wayne State University School of Medicine
Professor, Internal Medicine
Oakland University William Beaumont School of
Medicine
Visiting Professor, Internal Medicine
First Faculty of Medicine, Charles University
Prague, Czech Republic

Fred Kushner, MD, FACC
Clinical Professor of Medicine
Louisiana State University Medical Center
Clinical Professor of Medicine at Tulane University
Adjunct Professor of Medicine
New York University
Langone Health
New York, New York

Richard Dan Lumley

(Patient Representative)

Kansas, Missouri

Martha Nason, PhD

Mathematical Statistician

Biostatistics Branch

National Institute of Allergy and

Infectious Diseases, NIH

Bethesda, Maryland

Connie B. Newman, MD, FACP, FAHA, FAMWA

Adjunct Professor, Department of Medicine

Division of Endocrinology, Diabetes and Metabolism

New York University School of Medicine

New York, New York

David C. Robbins, MD

Professor of Medicine

Director, University of Kansas Diabetes Institute

University of Kansas School of Medicine

Kansas City, Kansas

Yves Rosenberg, MD, MPH

Chief, Atherothrombosis and Coronary Artery Disease Branch
Division of Cardiovascular Sciences
National Heart, Lung and Blood Institute, NIH
Bethesda, Maryland

Thomas J. Wang, MD

Professor of Medicine
Gottlieb C. Friesinger II Chair in Cardiovascular Medicine
Director, Division of Cardiovascular Medicine
Physician-in-Chief
Vanderbilt Heart & Vascular Institute
Vanderbilt University Medical Center
Nashville, Tennessee

ACTING INDUSTRY REPRESENTATIVE TO THE COMMITTEE

(Non-Voting)

Scott Wasserman, MD, FACC

(Acting Industry Representative)

Vice President, Global Development

Head, Cardiovascular, Metabolic, and Neuroscience

Therapeutic Areas

Amgen, Inc.

Thousand Oaks, California

FDA PARTICIPANTS (Non-Voting)

Mary Thanh Hai, MD

Director (Acting)

Office of Drug Evaluation II (ODE-II)

Office of New Drugs (OND), CDER, FDA

William Chong, MD

Director (Acting)

Division of Metabolism and Endocrinology Products

(DMEP)

ODE-II, OND, CDER, FDA

Lisa Yanoff, MD

Deputy Director (Acting)

DMEP, ODE-II, OND, CDER, FDA

Patrick Archdeacon, MD

Clinical Team Lead

DMEP, ODE-II, OND, CDER, FDA

Mahtab Niyyati, MD

Clinical Reviewer

DMEP, ODE-II, OND, CDER, FDA

C O N T E N T S

C O N T E N T S (continued)

P R O C E E D I N G S

(8:30 a.m.)

Call to Order

Introduction of Committee

DR. WILSON: Good morning. Thank you all for coming to this meeting. I'd like to remind you to please silence your mobile phones or any other devices that may make noise or might interrupt the proceedings.

We have a press contact, I believe, Amanda Turney. Is she in the room? If she could identify herself, if she's here. I don't see her. But if you want to reach out to her, you can come to me or my colleague here, Latoya Bonner, and we'll direct you to her.

I'm Peter Wilson. I'm the chair of the Endocrinologic and Metabolic Drugs Advisory Committee, and I'll be chairing the meeting, and we're now calling the meeting to order.

We're going to go around the table and first introduce ourselves, and we'll start with the FDA on my left.

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

DR. CHONG: William Chong, acting director, Division of Metabolism and Endocrinology Products.

DR. YANOFF: Good morning. Lisa Yanoff, acting deputy director of the Division of Metabolism and Endocrinology Products.

DR. ARCHDEACON: Hello. I'm Patrick Archdeacon. I'm an acting team lead in the same division.

DR. NIYYATI: Hello, my name is Mahtab Niyyati, same division.

DR. GRUNBERGER: I don't work for the FDA. I'm George Grunberger. I'm an adult endocrinologist, and I do diabetes for a living in Michigan.

DR. NASON: My name is Martha Nason. I'm a biostatistician at the National Institutes of Health, specifically in the National Institutes of Allergy and Infectious Diseases.

DR. KUSHNER: I'm Fred Kushner. I'm a

clinical cardiologist and a professor.

DR. LOW WANG: Cecilia Low Wang. I'm an endocrinologist at the University of Colorado and CPC Clinical Research.

DR. BLAHA: Hi. I'm Mike Blaha. I'm director of clinical research at Johns Hopkins Ciccarone Center for Prevention of Heart Disease.

DR. FRADKIN: Judy Fradkin. I'm an endocrinologist and a director of the Division of Diabetes, Endocrinology, and Metabolic Diseases at the National Institute of Diabetes and Digestive and Kidney Diseases at the NIH.

DR. EVERETT: I'm Brendan Everett. I'm a cardiologist at the Brigham and Women's Hospital and Harvard Medical School in Boston.

CDR BONNER: Good morning. I am LaToya Bonner, DFO for EMDAC.

DR. WILSON: Peter Wilson, endocrinologist, Emory University, also preventive cardiology and epidemiology.

CAPT BUDNITZ: Dan Budnitz, an internist and epidemiologist with the Centers for Disease Control

medication safety program.

DR. DE LEMOS: James de Lemos. I'm a cardiologist at UT Southwestern in Dallas.

DR. NEWMAN: Connie Newman. Good morning. I am an endocrinologist and adjunct professor of medicine at New York University School of Medicine.

MR. LUMLEY: Dan Lumley, patient rep from Kansas City.

DR. ELLENBERG: Susan Ellenberg, biostatistician, University of Pennsylvania, Perelman School of Medicine.

DR. WANG: Tommy Wang. I'm the chief of cardiology at Vanderbilt University.

DR. YANOVSKI: Sue Yanovski. I'm co-director of the Office of Obesity Research, the National Institute of Diabetes and Digestive and Kidney Diseases.

DR. ROBBINS: I'm David Robbins. I'm the director of the Diabetes Institute at the University of Kansas Medical Center and professor of medicine at the University of Kansas.

DR. ROSENBERG: Good morning. Yves

Rosenberg, a preventive cardiology clinical trialist. I'm the branch chief in the Division of Cardiovascular Sciences, NHBI, NIH.

DR. BURMAN: Good morning. Ken Berman. I'm chief of endocrinology at the MedStar Washington Hospital Center and Professor of Medicine at Georgetown University.

DR. WASSERMAN: Good morning. My name is Scott Wasserman. I'm a cardiologist and vice president, head of cardiovascular, metabolic, and neurosciences, global development at Amgen.

DR. WILSON: We're missing at the present time Anna Slipp, and we'll introduce her when she arrives.

For topics such as those being discussed at today's meeting, there are often a variety of opinions, some of which are quite strongly held, and our goal is that today's meeting will be a fair and open forum for discussion of these issues and that individuals can express their views without interruption.

Thus, as a gentle reminder, individuals will

be allowed to speak into the record only if recognized by the chairperson. We look forward to a productive meeting.

In the spirit of the Federal Advisory Committee Act and the Government in the Sunshine Act, we ask that 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, the FDA will refrain from discussing the details of this meeting with the media until its conclusion. Also, the committee is reminded to please refrain from discussing the meeting topic during breaks or lunch.

Now I'll pass over to Commander Latoya Bonner, who will read the conflict of interest.

Conflict of Interest Statement

CDR BONNER: The Food and Drug Administration is convening today's meeting of the Endocrinologic and Metabolic Drugs Advisory

Committee under the authority of the Federal Advisory Act of 1972. With the exception of the industry representative, all members and temporary voting members of the committee are special government employees or regular federal employees from other agencies and are subject to federal conflict of interest laws and regulations.

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

FDA has determined that members and temporary voting members of this committee are in compliance with federal ethics and conflict of interest laws. Under 18 U.S.C. Section 208, Congress has authorized FDA to grant waivers to special government employees and regular federal employees who have potential financial conflicts when it is determined that the agency's need for a special government employee's services outweighs

his or her potential financial conflict of interest or when the interest of a regular federal employee is not so substantial as to be deemed likely to affect the integrity of the services which the government may expect from the employee.

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

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

The agenda involves discussion of the "Guidance for Industry: Diabetes Mellitus, Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes" and the cardiovascular risk assessment of drugs and biologics for the treatment of type 2 diabetes

mellitus.

This is a particular matters meeting during which general issues will be discussed. Based on the agenda for today's meeting and all financial interests reported by the committee members and temporary voting members, no conflict of interest waivers have been issued in connection with this meeting. To ensure transparency, we encourage all standing committee members and temporary voting members to disclose any public statements that they have made concerning the topic at issue.

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

With regard to FDA's guest speaker, the agency has determined that the information to be

provided by the speaker is essential. The following interests are being made public to allow the audience to objectively evaluate any presentation and/or comments made by the speaker.

Dr. Robert Ratner has acknowledged that he has consulted for Novo Nordisk. He also holds stocks in Abbott, Johnson & Johnson, and Virta. As a guest speaker, Dr. Ratner will not participate in committee deliberations, nor will he vote.

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

FDA encourages all other participants to advise the committee of any financial relationships that they may have regarding the topic that could be affected by the committee's discussion. Thank you.

DR. WILSON: Thank you very much.

Next, we're going to have introductory remarks from the FDA, Dr. William Chong.

FDA Introductory Remarks - William Chong

DR. CHONG: Good morning. As acting director of the Division of Metabolism and Endocrine Products, I would like to welcome our committee members, our invited speakers, and members of the public to today's meeting.

Over the next two days, we'll be discussing the Guidance for Industry: Diabetes Mellitus, Evaluating Cardiovascular Risk and New Antidiabetic Therapies to Treat Type 2 Diabetes. For simplicity, rather than repeating this every time, I'm just going to call it the guidance.

Just about 10 years ago, this guidance was published, and you'll hear more about the reasons for the publication of this guidance later today. But briefly, it was to address the safety concern related to drugs that were designed to improve glucose control.

Some of the members of the committee today, and probably some members of the public, were here

in an advisory committee meeting held July of 2008 to discuss this issue, and we look forward to hearing some updated comments on the concern.

So briefly, the guidance was issued because of a safety concern, specifically that the establishment of cardiovascular safety was needed. In general, that is focused on atherosclerotic types of disease, and we generally accomplish this through postmarketing trials.

I want to touch briefly on some of the regulatory basis that we've been able to require these trials. In 2007, the Food and Drug Administration Amendments Act was signed. In that act, FDA gained additional authorities, which included the authority to require post-approval studies.

Reasons we could require post-approval studies are shown here. First, if there was a need to further assess a known serious risk; second, if there was a need to further assess a signal of serious risk; or as was pertinent to the cardiovascular safety concerns for diabetic drugs,

if there was a need to identify an unexpected serious risk when available data indicated the potential for a serious risk.

So since 2008, the drugs that we've approved to improve glycemic control have generally included a postmarketing requirement, or PMR for short, requiring that sponsors conduct a cardiovascular outcome trial. As I mentioned, this PMR has to be based on a safety concern that the available data indicated the potential for serious risk and the language in the approval letter reflects that.

An example of the language that has been included in the letter is shown here, and I'm just going to read it.

"There have been signals of serious risk for cardiovascular events with some medications developed for the treatment of type 2 diabetes mellitus, and available data have not definitively excluded the potential for this serious risk with," and you can insert the name of your drug here.

"We've determined that only a clinical trial, rather than a non-clinical or observational

study, will be sufficient to assess a signal of serious risk of major adverse cardiovascular events or the antidiabetic medications, including," again, inserting the name of the drug here.

In the 10 years since that conversation at the advisory committee and the publication of the guidance, we have now seen several trials completed. Some of them are shown here. There are some additional trials that have been reported recently. With the new data that we've learned over the last 10 years, it seems relevant and appropriate to revisit the discussion from 10 years ago. The question that we're faced with is, now what are we supposed to be doing going forward?

I want to turn briefly to our agenda for the next two days. This first day will include presentations by the FDA and some invited speakers. We'll start our FDA presentations by going back to 2008, and Dr. Lisa Yanoff will be revisiting all of the concerns and issues that were discussed at the advisory committee, as well as the concerns that led to the publication of the guidance. She'll

also be providing us with an overview of the recommendations of the guidance along with some of the reasons behind those recommendations.

Dr. Patrick Archdeacon will then follow, and he'll remind us of the overview and approach of assessing cardiovascular safety that was being used prior to 2008 as a reminder of what we used to do and to provide some context to help us understand how things have changed.

Then the FDA presentations will conclude with Drs. Tanya Condarco and Mahtab Niyyati, who will give us a rapid review of the cardiovascular trials that have been completed to date, discussing both some of the designs and results for the trials that have been conducted over the last 10 years.

After we get through the FDA presentations, we'll move on to some of our outside speakers. Dr. Robert Ratner from Georgetown University will share his thoughts on cardiovascular outcome trials for products to treat diabetes and discuss some of the alternative approaches to assessing cardiovascular risk.

We'll take a break for lunch after his presentation, and then after lunch, we'll have Dr. Marc Sabatine presenting on behalf of the thrombolysis in myocardial infarction study group. Dr. Sabatine will be discussing some of the work that goes into the design and conduct of cardiovascular outcome trials and will also share some of his thoughts as a cardiologist and things to consider as we think about the future of the guidance and these trials.

Our last speaker of the day will be Dr. Jennifer Green from Duke University, and she will round out our presentations and provide an endocrinologist's thoughts and perspectives on the guidance and discuss what she sees as the impact.

That will take us through the end of the first day. Throughout the day, there will be time for questions, and we're hoping for a lively discussion and look forward to hearing that.

On the second day is when we'll have our open public hearing, and we'll be gaining insight from the public comments. Then we'll turn to you,

the committee, and put you to work. That's when we're going to ask you for your thoughts and your recommendations on the discussion topics and the voting question. I'm going to go over those briefly over the next few slides.

One of the topics we'll want you to discuss will be the impact of the recommendations in the 2008 Guidance for Industry: Diabetes Mellitus, Evaluating Cardiovascular Risk and New Antidiabetic Therapies to Treat Type 2 Diabetes on the assessment for cardiovascular risk for drugs indicated to improve glycemic control in patients with type 2 diabetes mellitus.

The next discussion topic covers multiple parts. For each recommendation described in the guidance, we'd like you to discuss the value and the evaluation of the safety of the antidiabetic drugs. The recommendations we'd like you to consider are shown here.

First, we'd like you to consider the establishment of an independent cardiovascular endpoints committee for prospective adjudication.

We'd like you to discuss inclusion of patients at higher risk for cardiovascular events in phase 2 and phase 3 trials to obtain sufficient endpoints to allow for a meaningful estimate of risk.

We'd like you to discuss the exclusion of 1.8 from the upper bound of the two-sided 95 percent confidence interval for the estimated risk ratio prior to approval.

Lastly, we'd like you to discuss the exclusion of 1.3 from the upper bound of the two-sided 95 percent confidence interval for the estimated risk ratio to conclude that there is no unacceptable increase in cardiovascular risk.

Discussion topic 3 is shorter. We're going to ask you to discuss how the cardiovascular safety findings from members of a drug class should or should not be applied to all members of the drug class, and then that will bring us to our voting question.

The 2008 Guidance for Industry: Diabetes Mellitus, Evaluating Cardiovascular Risk and New

Antidiabetic Therapies to Treat Type 2 Diabetes, provided recommendations on excluding an unacceptable increase in cardiovascular risk for all new therapies to improve glycemic control in patients with type 2 diabetes. This was regardless of the presence or absence of a signal for cardiovascular risk in a specific drug's development program.

We'd like you to vote on whether an unacceptable increase in cardiovascular risk should be excluded for all new drugs to improve glycemic control in patients with type 2 diabetes, again regardless of the presence or absence of a signal for cardiovascular risk in the development program.

If you were to vote yes for this, we'd like to hear your rationale, and include in your discussions what changes, if any, you would recommend to the 2008 guidance and why, as well as what kind of assessment would be appropriate and when it should be conducted.

If you vote no, we'd like to hear your rationale again, and include in your discussion

what might constitute a signal of cardiovascular risk that would warrant conduct of a cardiovascular outcome trial or other form of cardiovascular risk assessment.

I want to thank our invited speakers and committee members for your service to this committee and this meeting. We look forward to hearing your thoughts, hearing the discussion over the next two days, and take that information to help us inform how we want to move forward with the evaluation of cardiovascular risk for diabetic drugs.

I would like to introduce Dr. Lisa Yanoff, but perhaps we should also have --

DR. WILSON: Thank you very much, Dr. Chong.

Anna Slipp, would you please introduce yourself? Your microphone, please?

MS. McCOLLISTER-SLIPP: Hi. I'm Anna McCollister-Slipp. I'm here as a consumer representative.

DR. WILSON: Thank you very much.

Now we will proceed with the presentation by

the FDA, Lisa Yanoff.

FDA Presentation - Lisa Yanoff

DR. YANOFF: Thank you.

As Bill said, I'm Lisa Yanoff, acting deputy director of the Division of Metabolism and Endocrinology Products at FDA. For my presentation, I'm going to provide a history of the 2008 cardiovascular guidance and the 2008 endocrine and metabolic advisory committee meeting that was convened to discuss this guidance and this issue, including a reminder of the current diabetes drug approval standard, a review of the data raising concern about drug-specific CV harm, and a summary of the discussion of the 2008 EMDAC meeting. The second part of my talk will be to provide an overview of the current CV guidance recommendations.

Just a brief introduction, it's estimated that over 30 million people in the United States have diabetes mellitus; at this time, about 95 percent of which is type 2 diabetes.

(Audio gap - microphone fades.)

DR. YANOFF: Type 2 diabetes is associated with a two- to fourfold higher risk of cardiovascular death compared to patients who do not have diabetes. Most of these deaths are due to cardiovascular disease and stroke, although there are other important long-term complications of diabetes, such as peripheral vascular disease, and importantly, microvascular disease, including retinopathy, nephropathy, and neuropathy, which can lead to blindness, kidney failure, chronic pain, gastroparesis, et cetera.

We have many treatment options for patients with type 2 diabetes, and since the 2008 EMDAC meeting, two additional classes of drugs, highlighted in red on this slide, have been approved; most notably, the SGLT2 inhibitor drug class.

A question was raised at the 2008 EMDAC meeting related to how much do we really need more therapies if we have so many, and how should we consider that need as it relates to how much excess CV risk might be acceptable?

We believe it is important to have many treatment options, for one because type 2 diabetes is a progressive condition, and a patient may start with one therapy but over time need more and more drugs to control their condition.

Development of treatments that target different parts of the pathophysiology are also important, and another relevant consideration is that class- or product-specific adverse reactions may limit use by a certain patient or group of patients such as metformin in patients with renal failure.

Patient acceptability of therapies is also a factor. Reportedly, weight gain is a major concern. And finally, hypoglycemia can limit the success of reaching glycemic goals. So we continue to think it's important to develop new therapies for type 2 diabetes.

Drugs for diabetes carry the indication as an adjunct to diet and exercise to improve glycemic control in whatever the patient population is; for example, adults with type 2 diabetes or adults and

children with type 2 diabetes.

In product development, efficacy for glycemic control was and still is established by demonstrating that the new drug is more effective than placebo at lowering hemoglobin A1c, usually at the end of a 6-month trial period. The new drug is usually also assessed in various treatment scenarios of, for example, monotherapy or very commonly add-on to metformin, sometimes also add-on to two or more agents, or as add-on to insulin.

Now, while it's established that patients with diabetes are at increased risk for both microvascular and macrovascular complications, drugs for the treatment of diabetes are approved based on hemoglobin A1c, which is a glycemic-lowering surrogate. Hemoglobin A1c, or A1c as I'll abbreviate it, is formed by irreversible attachment of glucose to hemoglobin. It is directly proportional to the ambient blood glucose concentration, and it correlates with the average blood glucose over the preceding 2 months. A standardized assay is available, making this

measurement reliable over time and across geographic reasons.

For drug development, we consider A1c reduction to be a surrogate benefit on microvascular disease. This is based on clinical trials that have established that glycemic lowering results in a reduction in the onset and progression of microvascular complications.

This slide, borrowed from Dr. Nathan's 2008 EMDAC presentation nicely summarizes the data demonstrating that by lowering glycemia, you can reduce long-term microvascular complications of diabetes.

In the Diabetes Control and Complications Trial, or DCCT, there was a 43 percent reduction in risk for every 10 percent reduction in A1c. In the United Kingdom Prospective Diabetes Study, or UKPDS, there was a 37 percent reduction in risk for every 10 percent decrease in A1c.

Note that the DCCT is a study in type 1 diabetes patients and UKPDS is in type 2 diabetes patients. But in the regulatory space, the A1c

surrogate is accepted for both drugs intended to treat both type 1 and type 2 diabetes. Further, there can be immediate symptomatic benefit from the treatment of more profound hyperglycemia, which is more of a direct clinical benefit than a surrogate.

Also of note, diabetes product labels do not explicitly state that they are indicated for a reduction in microvascular disease. In other words, microvascular benefit is not overtly claimed in the labeling based on the surrogate endpoint of A1c.

This slide is about macrovascular benefit or risk. We note that A1c is not considered to be a useful surrogate for macrovascular disease reduction. For diabetes drug approval, it's theorized that the robust risk reduction in CV events specifically attributed to glycemic lowering has not been shown in type 2 diabetes the way it's been shown for type 1 diabetes in the DCCT because the relationship between type 2 diabetes and CVD is perhaps too complex or with too many interactions, with traditional risk factors such as age, body

weight, renal function, hyperlipidemia, or even inflammatory status.

Also, CVD risk appears to begin even with glucose in the high normal range, or prediabetic range, with a more gradual increase in risk as higher glycemia is reached.

This relationship is illustrated here. Again, I'm using data from the UKPDS study. Macrovascular disease is in the red line and the microvascular disease is in the blue line. Risk for MI is already elevated, even with glucose in the prediabetes range, an A1c between 5 and 6, elevated above the microvascular disease complications. It has a more gradual increase in risk as higher levels of glycemia are reached.

Now compare this to the microvascular pattern, where substantial risk doesn't really occur until you get over about A1c of 7 and the rise is much more dramatic as you reach higher levels of A1c.

In addition to the lack of usefulness of A1c for macrovascular benefit, some evidence has been

emerging that certain antidiabetes therapies may increase the risk for CVD.

Concerns over the unintended increase in risk due to therapies for type 2 diabetes, the patient population vulnerable to CV disease, go back several decades. In 1970, tolbutamide, an antidiabetic therapy in the sulfonylurea class, was reported to increase the risk of cardiovascular mortality in the University Group Diabetes Program Study or the UGDP.

UGDP was a long-term prospective randomized clinical trial designed to evaluate the effectiveness of glucose-lowering drugs in preventing or delaying vascular complications in type 2 diabetes.

UGDP reported that patients treated for 5 to 8 years with diet plus tolbutamide had a rate of CV mortality approximately 2 and a half times that of patients treated with diet alone. The study results led to a new section in the Code of Federal Regulations on labeling for sulfonylurea drugs that required a, quote, "special warning" on increased

risk of cardiovascular mortality and specified that the patients should be informed of the potential risks and advantages of the sulfonylurea and alternative modes of therapy.

Even though the CFR language acknowledges that the findings have been controversial, this experience raised awareness of the issue of hypoglycemic drugs and CV risk.

In 1998, in a publication of the UKPDS, this particular paper reported on a subset of 537 patients who were inadequately controlled on maximum sulfonylurea therapy. These patients were randomized to additional metformin or continuation of the SU alone if the glucose was above 6.1 millimole per liter, which is about 110 milligrams per deciliter.

In this study, the median A1c over 4 years in the cohort with addition of metformin was 7.7 percent compared with 8.2 percent in those on the SU alone. The data had an unexpected finding of an increase in diabetes-related death with metformin add-on to SU.

Although the finding has never been fully explained, some have suggested patient factors or a chance unusually low rate of events in the group randomized to stay on SU alone. There were too few events to draw meaningful conclusions, but the study highlighted the uncertainty about the benefit of using metformin and SUs together.

Continuing on the story, in 2005, a dual PPAR drug called muraglitazar was being developed for the treatment of type 2 diabetes, and the drug appeared to have favorable benefits on endpoints of interest such as A1c, triglycerides, and HDL cholesterol, and appeared to have no adverse impact on LDL cholesterol, but there was a numeric imbalance in events suggesting CV harm such as MI, stroke, CV death, and heart failure.

The number of events in the overall muraglitazar development program was not high enough to be able to have clear evidence of harm. There were roughly about 40 events, if I remember correctly. But the trials were concerning, and FDA did not approve this product for marketing

authorization. Instead, FDA asked for further CV safety data to help inform the signal. And in 2006, development of the product was stopped by the sponsor because the additional crude CV safety outcome data confirmed the excess CV risk.

In 2007, a meta-analysis of 42 trials published in the New England Journal reported that rosiglitazone increased the risk of myocardial infarction and cardiovascular mortality compared to placebo and other antidiabetic agents. Again, there really weren't a sufficient number of events to draw reliable conclusions.

While acknowledging that conclusions drawn from post hoc meta-analyses can be unreliable, it's clear that the rosiglitazone experience highlighted some of the uncertainty in the premarketing assessment of CV risk of antidiabetic therapies.

Another example that's often cited is the ACCORD trial, which was stopped early by the DMC for an excess mortality signal for intensive A1c lowering versus standard A1c targets.

This trial focused more on the glycemic goal

rather than any specific antidiabetic agent, but the results contributed to the growing concern about determining the overall clinical benefit of diabetes drugs based on glycemic control.

FDA recognized the need to engage stakeholders about this concern, and on July 1st and 2nd, 2008, the Endocrinologic and Metabolic Drugs Advisory Committee met to discuss the role of CV assessment in a pre- and postmarketing setting.

After considering the discussion at this meeting, as well as other available data and information, FDA determined that concerns about CV risk should be more thoroughly addressed during drug development and issued the guidance for industry, which states that applicants of new antidiabetic medications for the treatment of type 2 diabetes should demonstrate their products are not associated with an unacceptable increase in CV risk.

I would like to remind you of the voting question that was posed back in 2008 to the committee about CV risk assessment. It should be

assumed that an antidiabetic therapy with a concerning CV safety signal during phase 2/3 development will be required to conduct a long-term cardiovascular trial. For those drugs are biologics without such a signal, should there be a requirement to conduct a long-term cardiovascular trial?

The committee was asked to vote yes or no. If yes, please discuss when such a study should be conducted, pre-approval, a post-approval. If a long-term CV trial is required post-approval, please discuss whether this study should be ongoing at the time of approval.

A majority of the committee recommended, yes, a more extensive, standardized assessment of CV risk in order to provide better information about the overall benefit-risk profile of the product in question. It should be emphasized, though, that the focus of this assessment was to evaluate risk, not to necessarily demonstrate CV benefit.

To put this another way, it was felt by a

majority of the committee that having a drug that can reduce the risk of microvascular complications and do no CV harm was still a good thing, although clearly a drug that also had CV benefit would be even better. However, that should not be the regulatory hurdle.

Nevertheless, it was noted by several members that if a trial to meet these requirements was conducted correctly, you could actually test for both lack of CV harm and for CV benefit.

In the committee discussion, the view that A1c is a surrogate for microvascular disease reduction was largely supported, and the acceptability of use of A1c as a surrogate for drug approval was upheld. It was recognized that a surrogate can be validated for one but not all clinical endpoints of interest; and in this case, validated for micro, but not macrovascular disease.

I'll just comment that in today, in 2018, FDA's view of the value of approving diabetes drugs based on glycemic control remains unchanged.

Now I'll describe a little bit about what's

in the guidance. To establish that a new drug for the treatment of diabetes did not result in an unacceptable increased risk, the guidance made several recommendations. For one, prior to marketing approval, it should be demonstrated that the upper bound of the 95 percent confidence interval for estimated hazard ratio for MACE versus a control group excludes 1.8 with a reassuring point estimate. And after marketing approval, it should be demonstrated that the upper bound of the 95 percent confidence interval excludes 1.3.

These goalposts were to apply to all drugs for type 2 diabetes regardless of their demonstrated benefit. So for example, a drug that may have had an enormous A1c lowering wouldn't necessarily have room for extra CV risk. The degree of what was to be considered unacceptable should be the same for all products.

So why and how were these goalposts selected? It was acknowledged that, pre-guidance, diabetes development programs for the most part did not have a sufficient number of CV events to assess

risk. But concerns were raised about feasibility and adversely impacting product development by introducing additional burden.

Just to remind you of the numbers of events to discern certain degrees of risk, if we assume the true relative risk is 1.0, to exclude a 1.3 risk margin for a 30 percent increase in risk, you would need 611 MACE events. For a doubling of risk, for a 2.0 margin, you would need 88 events.

As a general rule, dedicated CV safety trials are designed as event-driven trials, which are a special case of information-based clinical trial designs. A feature of information-based designs is that the statistical information is fixed in advance rather than using the number of subjects to determine the size of the trial.

For an event-driven trial, the statistical information corresponds to the number of events. Therefore, the trial will continue to enroll or follow patients until the prespecified number of events are observed.

To preserve statistical power, you need to

observe a certain number of patient-years. The expected number of patient-years can be anticipated by considering the likely rates of events being assessed, but the actual value will depend on the observed event rate.

When the question was asked, how big do these trials need to be, the best way to think of it is as the number of patient-years needed for a predicted event rate of MACE. In this table, the number of events in the left-most column is the number of events needed to have 90 percent power to rule out the margins in the second column.

The third column shows the maximum point estimate of the hazard ratio that could be achieved for each scenario, which is relevant because the guidance specifies that the point estimate of the hazard ratio should be reassuring. On the far right column, this shows the patient-years needed based on an assumed annual event rate of 3 percent.

Now, if this annual event rate of MACE is different than predicted, more or fewer patient-years could be needed to complete the

trial. In order to accrue events more expeditiously, trials have criteria that ensure higher-risk patients are enrolled so that the predicted event rate is met or even exceeded.

A little bit of a stats tutorial here; the following illustration depicts an event-driven trial objective of showing non-excessive risk. The dashed line corresponds to the risk margin, which represents the amount of risk to rule out.

If the upper bound of the 95 percent confidence interval is below the risk margin, the trial meets the non-excessive risk objective. And here, we see that scenarios 1, 3, and 4 meet the non-excessive risk objective.

Also of note is that the point estimate, shown by the black circles, does not have to be below 1 in order for the upper bound of the 95 percent confidence interval to be below the risk margin.

If we use this illustration to explore what happen if we are trying to rule out a risk margin of 1.3, for example, scenarios 1 through 4 now

represent hypothetical results from an event-driven trial that accrued 611 events. Scenarios 1, 3, and 4 would meet the guidance recommendation to rule out a risk margin of 1.3, while the second scenario would not because the upper bound of a 95 percent confidence interval is 1.33, and this exceeds the 1.3 cut-off.

Also of note, in scenario 1, the drug is also demonstrated to be superior to the comparator because the upper bound of the 95 percent confidence interval is less than 1. So you can show superiority or noninferiority in a similar trial.

To address the question of how the goalposts were selected, they were felt to reasonably balance the considerations of feasibility and how much risk should be considered unacceptable, given that there were many approved therapies for type 2 diabetes at the time of the meeting in 2008, and now even more so.

As the next FDA speaker, Dr. Archdeacon, will discuss, pre-guidance programs had fewer than

even 88 events, which is what you'll need to rule out a 2.0 margin. So a trial would need to be much larger or a longer duration to accrue enough events even to rule out the 1.8 margin. The requirement of a 1.8 risk margin at the time of approval was moving in a positive direction towards enhancing the certainty around CV risk at the time the drug would be marketed.

Post-approval, the requirement to meet a 1.3 risk margin needs over 20,000 patient-years. If you want to exclude a lower degree of risk, you're getting into very large numbers of patient-years needed.

Other recommendations in the guidance pertain to the goal that trial design and conduct should be optimized in order to allow trials to provide reliable and valid results. FDA has worked extensively with sponsors to help these trials to come to fruition, and this has typically involved multiple rounds of protocol review and discussion between the sponsor and the agency.

The guidance recommended that a blinded

independent adjudication committee review events to enhance sensitivity and specificity. And in 2008, this approach appeared to be favored by most of the committee because it was stated by some of the speakers that adjudication was useful and not terribly expensive or burdensome on sponsors.

Additionally, it was recommended that patients at higher risk for CV events be included in the studies. This recommendation in the guidance was for a number of reasons. It was intended to ensure enrollment of a patient population more representative of who would actually be using these drugs and also to ensure that a sufficient number of events accrued to allow for an assessment of risk.

Another important point is that these trials be longer studies than the pre-guidance safety and efficacy trials, which are usually mostly 6-month trials. And this is because it was noted that diabetes is a chronic disease, so long-term studies are warranted, and also that there could be an increased risk in the short term of some outcomes,

but overall favorable benefit-risk over the long term. An example of this phenomenon is retinopathy with intensive reduction in A1c, where early retinopathy risk did not translate into long-term harm.

Also recall that to preserve statistical power in an event-driven trial, we need at least a certain number of patient-years. While this could be accomplished by enrolling an extremely large number of patients, the results from such a trial would not be clinically meaningful in terms of identifying drug-related risk.

So essentially, the higher level goal of the guidance was to recommend that approval of antidiabetic therapies continue to rely on the A1c surrogate, but also improve the assessment of the cardiovascular risk, both pre- and postmarketing, to provide an informed choice of therapy with regard to overall benefit-risk.

We will continue the FDA presentations with information about CV safety data collection, both pre- and post-issuance of the guidance, and results

from CV outcome trials conducted, for the most part, to fulfill the guidance recommendations.

Now, I'm pleased to introduce Dr. Patrick Archdeacon, who will be discussing CV safety assessment before the guidance.

FDA Presentation - Patrick Archdeacon

DR. ARCHDEACON: Thank you, Dr. Yanoff, and thanks to the members of advisory committee for joining us today, and also to the public. Again, my name's Patrick Archdeacon. I'm one of the team leads in the division.

The division is eager to hear from the attendees of this meeting regarding evolving perspectives on the evaluation of products used to manage type 2 diabetes. However, before we dive into the data from the cardiovascular outcome trials that have been completed to date and the relative value of conducting additional CVOTs in the future, we think that it may be of some value to reflect on the previous era of diabetes drug development, as I think that will inform the conversation about best practices and best options

going forward.

When you compare the approaches to the cardiovascular safety assessments of antidiabetic agents before and after the publication of the CVOT guidance, the focus is often placed on differences between patient demographics, trial size, and trial duration.

That's not an unreasonable focus. It's undeniably important that the current trials have shifted towards including older patients, patients with a longer history of diabetes, and patients with established cardiovascular disease. Likewise, it's of obvious import that the studies are now much bigger and of a longer duration.

Perhaps less discussed but also very important, and I think possibly more interesting, has been the impact the guidance has had on the approach to collection, curation, and evaluation of cardiovascular safety data.

For the next 20 to 25 minutes, I'll attempt to illustrate the importance of these factors by summarizing three development programs that were

completed in the era before the CVOT guidance.

Exenatide was approved by FDA in April 2005, sitagliptin in October 2006, and saxagliptin was approved by FDA in July of 2009, supported by an NDA submitted in 2008 prior to the publication of the guidance.

The effect of the CVOT guidance on demographics and disease characteristics was fairly straightforward. That's because the guidance makes explicit recommendations in this area.

So quote, it states, "To obtain sufficient endpoints to allow meaningful estimates of risk, the phase 2 and phase 3 program should include patients at higher risk of cardiovascular events such as patients with a relatively advanced disease, elderly patients, and patients with some degree of renal impairment.

"Because these types of patients are likely to be treated with the antidiabetic if approved, this population is more appropriate than a younger and healthier population for assessments of other aspects of the test drug safeties."

So as Dr. Condarco and Dr. Niyyati will demonstrate in the talk following mine, more recent development programs have indeed shifted towards featuring trials that better reflect the populations of patients that actually depend on antidiabetic drugs. As recently as a decade ago, however, that was not the role.

For example, let's consider the demographics of the exenatide clinical development program. This overall program consisted of 27 studies that ultimately culminated in three pivotal phase 3 trials. Those three trials were control trials designed to establish the efficacy of exenatide when used in combination with metformin, a sulfonylurea, or both.

Of the patients that were enrolled in these pivotal phase 3 studies, only 18 percent were 65 years of age or greater, and only 1.5 percent were 75 years of age or older. The vast majority of these patients had lived with their diabetes for far less than 10 years, and their hemoglobin A1cs were no more than moderately elevated.

Importantly, none of the patients enrolled had experienced any macrovascular or microvascular complications of their disease. Specifically, I'm referring to the enrollment criteria for each of the three trials that explicitly excluded all patients who had any active cardiovascular disease symptoms within the previous 12 months and excluded any patients with any history of clinically significant renal disease.

So shifting to sitagliptin, the demographics of the pivotal trials of the sitagliptin development program mirror almost precisely those of the exenatide program. The mean age of the patients was 55, the mean duration of diabetic disease around 5 years, and the mean hemoglobin A1c at study entry was 8.

Again, all 4 pivotal phase 3 trials specifically excluded patients who had experienced signs or symptoms of cardiovascular disease, in this case for 6 months prior to enrollment, and again excluded all patients who had an eGFR less than 50. to be fair, there was one small study of

91 patients lasting 12 weeks that included patients with chronic kidney disease.

Saxagliptin was developed a few years after exenatide and sitagliptin at a time where there was growing interest in the effects of anti-hyperglycemic agents on cardiovascular risk. However, the baseline characteristics of the patient population of the 8 core phase 2 and phase 3 clinical studies was again largely similar to those of the exenatide and sitagliptin pivotal trials.

The composition of these trial populations was slightly different in that they did at least include a few patients who had a history of coronary artery disease. However, this representation remains small, on the order of 3 to 5 percent of the enrolled subjects, with the exception of one outlier study, where the prevalence of coronary artery disease was 13 percent.

So overall, these three programs are remarkably consistent with one another. They

largely enrolled middle-aged patients relatively early in the course of their diabetic disease with few comorbidities and no more than moderate elevations in hemoglobin A1c. Not surprisingly, as we'll shortly discuss, the demographics of these trials contributed to their limitations with regards to their potential to detect effects on cardiovascular outcomes.

Shifting to exposure, prior to 2008, the accumulated drug exposure in an original NDA application for a new anti-hyperglycemic agent was largely dictated by the ICH E1 guideline recommendations. Those guideline recommendations state that at least 1,000 subjects should have some level of exposure, at least 300 subjects should be studied for 6 months, and at least 100 subjects should be studied for a year.

The amount of exposure achieved in these programs is also related to the size of the studies needed to establish the efficacy of the product intended to improve glycemic control, as reflected by an effect on hemoglobin A1c.

In February 2008, FDA issued draft guidance on the development of drugs and therapeutic biologics for the treatment and prevention of diabetes mellitus. This is a separate guidance that FDA issued from the CVOT guidance that deserved primary focus, but that guidance also recognized the role for more extensive safety data collection for drugs developed for type 2 diabetes.

Specifically, that guidance called for phase 3 trial data from at least 2500 subjects exposed to investigational product with 1300 to 1500 of those expected to be exposed for 1 year or more and 300 to 500 exposed for 18 months or more.

So those recommendations were consistent with the statement that was contained in the CVOT guidance that noted that future controlled trials will need to last more than the typical 3 to 6 months' duration to obtain enough events and to provide data on longer-term cardiovascular risk, and then it says, quote, "e.g. a minimum of 2 years."

So what did the development programs for

anti-hyperglycemic agents look like in type 2 diabetes prior to these two guidances issuing in 2008 with regards to the exposures achieved in terms of total numbers of patients, the durations of the exposures, and the duration of exposure data for which there were adequate controls?

So differences in the design of the development programs complicate meaningful comparisons across programs with regards to the overall cumulative exposures achieved. For instance, some programs conducted large controlled phase 2 trials that contributed valuable exposure data, whereas other programs really relied exclusively on their phase 3 trials to conduct their integrated safety assessments.

Some programs included a broad range of dosing strategies through late-phase studies, while others were able to, early on, focus on data collection for the dose or doses that were ultimately approved. Programs exhibited some variability around the timing of their data collection, so it makes it complicated to compare.

Despite those caveats, however, I think it's still reasonable to consider and compare the composition of the populations that form the core safety assessments that were conducted by FDA during the NDA reviews to get an impression of the useful exposure data that were achieved by these programs.

So the clinical program for exenatide, as I mentioned, was comprised of 27 studies. In those 27 studies, 2,252 subjects participated; 1,857 of those received exenatide.

The early studies led to a selection of 5- and 10-microgram BID fixed-dosed regimens, and those were the regimens that were studied in the 3 6-month-long controlled studies. A total of 1446 patients were included in those trials, including 963 patients that were randomized to exenatide; 483 randomized to placebo.

So open-label extensions of these three pivotal phase 3 studies and a fourth open-label trial did collect some safety data from patients exposed to exenatide beyond 52 weeks. However,

there was no control data available for those patients because the control arms of the pivotal trials were discontinued after 6 months.

Shifting to sitagliptin, a slightly larger program, the total number of subjects exposed to sitagliptin in this development program was 3,276 with a cumulative exposure of 1339 subject-years. The integrated analysis of safety, however, relied primarily on two smaller populations: a pooled phase 3 population and a so-called long-term safety population. The pooled phase 3 population drew from 1538 patients randomized to sitagliptin and 778 randomized to a placebo comparator.

The trials that contributed to the pooled phase 3 safety data collected data over 6 months of exposure. And by the end of those 6 months, 966 patients randomized to sitagliptin remained on the drug. The safety assessments did include this long-term safety population that included data from beyond 6 months of exposure time. However, those assessments were limited, both because there were relatively few patients that remained on study drug

and because even fewer remained in the control arms of those studies.

The saxagliptin program, again a little bit larger, included a total of 4,000 subjects in the overall program, including 3,400 patients randomized to saxagliptin in the pivotal phase 2 and phase 3 trials. The trials also included 1200 patients randomized to control arms, which included either placebo or metformin.

2400 of the patients remained exposed to exenatide for more than 24 weeks and a little over 1,000 were exposed to exenatide for more than a year. Importantly in this program, the control arms also continued out beyond the year, so adequate control data was available for those in this program.

Overall, despite challenges directly comparing the program, I would say it's clear that the exposures achieved by these programs meet or somewhat exceed the guidelines provided by ICH E1, but only the saxagliptin program was consistent with the recommendations later included in the

February 2008 draft guidance.

So before shifting to a discussion of the methods of cardiovascular event data collection common in the pre-2008 era, I think it's worth considering the impact the CVOT guidance had on our current approach to CV data curation.

Recommendations in the guidance included the establishment of an independent cardiovascular endpoints committee to prospectively adjudicate, in a blinded fashion, cardiovascular events during all phases of phase 2 and phase 3 trials, so including events of CV mortality, myocardial infarction, and stroke, and also possibly hospitalization for acute coronary syndrome, urgent revascularization procedures, and other endpoints.

How was this done? While the guidance was not prescriptive about how the recommendations contained in it should be implemented, the CVOTs that have been conducted to date have adopted similar practices for identifying and gathering data for adjudicating cases.

So an informal review of multiple event

adjudication committee charters completed since the 2008 guidance suggest that CV data collection has dramatically changed as a result of the guidance. So standard practices now include protocols for the prospective collection and curation of cardiovascular data.

The triggers for CV data collection are not only investigative reports of MACE, but also potential MACE events that are detected by automated sponsored MedDRA queries, by out-of-range lab values, and by new abnormal ECG findings.

So each of these triggers now results in an immediate site query and follow-up. And during that follow-up, there's extensive source data collection, and this includes events that happened away from investigational sites; source data collection tools that ensure this complete data collection; make sure that we get admission notes, discharge notes, procedure and consult notes; event data from the eCRFs; complete EKGs; labs including biomarkers such as cardiac biomarkers; and even autopsy reports of death certificates where

applicable. These complete data then are assembled, organized, and submitted to the committees for standardized adjudication.

So how does this compare to what was going on pre-2008? And I would submit that there's a sharp contrast. Pre-2008, cardiovascular events were treated like any other adverse event. The events that were captured were those that were reported by the investigator, and we largely relied on the investigator's judgment for categorizing these events and rating their severity.

As I said, in general, in the previous era, protocols did not systematically collect data prospectively for cardiovascular events. The cardiovascular event capture in these studies relied largely on capturing data collection through coding and using the medical dictionary for regulatory activities, also known as MedDRA. This is a methodology that was developed in the late 1990s.

I know many in the audience will be familiar with this, but for those that are not, I thought it

would be helpful to walk through how this works. While the methodology standardizes language used to capture events, there are a lot of limitations.

Coding information using MedDRA begins with a verbatim investigator-reported term for the adverse event. Those investigator-reported terms are converted to a preferred term, of which there are around 22,000, and then those terms are organized under 27 different system organ classes. So let's walk through some examples.

An investigator may report Q waves detected or chest pain to rule out myocardial infarction. That's going to get coded into one or more preferred terms, perhaps myocardial infarction, perhaps EKG signs of myocardial infarction, and those ultimately get categorized into a variety of different SOCs.

How does that affect how a cardiovascular assessment occurred during our overall programs? We have a couple different factors now to consider: the demographics, the exposure, and limitations around CV data collection.

As we just reviewed, the trials supporting exenatide and sitagliptin included exclusively healthy patients and collected control data for less than a year for the majority of patients in the safety database. Perhaps it's not surprising, then, that even high-level queries at the level of system organ class did not identify many potential cardiovascular events.

The primary clinical review of the exenatide NDA notes only that the number of serious treatment-emergent adverse events were essentially balanced. So in the long-term controlled studies, 8 out of 963 patients receiving exenatide and 9 out of 483 patients receiving placebo had serious events that were mapped to the SOC category of cardiac disorders. The lack of event capture at the level of SOC really prevented any assessment at lower MedDRA levels for the exenatide program.

Similarly, the primary clinical review of the sitagliptin NDA noted only that there were too few events captured and mapped to relevant SOC terms to determine whether sitagliptin increases

cardiovascular events.

Although the saxagliptin development program was completed in advance of the publication of the cardiovascular outcome trials guidance, the review of saxagliptin overlapped the publication data of that guidance. For that reason, the division was particularly interested in evaluating the data in the saxagliptin NDA for evidence to support a conclusion that an unacceptable increase in MACE could be excluded. To this end, the MedDRA data were respectively interrogated using two strategies, and I'll describe those a bit more here.

The first strategy we called the broad SMQ MACE, and it was a composite of cardiovascular deaths that were captured specifically and also preferred terms in the standardized MedDRA queries, or SMQs, that we believe mapped to myocardial infarction or central nervous system hemorrhages in cerebrovascular accidents.

The second strategy we employed, we called custom MACE, and that comprised a subset of the

MedDRA terms that FDA reviewers were most likely to represent true events of myocardial infarction or stroke. This slide shows a sample of some of the terms that comprise the broad and custom MACE SMQs. To be clear, this is not all of the terms. To include all of them would have required me to go through four or five such slides.

The preferred terms captured by the custom MACE query included terms like acute myocardial infarction, myocardial infarction, cardiac failure, sudden cardiac death, et cetera, et cetera. In contrast, the preferred terms that drove the broad MACE query were terms like blood creatinine and phosphokinase increased.

As we'll see in the next slide, there were a total of 83 events detected by the broad MACE query, but 63 of those events mapped to the term blood creatinine phosphokinase increase. In the long-term analysis, there were a total of 141 events detected by the broad MACE query, but 88 of those mapped to the term blood creatinine phosphokinase increased. Here are the actual

numbers that I was referring to.

Although the saxagliptin program, compared to, say, exenatide or sitagliptin, achieved greater exposure to study drug in terms of the total number of patients, the duration of exposure, and the availability of the control data, the number of events captured by the more specific custom MACE SMQ query was still rather small. The number of events captured by the broad MACE query was larger.

I'd say with both strategies, but particularly for broad MACE, there were concerns about whether these events that were "ascertained" truly represented real cardiovascular events.

We took these numbers, and we presented them to the saxagliptin advisory committee, and that committee was posed the following question:

"For the custom MACE endpoint, the upper bound of the two-sided 95 percent confidence interval for the risk ratios and odd ratios was less than 1.3. These data involved a total of 11 cardiovascular events in the 24-week, double-blind, short-term study periods and a total of

40 cardiovascular events in the combined short-term and long-term study period, with a median 62-week exposure.

"Are these data adequate to conclude that postmarketing cardiovascular safety trials are unnecessary?"

That committee voted no 12 to 0 with no abstentions, stating that they would like to see detailed focused postmarketing studies conducted in a higher-risk population.

That's my summary of where we were pre-2008, and just a few bullet points I tried to jot down to summarize this. At that time, the studies were designed primarily with the intent of demonstrating an effect on hemoglobin A1c. As a consequence, the duration of those studies was typically on the order of 6 months. The types of patients that were enrolled were typically younger patients with limited comorbidities with a duration of diabetic disease under 10 years.

We often had limited control data that went beyond 6 months. And importantly, there was a lack

of systematic prospective collection of cardiovascular event data, severely limiting our ability to accurately ascertain cardiovascular events.

Thank you for your attention, and now Drs. Condarco and Niyyati will discuss the design and conduct of CVOTs in the current era.

FDA Presentation - Tania Condarco

DR. CONDARCO: Thank you, Dr. Archdeacon, and thank you, committee members, for joining us today.

Dr. Niyyati and I will present an overview of the premarket and postmarket cardiovascular assessment conducted to fulfill the 2008 cardiovascular guidance, then we will discuss the specific trial designs and results for cardiovascular outcome trials. Finally, we will compare and contrast examples of cardiovascular safety assessments preceding and subsequent to the cardiovascular guidance.

Let's start with an overview of the cardiovascular assessments conducted after the

issuance of the 2008 cardiovascular guidance. As Dr. Archdeacon discussed, before the guidance was issued, the clinical trial population enrolled in phase 3 trials excluded patients with established cardiovascular disease. Trial durations were relatively short and cardiovascular safety evaluations were based on investigator-reported adverse events, which were not prospectively specified.

In December 2008, the cardiovascular guidance was issued, which recommended that sponsors demonstrate that new antidiabetic therapies intended for the treatment of type 2 diabetes were not associated with an unacceptable increase in cardiovascular risk through a premarket and postmarket cardiovascular evaluation.

Let's start by discussing some of the different prospective approaches used to assess the premarket CV safety. Dapagliflozin's premarket CV risk assessment relied on a meta-analysis of 21 trials. Of these, 14 were phase 3 trials. The trials differed in their primary objectives,

designs, choice of comparators, populations of interest, and inclusion criteria. There were two trials in the dapagliflozin program which enrolled subjects with a history of CV disease. In total, 178 CV events were captured during the premarket period.

Canagliflozin's premarket safety assessment relied on a meta-analysis of the phase 2 and 3 trials in addition to a pre-planned interim analysis of their respective cardiovascular outcome trial. Over 200 events were captured for canagliflozin. 161 events were captured from the interim analysis with the remainder coming from the phase 2 and 3 trials.

Alogliptin's premarket CV safety assessment relied on an interim analysis of the CVOT EXAMINE. 83 events were captured for alogliptin at the time of interim analysis.

Another approach to the premarket CV safety assessment was to evaluate cardiovascular safety via a smaller cardiovascular outcome trial. Semaglutide captured 254 events using this

approach. A smaller CVOT is a CVOT that is designed to rule out the 1.8 margin without expectation of having sufficient events to rule out the 1.3 margin.

Regardless of the cardiovascular approach used, each program met the guidance's recommendation and ruled out an 80 percent excess premarketing cardiovascular risk as compared to control.

This slide shows the cardiovascular outcome trials that have been conducted or are being conducted as a result of the 2008 cardiovascular guidance. The trials are arranged according to their completion or expected completion date as reported on clinicaltrials.gov.

So far, cardiovascular outcome trials have been conducted with DPP-4 inhibitors, shown in purple, SGLT2 inhibitors, shown in orange, and GLP-1 receptor agonists, shown in pink. Participants in all of the studies, regardless of treatment arm, received the standard of care treatment in addition to the trial drug or

comparator.

Note the two trials shown here were not required by the FDA, but are included for completeness. These are TECOS, which pre-dated the issuance of the guidance, and CAROLINA, which used an active comparator.

Trials also varied in their respective primary outcome. Most trials had a primary outcome of the composite of 3-point MACE, which included nonfatal myocardial infarction, nonfatal stroke, and cardiovascular death.

Two trials, TECOS and ELIXA, had a primary composite endpoint of the 4-point MACE, which in addition to nonfatal MI, nonfatal stroke, and CV death, also included unstable angina, requiring hospitalization. The average follow-up of the trials varied and ranged from 1 and a half years for EXAMINE to almost 4 years for LEADER.

The average follow-up of the trials was dictated by their design, some of the trials being solely event-driven, while others were event driven with a prespecified minimum duration.

As you see, the outcome trials were large overall, but there was heterogeneity in the number of randomized patients ranging from 3,000 in SUSTAIN 6 to over 16,000 patients in SAVOR.

To ensure a sufficient number of events were accrued, these trials were enriched with high cardiovascular risk patients or patients with established cardiovascular disease. Specific criteria as to what was considered established cardiovascular disease differed across trials.

For example, EXAMINE included patients who experienced acute coronary syndrome in the preceding 3 months, while ELIXA enrolled patients who had an acute coronary syndrome event within 6 months. Other trials like LEADER and SUSTAIN 6 had criteria which specified the enrollment of a certain percentage of patients with established cardiovascular disease.

The mean age of the trial population was over 60 years for most CVOTs. The mean hemoglobin A1c at baseline was over 7 percent for all trials, and the mean diabetes duration at the time of

enrollment exceeded 5 years.

Most patients were obese and had a variety of cardiovascular and non-cardiovascular comorbidities. Here I show hypertension, which was common in most trials, and renal impairment, which affected over 10 percent of the enrolled patients in each trial.

Now, I will turn the podium to Dr. Niyyati, who will discuss the trial designs and results of the cardiovascular outcome trials to fulfill the 2008 guidance.

FDA Presentation - Mahtab Niyyati

DR. NIYYATI: Thank you, Dr. Condarco.

I will present the outcome trials according to their completion date as reported on clinicaltrials.gov. I'll start with discussing SAVOR.

SAVOR was an event-driven, randomized, prospective, double-blind trial comparing saxagliptin versus placebo in patients with type 2 diabetes mellitus with mostly established cardiovascular disease. Over 16,000 patients were

randomized through saxagliptin or placebo.

After a median follow-up of about 2 years, over 1200 primary MACE events were accrued and vital status was ascertained in 99 percent of randomized patients. A similar proportion of patients in both treatment arms experienced a primary endpoint.

The upper bound of the 95.1 percent confidence interval was 1.12, with a point estimate of 1.0, ruling out the guidance-recommended unacceptable increased risk for MACE with saxagliptin as compared to placebo.

The contribution of each component of MACE to the composite is shown here. The proportion of patients who experienced each component of MACE was approximately balanced between treatment arms. Some trials had unexpected safety findings. For example, SAVOR showed a possible increase risk for heart failure associated with saxagliptin.

The top table shows the time to the first occurrence of hospitalization for heart failure. The estimated hazard ratio for hospitalization for

heart failure was 1.27 with an associated 95.1 percent confidence interval that excluded the null value of 1, which suggested a potential increased risk of hospitalization for heart failure associated with the use of saxagliptin.

The figure below shows an explanatory analysis of the cumulative probability of the occurrence of hospitalization for a heart failure event obtained from Kaplan-Meier survival estimates.

EXAMINE was an event-driven randomized prospective double-blind trial comparing alogliptin versus placebo in patients with type 2 diabetes mellitus, with a history of acute coronary syndrome within 50 to 90 days from randomization. Over 5,000 patients were randomized to alogliptin or placebo.

After a mean follow-up time of about 1 and a half years, about 620 primary MACE events were accrued and vital status was ascertained in 95 percent of randomized patients. A similar proportion of patients in both treatment arms

experienced a primary endpoint. The upper bound of the 95 percent confidence interval was 1.16 with a point estimate of 0.96, ruling out the guidance-recommended unacceptable increased risk for MACE with alogliptin as compared to placebo.

The contribution of each component of MACE is shown here. The proportion of patients who experienced each component of MACE was approximately balanced between treatment arms.

The EXAMINE data suggests an increased risk of hospitalization for heart failure for alogliptin as compared to placebo based on a non-measuuring point estimate. The top table shows the time to the first occurrence for hospitalization for heart failure.

The estimated hazard ratio for hospitalization for heart failure was 1.19 with an associated 95 percent confidence interval, 0.9 to 1.58. The figure below shows the cumulative probability of occurrence of a hospitalization for a heart failure event obtained from Kaplan-Meier survival estimates.

ELIXA was an event-driven, randomized, double-blind, prospective trial comparing lixisenatide versus placebo in patients with type 2 diabetes mellitus with a history of a recent acute coronary syndrome event within 180 days from randomization. Over 6,000 patients were randomized to lixisenatide or placebo.

After a median follow-up time of treatment of about 2 years, about 800 primary MACE-plus events were accrued and vital status was ascertained in 99 percent of randomized patients. A similar proportion of patients in both treatment arms experienced the primary endpoint.

The upper bound of the 95 percent confidence interval was 1.17 with a point estimate of 1.02, ruling out the guidance recommended unacceptable increased risk for MACE-plus with lixisenatide as compared to placebo. A contribution of each component of MACE-plus is shown here. The individual MACE-plus components were similar between treatment arms.

The EMPA-REG outcome trial was an event-

driven, randomized, double-blind prospective trial comparing empagliflozin versus placebo in patients with type 2 diabetes mellitus with mostly established cardiovascular disease. Over 7,000 patients were randomized to empagliflozin or placebo.

After an average follow-up time of about 3 years, over 770 primary MACE events were accrued and vital status was ascertained in about 99 percent of randomized patients. The EMPA-REG trial met the 1.3 goalpost and also demonstrated a reduction in the risk of MACE as compared to placebo.

A breakdown analysis of the MACE endpoint components show that CV death is the main component driving the differences seen in the MACE results. Time to CV death showed an estimated hazard ratio of 0.62 with an upper bound confidence interval of 0.77.

TECOS was an event-driven, randomized, double-blind, prospective trial comparing sitagliptin versus placebo in patients with type 2

diabetes mellitus with a history of mostly established cardiovascular disease. Over 14,000 patients were randomized to sitagliptin or placebo.

The published results show that after a median follow-up time of about 3 years, about 1,060 primary MACE-plus events were accrued and vital status was ascertained in about 97.5 percent of randomized patients. A similar proportion of patients in both treatment arms experienced a primary endpoint.

Overall, TECOS showed that there was no unacceptable increased risk for MACE-plus with sitagliptin as compared to placebo. The published results of the contribution of each component of MACE-plus is shown here.

LEADER was an event- and time-driven randomized, double-blind prospective trial comparing liraglutide versus placebo in patients with type 2 diabetes mellitus with a history of mostly established cardiovascular disease. Over 9,000 patients were randomized to liraglutide or placebo.

After a median follow-up time of about 4 years, about 1,300 primary MACE events were accrued. Vital status was ascertained in about 99 percent of randomized patients. A Cox proportional hazards model was used to test for noninferiority against the prespecified risk margin of 1.3 for the hazard ratio of MACE and to test for superiority on MACE if noninferiority was demonstrated.

Liraglutide significantly reduced the time to first occurrence of MACE with an estimated hazard ratio of 0.87 and a 95 percent confidence interval of 0.78 to 0.97. The contribution of each component of MACE is shown here. The estimated hazard ratios for each of the components were consistent.

SUSTAIN 6 was an event-driven, randomized, double-blind, prospective trial comparing semaglutide versus placebo in patients with type 2 diabetes mellitus with a history of mostly established cardiovascular disease. SUSTAIN 6 was designed to rule out a hazard ratio of 1.8. Over

3,000 patients are randomized to semaglutide or placebo.

After an average follow-up time of about 2 years, about 250 primary MACE events were accrued and vital status was ascertained in about 99 percent of randomized patients. No increased risk of MACE was observed with semaglutide as compared to placebo. The contribution of each component of MACE is shown here.

The time to first event of diabetic retinopathy complication was prespecified. The top table presents analysis results for the composite endpoint of diabetic retinopathy complications. The estimated hazard ratio was 1.76 with an associated 95 percent confidence interval of 1.11 to 2.78. This analysis showed evidence of increased risk of diabetic retinopathy complications associated with semaglutide.

The figure below is a Kaplan-Meier plot showing the imbalance in diabetic retinopathy complications throughout the trial. The observed probability of diabetic retinopathy complications

was higher in the semaglutide arm.

The CANVAS program included CANVAS and CANVAS-R. The two trials had a similar design and population. Both were event-driven, randomized, double-blind, prospective trials comparing canagliflozin versus placebo in patients with type 2 diabetes mellitus with a history of mostly established cardiovascular disease. Over 10,000 patients were randomized to canaglifozin or placebo in the CANVAS program.

The published intake rate of the analysis showed that after an average follow-up of about 4 years, over 1,000 primary MACE events were accrued and vital status was ascertained in about 99 percent of randomized patients. A similar proportion of patients in both treatment arms experienced the primary endpoint.

The published paper reported the trial excluded the 1.3 risk margin with a 95 percent interval of 0.75 to 0.97. The published results of the contribution of each component of MACE is shown here.

The CANVAS program showed there was a twofold increased risk for lower limb amputations associated with canagliflozin. The amputation event rate in the CANVAS program is shown here.

EXSCEL was an event-driven, randomized, double-blind prospective trial comparing exenatide versus placebo in patients with type 2 diabetes mellitus with a history of mostly established cardiovascular disease. Over 14,000 patients were randomized to exenatide or placebo.

The published results showed that after a median follow-up time of about 3 years, over 830 primary MACE events were accrued and vital status was ascertained in about 98 percent of randomized patients. A similar proportion of patients in both treatment arms experienced the primary endpoint.

The upper bound of the 95 percent confidence interval was 1.0 with a point estimate of 0.91, ruling out the guidance-recommended unacceptable increased risk for MACE with exenatide as compared to placebo. The published results of the contribution of each component of MACE is shown

here.

In summary, all trials met the recommendations for ruling out the prescribed excess risk of the 2008 cardiovascular guidance. It's unclear how the different trial designs and demographic characteristics contributed to the cardiovascular findings among trials.

The trends in cardiovascular safety were generally consistent within each drug class. In regards to safety, these trials raise unexpected safety signals not previously identified in phase 3 trials.

I'll now turn to highlighting the differences of trial characteristics of pre- and post-guidance development programs. This comparison will give context to the evolution of assessment of cardiovascular safety in trials preceding the guidance.

As examples, I'll discuss the development programs for saxagliptin, liraglutide, and alogliptin, since these programs were ongoing when the 2008 guidance was issued.

Here, we can see some examples of the pre-guidance, phase 3 demographic characteristics of patients enrolled in the development programs for diabetes drugs. The programs ranged in size but were generally above 4,000 patients. Patients with established cardiovascular disease were excluded in most cases, and, in general, there were few patients with cardiovascular disease. Patients were younger with a mean age below 60 years.

Average hemoglobin A1c was above 7 percent. The mean diabetes duration was below 10 years, and patients with severe renal impairment were generally excluded.

Now, contrast these with the patient characteristics for their respective cardiovascular outcome trial for these drugs. As we've discussed, the cardiovascular outcome trials were large and enriched with patients at risk for cardiovascular events. Whereas the phase 3 programs evaluated the relatively healthier diabetes population, outcome trials enrolled sicker diabetes patients as noted by the longer diabetes duration and additional

comorbidities such as renal impairment.

The differences in demographics and trial designs between pre- and post-guidance trials clearly affected accrual of cardiovascular events. In the case of alogliptin, there were only 18 cardiovascular events detected over the 26 to 52 weeks' duration of the pre-guidance phase 3 programs, while EXAMINE, the cardiovascular outcome trial for alogliptin, accrued 621 events over 1 and a half years of follow-up.

Here again, we see the contrast in the accrual of cardiovascular events between the pre- and post-guidance period. With a mean follow-up of 2 years, SAVOR captured 30 times more events than the saxagliptin phase 3 trials.

Lastly, here we see the cardiovascular events accrued for liraglutide in their phase 3 development as compared to LEADER. LEADER accrued over 1,300 events as compared to 38 events in the liraglutide phase 3 trials. These three examples show that the significantly higher number of events accrued in outcome trials provide more reassurance

in their assessment of the cardiovascular safety for these products.

In summary, following the cardiovascular guidance, drug products and the DPP-4 inhibitors, GLP-1 receptor agonists, and SGLT2 inhibitor classes conducted outcome trials to evaluate cardiovascular safety in patients with type 2 diabetes.

Although there was some heterogeneity between trial design and conduct, all trials were enriched with high cardiovascular risk patients. All trials demonstrated no excess cardiovascular risk, while some trials showed a cardiovascular benefit. This benefit was unforeseen, but with an unintended consequence from the guidance, which allowed us to generate data for indications beyond glycemic control.

Unexpected non-cardiovascular safety findings not previously observed in phase 3 trials were also detected. The evolution of cardiovascular safety assessments is notable when comparing the pre-guidance period to the post-

guidance period.

In particular, we have seen the clinical trial programs preceding the guidance provided a limited assessment of cardiovascular safety as compared to cardiovascular outcome trials following the guidance.

Now, as we're nearing the 11th year since the issuance of the cardiovascular guidance, we need to consider what we have learned and what changes, if any, are necessary. To this end, we welcome your discussion, and thank you for helping us ensure continued cardiovascular safety in patients with diabetes.

Clarifying Questions to FDA

DR. WILSON: Thank you very much.

We have some other slides here right at the end of this. Dr. Chong, are you going to reiterate those or shall we move directly to questions?

DR. CHONG: Those slides are for tomorrow.

DR. WILSON: All right.

I think we're now open to questions for the FDA. Please identify yourself or I'll recognize

you and then go forward. Why don't we start with Dr. Ellenberg?

DR. ELLENBERG: I have several questions. The first one is the use of A1c as a surrogate for microvascular disease, which was apparently established on the basis of two trials that were done some time ago, I assume with older antidiabetic agents.

There's always an issue with surrogates as to whether a surrogate based on one class of drugs will in fact be a surrogate for another class of drugs. And it looked like, for at least a couple of the cardiovascular outcome trials, the issue of whether A1c is a surrogate for even the microvascular complications is adequate.

Apparently, I'm assuming that microvascular complications were collected in all of the outcome trials. So I'm wondering whether any of those programs actually showed a benefit in microvascular complications or whether there were one or two that showed a risk, but the rest of them may have been neutral, which would even more call into question

of whether A1c is in fact an adequate surrogate for anything, any of the complications of diabetes.

So that's one question. Do you want to just deal with it? Can I ask the other questions? The others are simpler.

DR. WILSON: You have no competition at this minute, so go ahead while you've got at least one more, and then Dr. Newman?

DR. ELLENBERG: So Dr. Yanoff had a slide that mentioned symptomatic benefits of A1c, but I don't remember hearing what those were, and I would be interested in knowing if they were symptomatic benefits.

My third question was the standard of care used in these outcome trials; how varied were they? What were those standards of care?

DR. WILSON: Are you going to respond to that? You pushed your button.

DR. YANOFF: What is your preference? Hold responses until all the questions?

DR. WILSON: We can have some short questions and short responses. Why don't we do

that? We have another 10 to 15 minutes. We'll take a break. We'll be revisiting questions to FDA later. So sure; why don't you respond now? Go ahead if you can.

DR. YANOFF: The cardiovascular outcomes trials are designed to compare the drug in question versus a placebo, but this is all really on standard of care. So we expect that, in the placebo group -- at least we had hoped to when we initially came up with this idea -- that the control group's A1c would not be any different than the drug group because investigators would treat the patient to the glycemic goals that were appropriate for that patient.

In that case, the design of the trial is not one where one could evaluate the effect of A1c on microvascular complications because we wouldn't be expecting differences in A1c between the treatment arms.

So the designs did not expect to see that difference, and so the trials weren't set up to detect that differences. So the endpoint

collection was not designed to look at microvascular from a benefit standpoint. Safety data were collected for things like acute kidney injury, for example. But the trials were not required to collect, let's say, improvement on diabetic nephropathy.

The impracticality, what ended up happening was, for every trial we've reviewed so far, the A1c was not very similar between the two treatment groups. That wasn't supposed to happen and it's not explained. But the differences between the two arms were still smaller than what you'd expect if you were doing, let's say, a monotherapy study where you're adding the drug on to a diet and exercise-controlled patient versus placebo.

So the difference from the control group is either reflective of the addition of anti-hyperglycemic agents in the control group, which we did see a lot more of. So in every trial, there were more other treatments for diabetes added to the control group. They just didn't get their A1cs down enough to match what the drug group would have

gotten.

So I don't think that those trials could either adequately assess microvascular outcomes or even provide any type of insight as to the use of a surrogate for microvascular disease.

The question about symptomatic profound hyperglycemia can be associated with polyurea, polydipsia, some of the common complications of having high blood sugars where sugar spilling into the urine, by trading it down to below that level, you can actually help patients feel better.

That may not be as applicable in an area where we have so many treatments that these trials aren't dealing with patients generally in that range, but that's a theoretical direct clinical benefit of A1c or any type of glycemic measure, fasting glucose, whatever, that goes beyond using it as a surrogate.

Then what was your third question?

DR. ELLENBERG: How varied was the standard of care in these?

DR. YANOFF: The various standard of care,

these are worldwide international trials, and they varied by region. But in our statistical analyses, we did do subgroup analyses to look at whether region or other factors could have affected the outcomes, and we didn't find anything notable.

DR. WILSON: Next question, Dr. Low Wang?

DR. LOW WANG: Thank you. I wanted to thank the FDA presenters for the great overview that you provided. It really gives us a good context and background for today's discussion.

I had two questions. One, just to go back to how the FDA chose the thresholds for exclusion of unacceptable risk of 1.8 and 1.3, that was the first question, was it mainly a statistical design consideration, just kind of the reasonable size of the trial, or were there other points that were also taken into consideration?

The second question was, I think one of the big differences, before and after the guidance came out, was really the risk of hypoglycemia in the newer drug classes and the newer drugs. In the overview that was presented, there was no

information that was presented about hypoglycemia in these trials. So I wondered if you could comment on those two issues.

DR. YANOFF: The choice of the goalposts was generally what I presented to you with the addition, I think, of some contemporary studies that were going on at the time 10 years ago with NSAIDs and other areas that we're using similar numbers.

So the comfort level at that time with those numbers was also considered. It seemed to fit this whole idea of assessing cardiovascular risk for drugs that really had unexpected concerns.

Then the other, I already mentioned, the idea that we already had so many, we maybe shouldn't be accepting a doubling of the risk, although I know that was brought up in the committee discussion in 2008.

Beyond that, I will see if Dr. Thanh Hai has any comment, but I wanted to answer your second question first about hypoglycemia.

Could you rephrase your question, please?

DR. LOW WANG: Yes. The reason that I wanted to bring up the issue of hypoglycemia is that I think there's a big -- well, first going back to that surrogate idea or the average blood sugar that's indicated by A1c is that it really doesn't incorporate much information about hypoglycemia.

So I think prior concerns about risk associated with antidiabetic drugs may have had something to do with the incidence of hypoglycemia with those drugs, so both with sulfonylureas as well as insulin. So I think that one of the big differences in the newer classes is the reduced rate of hypoglycemia and the association of -- or maybe actually the causality of hypoglycemic episodes with a cardiovascular event; sudden death, for example.

DR. YANOFF: Yes. As you say, there were a few -- the newer products tend to have mechanisms of action that reduce the risk of hypoglycemia as compared to some of the older products and insulin. So those were not factors that were found

statistically to be related to the CV outcomes in the trials.

Does that answer your question?

DR. LOW WANG: I think we'll come back to this later. I do think it's an important issue.

DR. WILSON: Dr. Newman?

DR. NEWMAN: Thank you. I wanted also to thank the FDA for their presentations. I had two questions, and the first is about adjudicated non-cardiovascular adverse events. There have been many cardiovascular outcome trials, not just with these medications for diabetes, but medications, for example, for lipid reduction. And I wonder whether the FDA has done an analysis of investigator- or site-reported cardiovascular adverse events compared to adjudicated cardiovascular adverse events.

The second question is about Dr. Yanoff's slide 24, which lists some patient-years needed for trials to show cardiovascular safety. And I'm wondering what patient population you're using in this table.

DR. YANOFF: So I'll answer your second question first. It's a hypothetical population with a predicted 3 percent event rate. So I suppose that event rate could be generated by certain risk factors, either age or previous events. But really the point is that's how many events per year you would expect for that group, so it's a hypothetical population.

DR. NEWMAN: Not with patients with heart disease or a long duration of diabetes?

DR. YANOFF: It could be, but if you had, let's say, hypothetically, a high level of heart disease, it might increase your event rate over 3. So it's about how you balance all those factors to get an event rate of 3.

If you wanted to get a patient population that had a very high risk and a long level of disease, you might be able to get a higher annual predicted event rate in maybe fewer years, but this is just a hypothetical scenario.

DR. NEWMAN: Thank you.

DR. CHONG: To speak to your first question,

I just want to make sure I remember it correctly. The question was, have we looked at comparing results looking at investigator-reported terms compared to the adjudicated results?

DR. NEWMAN: Yes.

DR. CHONG: I'm not sure that we have done that detailed analysis across all the trials.

DR. ARCHDEACON: So I don't think we've done an analysis as you suggest, but I think it does get to a point that I was trying to make in my presentation. To me, the biggest difference is not the value of the adjudication committees, but just the approach to data collection and data curation.

Certainly, you could do the calculation that you're talking about. A difference would be, so now, investigator-reported potential MACE events are only a subset of the total number of MACE events. I guess you would say, okay, we could calculate sort of the truth table for if an investigator has reported an event, then it would also be included in the events that the committee would adjudicate, and we could see how often.

So we haven't done that; we could. I suspect from having looked at some of these data, that oftentimes those will be similar, but I think some of that has to do with the fact that now there's been better data collection and better attention in general. But it would be impossible to do a complete truth table because not all the events that ultimately get adjudicated were reported by the investigator.

DR. NEWMAN: Sometimes they were reported by the site?

DR. ARCHDEACON: So to be fair, in my slide, I think I didn't have things on the slide that speak to this that I'm addressing. So there could be an event, for instance, where the sponsor may identify a potential MACE event based on some terminology or based on an out-of-range laboratory, something like that, so that would have them then go and collect additional data. So there wouldn't be an investigator-reported term for that event. Now, admittedly, that's probably a minority of events.

DR. NEWMAN: Thank you.

DR. YANOFF: I can say, from my personal experience in one application looking at adjudicated heart failure events versus just MedDRA-reported heart failure events, they were roughly the same. That's one example of one program conducted a certain way, so I wouldn't want to make any generalizations about how successful that would be among any other trials.

DR. NEWMAN: I understand. Thank you.

DR. WILSON: We're going to take a break now. It's 10:30. We'll be back at 10:45. We'll come back to further FDA questions, but next up will be Dr. Ratner making a presentation. Thank you.

(Whereupon, at 10:31 a.m., a recess was taken.)

DR. WILSON: If we could have the committee members take their seats so we can be sure we're all ready to go. I won't mention names. I'd rather just say committee members.

Dr. Chong, you're going to introduce our

speaker? Thank you.

DR. CHONG: I'd like to welcome our next speaker. Dr. Robert Ratner is a professor of medicine at Georgetown University Hospital. Prior to that position, he served as the chief scientific and medical officer at the American Diabetes Association, where he played a key role in publishing clinical care guidelines, consensus opinions, basically everything.

His career has been spent studying patients with diabetes, treatments for diabetes, and on a side note, he was also involved in the 2008 advisory committee meeting.

Dr. Ratner, we welcome you to come back and kind of come full circle.

Guest Presentation - Robert Ratner

DR. RATNER: Thank you, Dr. Chong. Mr. Chairman and members of the committee, it's a pleasure and an honor to be back here with you. I am not representing any group or institution, and no one has provided any financial support to me for participation in this meeting. These are my

financial disclosures, which you heard earlier.

I go back to this slide that Dr. Yanoff showed. I did have the honor of participating in that 2008 meeting in July, as several members of this committee did as well. And I think that it's important to understand that any well-designed research that is well executed will give us useful information. That's a given.

The design of the guidance in 2008 really established, as Dr. Yanoff said, the goalposts, 1.3 and 1.8. Having been at that meeting, that debate was a bit arbitrary, but in fact, it was really set from practical reasons. How many events would it take to really see those events, and what could we actually accomplish during that?

Now, I was outspoken at the 2008 meeting, talking about the relationship between glycemic control and microvascular complications. That is an established fact. We clearly know the relationship from a number of different studies, not just UKPDS and DCCT, but a number of different studies where glucose levels were the

differentiating factor and a difference in microvascular outcomes.

The difficulty in those studies with macrovascular outcomes was the very low event rate and the fact that it took 30 years to see sufficient events to show a difference, but by that point, the randomization had long been broken.

But this isn't my first time speaking about cardiovascular outcome trials. In April of 2015, I also spoke at the EMDAC meeting. At that point in time, these were the points that I made. Over 130,000 subjects with diabetes entered into placebo-controlled clinical trials in the absence of a problem.

These were safety studies. These patients had to be at very high risk for cardiovascular outcomes so that the studies could be done in a reasonable period of time with an achievable number of subjects. And the simple fact was these were not patients that reflected the typical patient with type 2 diabetes.

The safety studies didn't test the

hypothesis. They simply said we're looking for unacceptable risk. And I posed the question of are we asking the right questions? Maybe we need to be looking somewhere else.

Now, these are the cardiovascular outcome trials that have been done and are currently underway. We are now up to 26 cardiovascular outcome trials, in a variety of different drug classes, utilizing over 190,000 patients randomized to placebo by design. So we really need to stop and think whether or not this is worth it. Is this the direction we want to continue going?

Now, recognizing that we do learn from any well designed, well-executed trials, what have we learned from the CVOTs? The CVOTs have clearly demonstrated that all of the drugs that have currently completed their studies show no increased risk of cardiovascular events. All of the studies since 2008 have been either neutral, or in fact some of the drugs have demonstrated reduced cardiovascular risk.

Well, that's good. It's nice that we can

show a positive outcome as opposed to a noninferiority outcome, which is the way all of these studies were initially powered.

We've also discovered some very interesting and very useful information in terms of side effects. Early on, the DPP-4 inhibitors and the GLP-1 receptor agonists were saddled with the concern of pancreatic safety, both pancreatitis as well as pancreatic carcinoma.

With the long exposure and the high numbers of individuals that have now been exposed, it's really quite clear that the risk lies with having diabetes, not with which drug you're being treated with. But as Dr. Yanoff noted, we also discovered some side effect issues, things of concern, whether it's congestive heart failure for the DPP-4 inhibitor class, or whether it's amputations for the SGLT2 inhibitors.

These become important learning lessons that we've gotten from these studies, and there's no question, the CVOTs have been valuable additions to our knowledge base.

So the question that was posed to me was should cardiovascular outcome trials continue to be mandatory? Understand there has never been a prohibition for a company to undertake a cardiovascular outcome trial. What the 2008 guidance did was to make it mandatory to perform these. And again, going back to the original guidance, it was to identify an unacceptable harm.

So the question is, is there a signal of harm from these classes of agents to treat diabetes?

This is a collection of the published trials. The one that isn't in here is the recently presented HARMONY trial. And what you can see is, whether you're looking at DPP-4 inhibitors on the left, GLP-1 receptor agonists in the center, or SGLT2 inhibitors on the right, the point estimates are always unity or to the left, the benefit of the drugs being studied. In no cases are the point estimates in favor of placebo. With all of these studies, there is absolutely no evidence of harm.

So the next question is, are the current

cardiovascular outcome trials generalizable? Let's take a look at the population with diabetes in the United States. When you look at cardiovascular disease, it occurs in about 21 percent of patients with diabetes. What's interesting is chronic kidney disease is even more common and in fact is easily the highest risk factor for the development of a cardiovascular event. But in fact, you're looking at 79 percent of individuals with diabetes who don't have cardiovascular disease.

So how does this play out in terms of the cardiovascular outcome trials? Well, here are the 4 GLP-1 studies. We're assuming a population with type 2 diabetes based on the previous CDC report about a year ago of about 24 million individuals.

If you look at the percent of individuals in the diabetes community who would meet the inclusion/exclusion criteria for participation in these studies, you're looking at 6.4 percent for ELIXA, 12.8 percent for LEADER, 11.8 percent for SUSTAIN, and the one study that didn't show statistical significance had 47.4 percent. This

was the percent of the diabetes population who might otherwise qualify.

What becomes even more important is if you look at those who have cardiovascular disease versus those who just had risk factors, that met the inclusion/exclusion criteria but they didn't have established cardiovascular disease, the benefit in LEADER and SUSTAIN was exclusively in those with established cardiovascular disease. It was not seen in those with just risk factors.

So the next question is, given the fact that we've got some positive outcome studies, actual statistical benefit when it comes down to cardiovascular outcomes, is it now ethical to withhold empagliflozin, or liraglutide, or canagliflozin, or albiglutide from the control arm?

In fact, should one of those drugs be included in the treatment arm as well, since the hypothesis is we don't know whether or not the treatment is going to be effective? How does that affect study design?

What is the standard of care? That was

asked earlier today. The standard of care is typically assessed by organizations who write medical guidelines. The American Diabetes Association has been doing it for 28 years.

Recently, the American Diabetes Association, together with the European Association for the Study of Diabetes, came up with new recommendations for the treatment of type 2 diabetes based upon the cardiovascular outcome trials.

Clearly, they learned a lot, and what they came up with was, among patients with type 2 diabetes with established atherosclerotic cardiovascular disease, SGLT2 or GLP-1 receptor agonists with proven cardiovascular benefit are recommended as part of glycemic management.

To withhold these therapies in essence goes against the standards of care. Not only that, the ADA and EASD identified which agents in which order based upon the cardiovascular outcome trials. If atherosclerotic cardiovascular disease predominates, either GLP-1 or SGLT2 should be used.

If you're using a GLP-1, liraglutide is

considered to be the first choice at the present time with semaglutide being next, but it doesn't have a label indication for cardiovascular benefit. And exenatide LAR is last because it didn't achieve statistical significance.

Within the SGLT2 class, the proven therapies are empagliflozin and canagliflozin. And the statement was made that empa appears to have a greater beneficial effect, and so that was put first.

However, there are some caveats. Number one, there is no evidence of cardiovascular benefit in those at lower cardiovascular risk, those who only have risk factors as opposed to having underlying cardiovascular disease. And number two, the combination of an SGLT2 inhibitor and a GLP-1 receptor agonist has not been tested in cardiovascular outcome trials at all.

So we don't know all of the answers there. However, we have to keep moving on. So the next question is, if it's unethical to withhold therapy in either the control arm or in the treatment arm,

what's the impact of allowing use of a proven effective therapy?

The study design might look something like this. In the control group, you've got standard of care that includes, by mandate and by study design, either liraglutide, albiglutide, semaglutide, empagliflozin, or canagliflozin. And one could even argue if semaglutide should be in there.

In the treatment arm, you'd have your new or old drug, whatever it is you're testing, and then the question is, do you have to also include one of those proven effective agents from an ethical standpoint?

So what's the impact of that sort of study design? Dr. Yanoff showed this power calculation based on the noninferiority margins that were decided in 2008. These are identical numbers to what she showed, so that the number of events required to have a noninferiority margin of less than 1.3 was 611.

Now, that was a placebo-controlled trial. What happens if you're now doing an active-

controlled trial and you're now looking at noninferiority? Because you clearly don't want to be inferior to an active agent that's been proven to be effective. So typically what happens is you cut the effect size by half. What that does is it takes your noninferiority margin down to 1.1, and that's 1800 events, almost a threefold increase in the number of events to be able to show noninferiority against an active agent.

Next question, should CVOTs be undertaken for primary prevention? Over 70 percent of patients with diabetes don't have cardiovascular disease. If we're going to treat the primary population, this is the primary population. Now, the data I'm going to show you are all post hoc, but I think they're informative in driving hypotheses moving forward.

Here are LEADER, SUSTAIN, and EXSCEL, and you're looking at the benefits seen in those individuals who have established cardiovascular disease at baseline versus those individuals who were older that only had risk factors. And what

you can see is that the point estimates for those individuals with established cardiovascular disease are all in favor of the treatment. EXSCEL failed to meet statistical significance, but the point estimate was clearly beneficial.

On the other hand, in all 3 studies, what you see is either a neutral effect or a point estimate that actually favors placebo in those individuals who were, number one, older, but number two, only had cardiovascular risk factors as opposed to cardiovascular events.

So if you begin to look at these sorts of post hoc data a little bit more closely, and now we include the SAVOR, which looked at a DPP-4 inhibitor, the DEVOTE, which looked at insulin, and the SUSTAIN 6 study, and the CANVAS study, which looked at SGLT2, what you can see, again, is that the effect appears to be on secondary prevention on the point estimate, even in those studies in which the statistical significance was not shown.

When you get down to primary prevention, individuals with no preexisting cardiovascular

disease, the closest you get is a point estimate of 0.98 or 0.99, and all of the confidence intervals exceed 1.3. When you're looking at primary prevention, part of the difficulty is the event rate. So if you look at these studies and you look at the annualized 3-point MACE in the primary prevention cohort, it's not surprising that you're not seeing statistical significance and you're not seeing benefit because the event rates ranged from 1.3 to 2.7 percent per year.

Now, that's really a far cry from where the event rates are in all the cardiovascular outcome trials. In the placebo groups of the cardiovascular outcome trials that the FDA just presented, the range of event rates was 7.4 percent to 14.9 percent. Those were the placebo event rates. And yet, in the primary prevention, what you see here is that you don't even come close to that.

But that's not surprising. If you look at CDC data, what you see is that cardiovascular events, acute MI and stroke, have been falling

precipitously since the mid-1990s. Part of this is the development and utilization of statins. Part of it is the use of ARBs and ACE inhibitors, but the fall actually coincides with the publication of the DCCT and the UKPDs as well.

Improved care has dropped acute MI by 68 percent and stroke by 53 percent, so that if you look at the rates now of acute MI and stroke together, it's 1.5 percent per year, exactly what was seen in the primary prevention cohort in those studies.

More recent CDC data also emphasizes the fact that patients with diabetes are living longer. Death rates from any cause have fallen, although they've stabilized over the last 5 years. Death from cardiovascular disease has fallen, coronary heart disease has fallen, and hospitalization for cardiovascular disease has fallen.

It's still greater than a matched control; I will grant you that. But people with diabetes are living longer and are living healthier than they have in the past.

So from these data, the CDC has stated that individuals with diabetes have increased longevity compared to the past and are now within 5 to 7 years of the life expectancy of a matched non-diabetic cohort.

That's good news. People with diabetes are doing a whole lot better, no question about it. What does that mean for where we go from here? What should the FDA do from here? I have a very biased view, but I'll share it with you.

There are lots of different options. Number one is return to the pre-2008 regulatory approach. That probably isn't going to happen and probably shouldn't happen, but in fact, there are a whole lot of very simple tweaks to the pre-2008 regulatory situation that would get us to where we want to go, and we'll talk about some of those as we move forward.

Number two, there's no question that if there's a signal for cardiovascular, or for that matter any other safety signal, in phase 2 or phase 3, the FDA has the right and the responsibility of

demanding the appropriate outcome trial; no question.

Number three, if a company wishes to get a label indication for cardiovascular benefit, then it is absolutely within the FDA's purview to require the appropriate study. If on the other hand, a company doesn't care -- and I would guess that those companies that are dealing with DPP-4 inhibitors probably at this stage of the game don't care -- then they shouldn't be mandated to do a study that's going to show a neutral effect.

So if we're going to be doing these trials, what can we do to make them more effective, learn more from them, and get through with less money? Well, there are lots of different trial designs that can be done, some of which are already being done by the FDA for other signals that have been seen.

These are the studies. The traditional RCT, as you can see, is far and away the most expensive, but it is considered to be the gold standard for comparative effectiveness. But for rare events,

what the FDA has done is to use registries, the observational studies. Whether it's looking at medullary carcinoma of the thyroid with GLP-1 receptor agonists or a number of others, there are ways you can collect that information if there is any signal whatsoever.

You can do registry-based RCTs, which are a whole lot cheaper and easier to come by, and you can do large numbers of outcomes. You can do pragmatic studies. The EXSCEL study was thought to be a pragmatic trial. It failed because it wasn't so pragmatic. They had a number of drop-outs and a number of drop-ins.

But there are choices in terms of study design that get you to the same point without the need for doing a large RCT. And one of these is an adaptable study design. An adaptable study design has predefined points where you can make changes in the protocol based upon the information that comes out. These changes have to be pre-defined, otherwise you lose the statistical value.

In fact, this is the direction that the

Patient-Centered Outcome Research Institute, or PCORI, has taken to be the primary methodology for really looking at comparative effectiveness, so that you can begin to see a lot of different alternative clinical investigation techniques to really try and get to the answer.

Then you can try differential statistical approaches. We've been stuck using Fisher exact tests and parametric tests forever. But that's not how we practice medicine. We can say that study A has a benefit of 45 percent and study object B has a benefit of 25 percent. Therefore, A is better than B. But 25 percent of the patients on B got a benefit, and that's how we begin to weigh judgments.

One of the statistical techniques that can be utilized to improve power to really get us to where we need to be is to introduce Bayesian statistics into our therapeutic trials. Here, what you're doing is you're taking priors, what do we know, what do we expect, and building that into the power analysis.

Now, what I want to do is I want to have you focus on the decrease in probability of the null hypothesis. From is before Bayesian introduction to no less than is with Bayesian statistical analysis.

What you're seeing is a remarkable improvement in the power of a study when you introduce the use of primers in study design. This reduces the number of individuals that would need to be studied and gives you a more powerful interpretive value to the results that is clinically applicable.

Finally, going back to my question in 2015, are we asking the right questions? What is it that's really important to people with diabetes? This was alluded to a bit in the earlier presentation, but in fact the diaTribe group has actually surveyed almost 3500 individuals, type 1's/type'2s not on insulin type 2's on insulin, asking them what's important to you.

What you see is that cardiovascular disease isn't even in the top five. Cardiovascular disease

is important when you get it, and there's no question that everybody is concerned about a stroke, and a heart attack, and dying. But with a life expectancy that's now significantly improved, how you live your life is becoming far more important than it has been in the past.

So looking at hypoglycemia, looking at glucose in range, A1c, how you have to treat the diabetes is clearly playing a much more important role.

Now, that's clearly been included in how the ADA and the European Association for the Study of Diabetes have approached their standards of care. The consensus recommendation as the choice of medication added to metformin is based on patient preference and clinical characteristics.

Important clinical characteristics include the presence of established atherosclerotic cardiovascular disease, other comorbidities such as heart failure or kidney disease, and risk for specific adverse medication effects, particularly hypoglycemia and weight gain, as well as safety,

tolerability, and costs.

These are the issues that are important to people with diabetes. These are what the standards of care are now focusing on. And you can see from the standards of care that they isolate all of these. If you have established CKD or heart disease, then you're on this left-hand side of the algorithm, turning towards either GLP-1 receptor agonists or SGLT2 inhibitors.

If you don't have atherosclerotic vascular disease and the concern here is hypoglycemia, then these are the treatment options that are available. If you don't have heart disease or kidney disease but weight is the major concern, then these are the appropriate treatment options. And if cost is the major factor, then you change your therapy once again. This is true patient-centered care that takes into account all of the different characteristics.

Final thoughts, diabetes is a chronic disease, and people are now living longer and are having to put up with the therapies, not just

waiting to die. All-cause mortality and cardiovascular mortality are falling among people with diabetes, with a marked increase in life expectancy.

Increased time living with diabetes puts a premium on quality of life issues, and the goal is to improve the lives of people with diabetes, not just to increase their longevity.

There are several follow-ups that the FDA can take from this meeting. Number one, widen the inclusion and narrow the exclusion criteria on regulatory trials so that you have older individuals at higher risk. There's no question that cherry-picking healthy patients for regulatory trials clearly skews the results.

Two, demand outcome trials for whatever endpoint if there is in fact a signal and require outcome trials in those situations in which a sponsor wants a label indication. Thank you very much.

Clarifying Questions for Dr. Ratner

DR. WILSON: Thank you.

Questions for Dr. Ratner? Yes, Dr. Grunberger?

DR. GRUNBERGER: Yes, Gerry Grunberger.

Bob, just a clinical question; and I know you were not the coauthor of this statement, but the ASCVD versus CVD versus peripheral artery disease is all the same? How specific is atherosclerotic cardiovascular disease versus cardiovascular disease, and does it include also PAD?

DR. RATNER: Dr. Grunberger is correct. I was not part of the ADA EASD writing group. I had nothing to do with it. I don't have a complete answer for you. In general, CVD is the most restrictive. ASCVD tends to really focus exclusively on coronary disease, and PAD is usually not included at all in these discussions.

DR. EVERETT: This is Brendan Everett, cardiologist. I think we should clarify that to the cardiologists in the room, cardiovascular disease is actually the broadest and typically includes heart failure as one of its key endpoints.

Atherosclerotic cardiovascular disease does include coronary disease, but also includes atherosclerosis of other vascular beds, including, for example, the carotid and the lower extremities. So it would include typically PAD as well.

DR. RATNER: Thank you.

DR. WILSON: Thanks, Dr. Everett.

Who was next? Dr. de Lemos?

DR. DE LEMOS: I want to follow up on a question that was raised earlier. We're talking potentially about rolling the bar back, but I still want more discussion about why an A1c reduction for a safe drug is sufficient to enter the market for diabetes in 2018 in a crowded field with lots of drugs.

Then in a second, I want to come back and pull your slide up, where you talk about the active comparator trials and talk about that a little bit.

DR. RATNER: So going back to the issue of A1c as a surrogate, it's important to understand exactly how the DCCT and the UKPDS were done. In DCCT, there were 1440 individuals with type 1

diabetes who were followed in two stratifications: one, those individuals who have no evidence of any complications at baseline, and the second strata was those individuals who had some indication of microvascular disease.

The study went 9 years to show a statistically significant benefit in 3-point progression of retinopathy as measured by fundus photography and showing a reduction in the progression of renal disease with a doubling of serum creatinine. It took 1440 patients 9 years to show those statistical changes.

At that point, it was unethical to maintain the control group, so they converted this to an observational trial, and both groups actually came together in terms of hemoglobin A1c. To the benefit of the NIH and the investigators, that's taught us an enormous amount. It took us 30 years to see sufficient cardiovascular events in the DCCT, to show a difference between the two groups.

The difference in microvascular events persisted even though the A1cs came together, and

that's been termed a memory effect or, in my mind, it's the area under the A1c curve. It's the exposure of the beds to glucose. But it took 30 years for that.

The UKPDS actually went 17 years to show a microvascular benefit, and then they followed for an additional 8 years, out to 25 years, to show a cardiovascular benefit.

So if in fact you want hard outcomes in a chronic disease, you're talking about huge numbers of patients in a randomized trial over their lifetime because events don't occur quickly.

There is an empirical need for a surrogate, and A1c has proven time and again in all patient populations, for Asians to northern Europeans to Australians to Indians to the U.S., to correlate in controlled trials a reduction of A1c results and a reduction in microvascular complications. To go back on that at this point would be impossible to generate sufficient data to approve anything.

DR. WILSON: Was that a satisfactory answer?

Dr. Robbins?

DR. ROBBINS: Bob, thank you for a really great talk, very thought-provoking and well done. My question stems from the issue of the increased longevity and longer life of type 2.

This is a dynamic of, on the one hand, having better treatment of atherosclerosis, and hypertension, and so on, but on the other hand, we have really a tsunami of kids that are coming down with type 2 diabetes, and we also have fatty liver disease, which is really just beginning to be talked about as yet another epidemic that can bring the healthcare system to its knees.

I'm not sure what the answer is to that, but I don't think we should get up and celebrate yet that the epidemic is over. I think we're seeing it morphed, and I'd love to hear your comments about that.

DR. RATNER: Dr. Robbins is absolutely correct. The number of individuals with diabetes is still going up. We should take pride in the fact that we've learned how to better care for them and to improve the quality as well as the longevity

of their life, but we're not done by any means.

The issue of kids with type 2 diabetes is unique. The studies that have looked at type 2 in kids have been really depressing. The TODAY trial shows that basically nothing works. The best predictor of glycemic control in adolescents with type 2 diabetes was whether or not they were incarcerated. And that really goes to the heart of how difficult it is to deal with this patient population.

So we need to keep working. I think it raises the issue, though, of opportunity costs. Where should we be putting our money? There's a limited amount of money for clinical research, whether it's with NIH or whether it's within sponsors' budgets. We need to be asking questions like how do we deal with NASH and NAFLD? What do we do about childhood obesity and trying to stem it early on? What's going on in utero in terms of establishing satiety points in the hypothalamus?

Those are opportunity costs that are lost if all of the money is going to cardiovascular outcome

trials.

DR. WILSON: Next, Dr. Everett?

DR. EVERETT: Thank you for a very thought-provoking talk, Dr. Ratner. I want to see if you can maybe make the case to me -- and this is perhaps a follow-up on Dr. de Lemos' question -- for why hemoglobin A1c, given what you've said about it, should be the only surrogate endpoint required for the approval of a diabetes medication, specifically because we all know diabetes, particularly the cardiologists here, as a disease that we often talk about the microvascular complications, which is really related to the A1c, as you just outlined, and also extensive macrovascular complications, and in particular, not just the atherothrombotic ones that we are used to thinking about in which the 2008 guidelines focused on, but on heart failure as well, which I think is increasingly prevalent and is a tremendously morbid condition for patients with diabetes.

So why is it adequate to use a surrogate endpoint to approve a drug to treat this syndrome,

diabetes, or this disease process, that has multiple manifestations, one of which, and an important one of which, is microvascular disease. But certainly is not all of the manifestations or even the one that causes the most mortality from that disease.

Why should we just go with A1c and not with other endpoints, be they surrogate or hard endpoints?

DR. RATNER: So first of all, I am entirely in favor of expanding the endpoints because A1c is not sufficient. As Dr. Wang mentioned earlier, hypoglycemia, I think, needs to be a clinical endpoint in the regulatory process; that I don't believe any drug that increases the risk of hypoglycemia can get approved or should get approved moving forward. That's an unacceptable risk factor. And we now know that hypoglycemia, even mild to moderate hypoglycemia, has major impact.

Individuals who have measured glucose values independent of symptoms below 3 millimolar have

increased morbidity and mortality. That's very clearly defined. What the mechanism is remains to be seen, but we know that it is a marker for morbidity and mortality. So yes, I think hypoglycemia needs to be part of that as well.

I think it's probably unacceptable for a diabetes drug to exacerbate the underlying obesity that's resulting in the expansion of the population. So weight change probably ought to be an independent characteristic in the evaluation of a drug.

The question of what about heart failure, what about atherosclerotic cardiovascular disease that causes death, clearly, they're important. I'm not trying to minimize their importance. But even when you begin to look at the prevalence of these, you're talking about less than a quarter of the patients with the underlying disease.

In those patients, it's critically important, and that's why I fully agree that anybody who wants to get a label indication, whether it's for heart failure or atherosclerotic

cardiovascular disease, needs to do these studies.

The question is the mandate. The question is should everyone have to do this for a population that encompasses less than 25 percent of the population with diabetes.

DR. WILSON: Dr. Chong, you had a comment or a question?

DR. CHONG: Yes, I just had a quick comment. I'm getting the sense that there's a lot of discomfort with A1c as a surrogate, and I think it might be worth hearing from some of the endocrinologists on the committee to get their perspective on the value of A1c.

I don't know if Dr. Grunberger, Dr. Fradkin, Dr. Low Wang?

DR. GRUNBERGER: If I can comment, Bob has been part of that party. And as you know, FDA has been very interested in expanding this hard area, going beyond A1c. And Bob showed a diaTribe survey, and it showed that patients are more interested in time and range, and of course by definition also means reducing time in

hypoglycemia.

So there's no question that is the mind of the endocrinologist, if I can speak to them, that the time should be the preferred way to look at the glycemic part of diabetes mellitus. But what you brought up, and other people I guess are thinking about, is our definition of diabetes mellitus wrong?

As long as the definition is death of disease and hyperglycemia, you are stuck by definition in looking at glycemia. Whether A1c is the proper measure, you can debate. For the life of a patient, I don't think A1c is what matters. It's the quality of life, i.e., lack of hypoglycemia, maximize time and range.

A question is should we broaden the definition? This is beyond, obviously, this meeting, but I think that's debatable, is the current definition of diabetes mellitus the proper one?

DR. WILSON: Any of the other endocrinologists? Dr. Fradkin? Judith?

DR. FRADKIN: I just want to remind people that it's not only the long-term complications that are important in terms of glycemia, but people whose A1c is over 9 are really ill. I mean, they have lack of energy. They're dehydrated. They have increased risk of infection. So I mean, there is value per se to lowering glycemia.

I think that the issue then in terms of meaningful outcomes relates to the time course of complications. The benefits in terms of being well-being and acutely feeling well, that you see immediately when you lower somebody from a very high glucose down into something closer than the normal range.

But I think that different interventions have different time courses, and most people now are going to live 20 years with diabetes. So you're going to see people developing microvascular complications. Diabetes is the leading cause of end-stage renal disease. That continues to be the case even though the care of diabetes has improved.

Renal disease is also a contributor to

cardiovascular disease, but the question is, we have very strong evidence that Bob talked about in terms of long-term control reducing microvascular complications. To re-demonstrate that in every clinical trial is not going to be feasible. You would have to follow people for at least 10 years in a clinical trial.

But even DCCT and UKPDS, which looked at the benefits of taking somebody from a relatively high A1c to a more normal A1c, for those studies, the benefits were very, very large. But in some of the cardiovascular trials that looked at glycemic control for cardiovascular disease and didn't show a benefit, per se, on cardiovascular disease, when you followed out the people in ACCORD or ADVANCE, you saw benefits even lowering from an A1c of about 8 to an A1c getting below 7 in terms of microvascular disease.

So I think it's not just the studies that used insulin or sulfonylurea to lower A1c that have shown microvascular benefit. It's studies like ACCORD, which used all the classes and did

demonstrate -- and these were fairly short-term studies, unlike the DCCT and the UKPDS, but we saw benefits.

DR. WILSON: So thank you both.

We have Dr. Ratner behind the podium. Let's have the questions directly toward his presentation. We're going to have plenty of time for discussion, similar to what's happened in the last few minutes.

So a specific question, Dr. Robbins, do you have for Dr. Ratner?

DR. ROBBINS: I want to address the issue of the A1c as the marker. Is that okay?

DR. WILSON: Can we come back to that one maybe? Let's get the things expressly on his presentation.

DR. THANH HAI: Dr. Wilson?

DR. WILSON: Sorry. Dr. Burman, go ahead.

DR. BURMAN: Dr. Ratner, thank you for an excellent discussion. The studies that you reviewed were quite extensive and quite nice, and had a quite nice review. But in many other

diseases, we're trying to compare real-life experience versus people who participate in a trial, and there's usually or frequently a marked difference in the baseline and follow-up demographics of those patients.

Are there really any studies looking at real-life patients in diabetes to see how they compare with the baseline demographics in the original studies? In our practice, my practice, the hemoglobin A1c seems to be a lot higher than were entered into the study, and the cardiovascular incidence rate seems to be a lot higher.

DR. RATNER: Probably the best study in that regard is the CVD REAL study, which is a real-life clinical trial looking at SGLT2's --

DR. WILSON: Are you finished, Dr. Ratner, or were you going to say more?

DR. RATNER: -- looking at canagliflozin in a real-life population. The outcomes of that have looked to be very similar to the CANVAS trials. So the demographics are different, but at least in that real-life trial, the outcomes are consistent.

I know that Novo Nordisk is also undertaking a real-life experiment looking at liraglutide as well.

DR. WILSON: Dr. Kushner?

DR. KUSHNER: Yes. Thank you for a very thought-provoking talk. Two things; one, and follow up of this question about the use of adaptive trials or a registry type randomized trials, have any of these been done in parallel with randomized clinical trials? In other words, is there any incidence of a registry going simultaneously and seeing what these differences are?

DR. RATNER: Again, the CVD REAL trial would be the only one that I know of.

DR. KUSHNER: The other problem is, there was a class of drugs, one drug, that the hypothesis was testing whether raising HDL would limit outcomes, and torcetrapib was the drug and went all the way through. It even lowered LDL by 25 percent. And then suddenly, there was an increase rather late in the trials. There was an

increased incidence of death. And that wasn't the case necessarily with the other drugs in the class.

How do you pick up the next drug that comes up that may be in the same class or similar class, that may in fact have a completely different set of outcomes if you drop the outcomes trials?

DR. RATNER: I think we return to the safety component of all FDA registry trials. You're looking for signals early on. If there is any sort of a signal, then it is absolutely within the purview of the FDA, and it is their obligation to request further data in that regard, and there's nothing to preclude that.

I think that requiring extensive studies of drugs that have absolutely no indication of a problem has become problematic. I think to Dr. Everett's point, looking at a hypothesis-driven trial to show benefit is absolutely worthwhile. The question is, should it be mandated for everyone?

DR. WILSON: Mary Thanh Hai, you want to make a comment?

DR. THANH HAI: Thank you. Mary Thanh Hai, FDA. I just want to revisit or bring up these questions about hemoglobin A1c and its surrogacy in the approval of drugs. It's really not the objective at this advisory committee to discuss that. We can go back and look at the discussion points and the voting question. It's about looking at the CV guidance and what direction do we go here.

Antidiabetic therapies are approved for the indication, as Dr. Yanoff put in her slide, an adjunct to diet and exercise to improve glycemic control. Hemoglobin A1c is a measure of glycemic control, and you've already heard up to this point why it has gotten to the point of being accepted as a surrogate for glycemic control. That indication won't change.

I think that the topic that's been raised, the question, is really important, and the FDA has not dismissed that. As you've heard, we've had workshops outside of this to talk about measures beyond hemoglobin A1c, but for the purposes of this

meeting here, it is beyond the scope.

DR. WILSON: Thank you very much.

We have a whole bunch more questions for you, Dr. Ratner. Are you doing all right up there?

DR. RATNER: Doing just fine.

DR. WILSON: Can we get you a glass of water? We're not going to let you off the hook here very soon? Are you all right?

DR. RATNER: I'm good. Thank you.

DR. WILSON: Next, we have Dr. Low Wang.

DR. LOW WANG: Thank you so much for that presentation. I really thought you articulated some very important points. I had a question in terms of trying to incorporate more primary prevention patients with diabetes, so the question of pragmatic trial design.

One of the points that you made was that possibly the EXSCEL trial was not so good or didn't turn out so well because there are a lot of drop-ins and drop-outs. My kind of understanding of pragmatic trials is that's the whole points. We expect drop-ins and drop-outs in real life.

So is the only answer to expand the trial size, the sample size, or are there other thoughts that you have, or comments, on the pragmatic trial design?

DR. RATNER: The issue surrounding a pragmatic trial really is going to the initial study design and what the primary characteristics are going to be that could alter that trial. Those need to be pre-defined, and they need to be binary decisions, essentially. If you hit this endpoint, if you have this outcome, then you will do something, and you have to pre-define what that something is.

The difficulty with the EXSCEL trial -- there were a lot of problems with the EXSCEL trial -- was that they really hadn't dealt with all of the other issues that could come up. And one of the concerns is something that all of the trials moving forward are going to have to face, and that's drop-in of active agents into the control arm.

So if in fact you've got the dapagliflozin

study that's going on right now, if the control arm has a disproportionate number of individuals who get treated with liraglutide, that's going to decrease the event rate in the control group and may render the outcome noninferior. But that's the difficulty when you start including an active agent into the trials. And I tried to point out what the impact on the event rate would be there.

DR. WILSON: Thank you.

Dr. Fradkin? No? Dr. Ellenberg?

DR. ELLENBERG: Yes. Thank you very much. That was a very, very thorough and informative presentation. My question has to do with the safety signal issue. You had said that, certainly, if there are safety signals that arise in the standard trial, they should be followed up. And I doubt that anybody would disagree with that.

But the size of the trial itself depends on how readily you will be able to even see a safety signal. So it looked to me, from Dr. Archdeacon's presentation, that prior to the guidance, these trials might have been only a few hundred patients

each, which would be adequate to detect a difference, or I guess they were probably active controlled trials, in hemoglobin A1c, which is a continuous endpoint.

When you have a continuous endpoint, you have lots of information and you need smaller trials. But a trial of only several hundred people, you are going to limit the kind of signals you can detect. And in a disease like type 2 diabetes, where there are 1 and a half million new cases every year, a question is what kind of signal ought we be able to detect?

Is it adequate to detect only a signal that might suggest increasing something by a factor of 6 or 7, which is what you might get for a not super common condition; or do you think there would be some value to thinking about how many people we need to be able to detect certain signals of importance? That wouldn't be definitive, but it would be a signal to carry on and do further work.

DR. RATNER: I've not done the calculations to give you a number needed to harm in that regard,

and I really can't comment on that. I think, though, that signals can come from a lot of different places.

For example, there's a requirement for liraglutide to have a registry of medullary carcinoma of the thyroid based on preclinical findings, nothing that was ever seen in a human being. And in fact, the relevance to human beings in terms of the biology has even been raised.

I think leaving the FDA with the latitude of defining and identifying signals is critically important, and I would trust them to really identify what is sufficient to require another study specifically to examine that outcome

DR. WILSON: Next is Ms. McCollister-Slipp.

MS. McCOLLISTER-SLIPP: I'm going to try to say this staring at you while also talking into the microphone. One question I have -- and just again, for context, for those on the committee who don't know me, I'm a type 1 diabetes patient. I've had it for 32 years; have all the microvascular complications. But I'm also the daughter of type 2

diabetes patients, one of whom is in pretty severe vascular dementia following a series of strokes. The first stroke that he had was as a result of induced by severe hypoglycemia.

One question I had, given the data that we have from ACCORD and ADVANCE and some of the data that's continuing to be collected or generated connecting hypoglycemia events and cardiovascular events, are we at a point, at this point and the research that's been conducted, in concluding that -- are we at a point where we could potentially use hypoglycemia as a potential surrogate marker for cardiovascular risk?

DR. RATNER: There's no question in my mind that hypoglycemia, even asymptomatic below 3 normal, is clinically and statistically critical. It's important. It needs to be identified. It needs to be dealt with, the cause of the associated morbidity and mortality.

That morbidity and mortality is epidemiologic. It's not directly causative. There are a few cases where causation has been shown.

The best is a dead-in-bed syndrome, an individual with type 1 diabetes who is wearing an insulin pump and wearing a CGM, a continuous glucose monitor simultaneously, and he was found the following morning dead in bed. You can see the fall in his glucose during the night and track it perfectly to his demise.

There is good evidence from experimental studies from stepped hypoglycemic clamps, looking at simultaneous EKGs and looking at arrhythmogenicity and setting up hypoglycemia-induced arrhythmias, that are, interestingly enough, different during the day and during the night. They're completely different.

So there's a lot of information that is unknown. The relationship is clearly there. I don't think that we're at a point in time where we can say that hypoglycemia is a surrogate for cardiovascular outcomes. We would need to have simultaneous CGM in all of those subjects in order to really be able to prove that.

DR. WILSON: Thank you.

We're going to have time for a couple more questions, but one for Dr. Ratner.

Are you available this afternoon after Dr. Green's, and we have you back up for further questions?

DR. RATNER: I'm going to be available.

DR. WILSON: Next is Dr. de Lemos.

DR. DE LEMOS: (Inaudible - off mic). I'd like to push back on two things, one is feasibility, and just get your response to the fact that 200,000 patients that have been enrolled this fast seems to argue that these trials are feasible.

The point I want to get your thoughts on is you sort of made a case that you thought the active control studies would be almost impractical because of sample size and event rate. But that's setting a completely different bar than the agency does now.

Why couldn't the agency require an active control, but set the noninferiority margin -- if you showed noninferiority versus a known efficacious active control, then by definition that

would be a higher bar than is present already, and you could add a separate bar if you wanted the cardiac indication.

So it seems like it would be really feasible to me, within the context of what's already being done, to simply require that the control arm has drugs that we know have a cardiovascular benefit. When you go down to that noninferiority -- you don't have to go down to a noninferiority margin of 1.1.

DR. RATNER: This is the slide that you were talking about.

DR. DE LEMOS: Yes. That obviously creates a completely different construct, but that's not necessary. If you were noninferior to empagliflozin, you've already met a higher bar than you did against usual care of placebo in the current era.

If you wanted to demonstrate that you maintain the effects of empagliflozin, then you could have a higher bar for that, and then you wouldn't have anywhere near the sample size

requirements that are scaring people off from these sorts of studies.

DR. RATNER: I'm not a statistician. I don't play one standing at the podium, either. My understanding is that noninferiority studies are actually more difficult to power than superiority studies are. Even if you're looking at an active comparator that is quote, "raising the benefit," that all you have to do is meet noninferiority, you still have to have enough events to have that confidence interval to say there's no difference; that it's not noninferior.

I think that unlike cardiovascular studies in which you recruit out of the CCU or you recruit out of the cath lab, and you put them on one drug, and you kind of leave it alone, the dynamics of diabetes in terms of day-to-day fluctuations in glucose, in terms of complex meds, would make that very, very difficult to be able to recruit and retain those patients. It's the retention that becomes the biggest problem because of the drop-ins and the drop-outs.

DR. WILSON: We're going to have one more question before lunch. Tommy Wang?

DR. WANG: Just two brief comments or clarifying questions about the pragmatic trials. The first clarification; adaptable, despite the name of the trial, I don't believe that's an example of an adaptive study design. That's a good example of a pragmatic trial, but there's nothing particularly adaptive about the study design. That's just a clarifying point.

My question, your comments about the potential virtue of pragmatic trials and how they may be a way to lower the cost are appreciated. In your view, would you agree there's nothing in the 2008 guidance that prohibits companies from doing pragmatic trials, perhaps with the exception of the first point about prospective adjudication committees?

Is that an accurate statement?

DR. RATNER: I believe so, particularly in view of the fact that it EXSCEL was proved by the FDA to meet the CVOT requirement. And again, they,

whether correctly or incorrectly, are calling it a pragmatic trial.

So yes, I think it would be within the purview of the 2008 guidance.

DR. WILSON: All right. Why don't we take a break for lunch? We're going to come back at 1:00. We're on schedule.

DR. ELLENBERG: Can I just say something about active controls, the active control trials?

DR. WILSON: Dr. Ellenberg, you'll get the last comment before lunch.

DR. ELLENBERG: It's very short, but I don't think we're talking now about active control trials versus placebo-controlled trials. All the prior trials were active control trials. Nobody had just a placebo. They had placebo plus whatever was the standard of care at the time. And if we add new agents to the standard of care, then we'll still have active control trials.

So we're not talking about going to some new paradigm of active control versus placebo control. They've all been active control trials.

DR. DE LEMOS: I think we are.

DR. WILSON: Dr. de Lemos, can you wait? He can wait.

What we're going to do is we're going to break for lunch. Then Latoya, after lunch, it's Dr. Sabatine first? Is that right?

(Dr. Bonner gestures yes.)

DR. WILSON: Then Dr. Green, and then we'll come back to Dr. Ratner for any follow-up questions. We have your name down if you raised it for Dr. Ratner, so we'll come back.

Thank you. 1:00.

(Whereupon, at 12:01 p.m., a lunch recess was taken.)

A F T E R N O O N S E S S I O N

(1:00 p.m.)

DR. WILSON: We're going to go ahead and start. Our next presentation is by Marc Sabatine. No, excuse me.

Oh, you're going to introduce him? Dr. Chong is going to introduce our next speaker. I apologize. Go ahead.

DR. CHONG: I'd like to welcome our next speaker. Dr. Marc Sabatine is presenting on behalf of the Thrombolysis and Myocardial Infarction Study Group or TIMI Study Group. He is a professor at Harvard University and the chair of the TIMI Study Group, and his career has been dedicated to studying and improving the quality of life of patients with cardiovascular disease. To that end, he's led several cardiovascular outcomes trials and today is going to be sharing his perspectives and expertise.

Thank you, Dr. Sabatine.

TIMI Presentation - Marc Sabatine

DR. SABATINE: Fabulous. Thank you for that

kind introduction. It's a pleasure to be here and see many friends and colleagues here. I should note that the opinions I present are my own. I haven't received any financial support to come here. And here are my relevant disclosures. We have many dance partners.

I was asked with covering a few topics. I am showing them here in outline format. I do want to make sure I leave plenty of time for question and answer given it's such a dynamic group. One issue will be just to outline the work that goes into a traditional randomized, controlled cardiovascular outcomes trial.

I will touch on some other sources of data that the groups may look at in terms of randomized-controlled trials, but not dedicated CV outcome trials. I'll talk a little bit about observational data, a topic that came up in one of the earlier talks, then touch on streamlined or pragmatic cardiovascular outcomes trials, and maybe try to tease out a little bit what those terms might mean.

Then end with two challenging issues that were nicely covered in Dr. Ratner's talk as well; what's the equipoise? How does that shift as data emerge in a drug class? And then as we think about the next wave of trials, what would be the optimal appropriate safety and efficacy endpoints?

Let me start just by talking about the work that goes into these trials. There are obviously a lot of people who come together to do it. Obviously, there's the sponsor for the trial who has developed the drug. There typically is academic leadership, and that will include a global principal investigator and executive and/or steering committees consisting of a mix of experienced thought leaders in the field, as well as perhaps national lead investigators from the various countries that are contributing to the trial.

There will be a clinical events committee if there's going to be event adjudication in the trial. There will obviously be an independent data-monitoring committee to look after the safety

and well-being of the participants.

There will typically be some contract research organization that will play a role, and what that is might vary from trial to trial, but for example, monitoring or other aspects, and many other vendors who may play a role in the trial from things like data management and laboratory assessments, et cetera.

There are obviously a whole series of key documents and familiar to everyone here, I think, in the room. And all of these documents then undergo review and feedback from regulators. Regulators on different sides of the Atlantic may not have the same view for what would be an optimal analytic approach. And then there's also review from the ethics point of view from the IRBs and all the investigators opining as well.

The start-up for any of these trials then involves a series of decisions, what country one might do the trial in. Obviously, it's important to have good representation in the United States. The challenge, of course, is that the cost of

living is very high in the United States, and therefore doing the trial here is very high and other factors that make conducting trials in the United States challenging.

There will be other countries involved. Within each country, then one needs to pick the sites that one thinks will do a good job and evaluate them.

People have talked about some of the metrics for the trial in terms of patients; for example, staying on study drug or not being lost to follow-up, and a lot of that comes from picking dedicated investigators.

So there's a lot of work that goes into that. All that needs to be, then, set up with site contracts, and then the regulatory and ethics approval, and developing all of the other various parts of the trial to start up the trial.

The sites obviously during the trial do a huge amount of work; the core of the trial for screening, and enrolling patients, and dispensing study drug, and bringing back patients for

follow-up visits. I'll talk in a little bit about some other models for getting data on patients, which might be embedded within a given healthcare system.

Then assessing specifically for efficacy and safety outcomes, this topic has come up earlier. Can you just rely on healthcare system data that may be useful in certain situations? But if there are specific questions that you wanted to ask and answer, you may want those questions to be specifically asked of the patients by people who understand the questions and how to interpret the answers.

A ton of data entry, compiling the efficacy outcome packets for the CEC to have centralized review; all the necessary expedited safety reporting; all the monitor visits to make sure the sites are doing all of the bullet points above; and then periodic IRB submissions.

At the central end, what do we do at the trial, it's not the watchmaker approach. It's not winding it up and letting it go, and then 5-6 years

from now seeing what happens. We obviously track the pace of enrollment into the trial. We assess who are the patients going into the trial. There were several points raised about the types of patients and their characteristics.

Obviously with blinded aggregate data, we look at the efficacy event accrual. We look at safety outcomes to see if there's some signal there that, even though, again, we're not seeing it by treatment arm, we have a concern if the rates are high, and we might need to make sure we have greater granularity for that.

There's a huge amount of data cleaning that goes into place, monitoring the sites, both remote and on site, putting the packets together, adjudicating them, getting DMC reports done, and huge work on retention efforts.

Then I'll get to this point towards the latter part of my talk, assessing external events relative to the trial. These trials obviously don't happen in a vacuum. You heard very nicely in Dr. Ratner's talk about the data that would

accumulate. That informs us not only in terms of what the next trial should be, but also is relevant for ongoing trials. Then to close that, we try to bring all of this data together, trying to minimize missing data, and eventually then locking the database.

So all of that can be summarized by saying it's a huge amount of work. Many thousands of subjects, thousands of researchers, spread out among 1,000-plus sites, depending on the size of the trial, in dozens of countries, thousands of staff at the sponsor and the CRO, and this can obviously cost hundreds of millions of dollars and takes many years.

Is that a good return on investment? Because there's going to be a balance here in terms of, obviously, for sure, making sure we protect the safety of individuals, but if we want there to be innovation in this field, these are for-profit companies, so they have a fiduciary responsibility to their shareholders to make sure that the money they would invest in, for example, the diabetes

space, is a good return on their investment. And if it isn't, then those resources will likely be shuttled elsewhere.

Fortunately, an effective and safe new drug is a win for everyone. It's a win for the patients and it's obviously a win for the company that develops the drug. If the burden to get approval is too high, investment in new drugs might diminish.

I'd note here that I think the guidance that was put out in 2008 is a good example of thinking about that balance in terms of stages of evidence to allow a drug to come on the market, but still ultimately getting data that will be important to clinicians for therapeutic decision making. Of course, if the burden is too low, then patients could be exposed to ineffective or unsafe drugs.

Let me pivot to potentially other sources of information, how we might consider them or integrate them.

What about non-cardiovascular outcome trials that are still randomized controlled trials?

Meaning there will be cardiovascular outcomes that will occur in the phase 2 and phase 3 trials. The numbers of course will be very small. The trials are small. Usually, they are done in a healthier population, so the event rate itself will be low.

But some programs have fairly extensive phase 2 and phase 3 programs, different types of patients, different background therapy, so could one aggregate such data and use it to inform on safety and efficacy?

You've seen versions of this slide before. You see the number of events, that if one were to have 90 percent power, how many events you would need to pass the thresholds there. To get on the market, a little over 120 events, and then, if you will, to stay on, then slightly over 600 events.

There was an interesting finding here. This was a meta-analysis for the DPP-4 inhibitors of the non-dedicated cardiovascular outcome trials. The print is small, so it will be hard for people to read the details. But what I've highlighted up there at the top is if you look across multiple

different drugs that were pooled together and you look at a composite outcome for MACE, you actually had about 500 events. That's a pretty decent number of events.

The odds ratio there for MACE is 0.71, 95 percent CI is 0.59 to 0.86. It's a highly significant reduction in MACE, pooling together all of the data for the DPP-4 inhibitors, again with about 500 MACE events.

When those results came out, that led to many papers saying, "Of course, it's obvious these drugs reduce MACE events," and let me show you the half-dozen or so mechanisms that explain this observation. Of course, the dedicated cardiovascular outcomes trials were done, and you can see here for MACE -- and these were obviously the larger trials ranging from 5300 to 16,000 patients with many, many more MACE events. But if you look and examine the number of MACE events, at 621, it's not that different, the number of MACE events in that meta-analysis. But here, very clearly, no reduction in MACE, all of them

essentially with a hazard ratio of 1.0?

If you look here -- and this is just sort of a back-of-the-envelope, if you will, meta-analysis -- you see the non-cardiovascular outcome trials, the meta-analysis of all of them, the hazard ratio -- actually, it was an odds ratio, but we'll let that pass amongst friends here -- of 0.71, and then for the RCTs, 0.99, and clearly significant heterogeneity for those results.

I find that I have to say somewhat troubling. Why aren't those studies at least giving us a signal that should be closer to what the truth should be because there are randomized trials.

Recently, same issue for linagliptin, a prespecified patient-level pooled analysis of all the data, so well thought through, well done for sure, small number of events because now we're talking just one drug, just 122. But again, a hazard ratio for MACE is the same, 20 to 30 percent reduction seen for all those other drugs in the class. And as we saw the recently announced

results at EASD in the CARMELINA study, a hazard ratio of 1.02.

Similarly, for dapagliflozin, a compound that we've been involved with, with AZ and the DECLARE-TIMI 58 trial, again, a very carefully done analysis of all the MACE events that were carefully collected in the trials. And you see there the effects on CV death and on MI and hospitalization for heart failure.

The results for DECLARE-TIMI 58 will be presented at the AHA. But here are the hazard ratios in dedicated outcomes trials for other SGLT2 inhibitors. And just looking at the range, so the range is the range of the point estimates for those trials, maybe the CV death is in the right range there, but the effect for MI looks like it's twice as great, and again, for hospitalization for heart failure, twice as great.

So it's not that the results are qualitatively wrong. They certainly seem to be on the right side of the line of unity, but quantitatively, I'm troubled by the fact that they

are so disparate from what the dedicated trials show. Why is that? I actually don't have an answer for it. I will point out a couple of points to think about.

One is you could say, look, there's a small number of events, and therefore, the confidence intervals are wide, and it's just a question of chance of where these point estimates fall. And that's an absolutely valid point, but that would not explain why the point estimates almost always seem to be more favorable in those studies.

It might be there's some sort of development bias, that you say, you know what? The only compounds that are going to move on to the big outcomes trial are ones that look good in these meta-analyses, so therefore there's going to be sort of a winner's curse, that the ones that don't look good won't get developed, but they would also be 1.0; and the ones that do look good, there's these meta-analyses, and they go on and ultimately have a hazard ratio of 1.0.

But this has been seen for multiple drugs in

the class. If you remember that meta-analysis for all the members of the DPP-4 inhibitor class, they all were well to the left of the line of unity. Maybe there's a publication bias to only publish if the meta-analysis shows favorable results, but this has been seen now for multiple classes of drugs.

It could be a different patient population. Certainly, these phase 2 and phase 3 trials don't have patients at the same level of cardiovascular risk that are typically in the outcomes trials, but there's really no data to suggest such an effect modification, meaning that there's greater benefit in less sick patients. In fact, Dr. Ratner very nicely showed, if anything, there appears to be greater benefit in the sicker patients, so that doesn't explain it. So I'll move on because I don't have an answer for that.

The next issue is observational data, and we've talked a little bit about how that might or might not inform us. So I'll say my bias obviously as a clinical trialist is to be a fan for randomized-controlled trial data. Obviously,

observational data are hugely important to give us insights into many important aspects of care, what patients look like, what are the types of medicines that they're on, and what are the event rate in perhaps a less carefully curated patient population.

Certainly, the data are much easier to obtain, but of course the lack of randomization raises concern that analyses, despite attempts at adjustment, are really going to be hopelessly confounded.

This is a classic example, probably familiar to many here, so I won't dwell on it, the Nurse's Health Study, very large, prospective observational study looking at patients, categorizing them on the basis of hormone replacement therapy use.

You can read the small print, after a careful multivariable analysis for the covariates listed in the footnote there, an adjusted risk ratio of 0.30; wow, a tremendous effect. Then Women's Health Initiative, a study done in half as many women, but randomized and in fact showing

evidence for harm.

Taking another page from the cardiology realm, this is for antiplatelet therapy, so bear with me for a minute, but there's a reason I'm showing this example. We did a randomized-controlled trial, TRITOM-TIMI 38, that looked at prasugrel versus clopidogrel, and it decreased the risk of ischemic events. The more potent antiplatelet drug increased the risk of bleeding, just as one might thing. There was parallel observational data from the TRANSLATE-ACS registry.

In a paper that was published in JAMA Cardiology, if you look at the unadjusted effects there, the MACE results appear to be pretty consistent there, but the bleeding is completely wrong. It's on the wrong side of the line of unity.

That's unadjusted, and then they did a couple of different analytic approaches, inverse probability of treatment weighting. You do that, and now the bleeding starts to get to be in the

right neighborhood for what we think is the truth, but suddenly the MACE benefit disappears. And then you do instrument variable, and suddenly the MACE benefit looks bigger than what was seen in the randomized trial and the bleeding signal disappears.

The most important point for that paper was what the authors concluded, that any conclusions regarding the safety and efficacy in this case of antiplatelet therapy varied depending on the analytic technique, and none were concordant with results from the randomized trial. That to me is a major prong. If the results depend on the particular adjustment, then you can't have confidence in any one particular approach for it.

So similar concerns then exist in the diabetes world. Here are data again for the DPP-4 inhibitors. Let me orient you to this chart here. On the left are the randomized-controlled trial data from 4 different trials of 4 different DPP-4 inhibitors. And you can see the effect on MACE and on hospitalization for heart failure there.

Then observational data are on the right, first looking at MACE, where we have data from U.S. Claims Service. Again, like some of the other non-dedicated cardiovascular outcome trials having favorable effects for MI, 0.85, for stroke, 0.88; and then looking for hospitalization for heart failure, again, all of them to the left of the sign of unity.

Now, a couple of caveats; the comparator here wasn't placebo or it's not going to be the nature for these observational studies. So you're comparing against people who aren't on these drugs, and they will be on different types of drugs. And some of these other glucose-lowering agents might have harmful effects, but the ones that typically do would have been a minority and I think would not explain that effect.

It's not that these observational data were necessarily all in the analysis showing statistical significance, but that's simply a numbers game. So if they didn't have 120,000 but had 240,000 patients, results that weren't

statistically significant would become so.

The concern is that the point estimates here with very large numbers of individuals appear to be very different from what we have from the gold standard. And very similar data for incretin-based drugs and heart failure here for randomized-controlled trial data and, again, a quick and dirty meta-analysis of the data that exist.

For the DPP-4 inhibitors for heart failure, your no signal for benefit may be depending on the agent, and we'll get back to that. That's a signal for harm; and maybe for the GLP-1 analogues, a trend towards a modest degree of benefit. Observational data close to a half-million patients, carefully adjusted; and there are the results.

Now, neither one is significant, but what I worry about is that if you just had these large observational data, you would walk away saying, for my patient, who I'm worried about hospitalization for heart failure, if I had to pick a drug for them, I'd probably pick a DPP-4 inhibitor based on

the data I have in front of me. Yes, the limit crosses 1.0, but as a clinician, you make do with the data you have in front of you. I'd be less likely to pick a GLP-1 analog where, in fact, based on the randomized controlled trial data, you would actually come to the exact opposite conclusion.

Similarly, for the SGLT2 inhibitor data, this question was asked. We have carefully-done, randomized-controlled trial data, showing impressive reductions in hospitalization for heart failure consistently, and it looks like some benefit for death -- it varies a bit depending on the agent -- and then carefully-done observational data, carefully-done propensity score adjusted hospitalization for heart failure, that lines up pretty well.

So that's nice. That is reassuring. What's troubling is that the effects on all-cause mortality are way greater than what was seen in the trials. So again, some elements may be concordant, but in my mind, the problem is other elements have data that are not.

So why can observational data be unreliable? So the answers are confounding, confounding, and confounding. But to take that a step further, I think in this particular case the issue, I think, is that the magnitude of any plausible treatment effect, which by and large, at least for MACE, what we're seeing is on the order of a 15-ish or so percent reduction depending on the drug and the class.

It's very likely to be outweighed by measured, but incompletely adjusted for, and unmeasured confounders related to the patients and/or the physicians who are more likely to either be cognizant of this new class of drugs, want to take the drug, want to be prescribed the drug, think to prescribe the drug to their patients, a patient able to afford it, and ultimately kind of have the wherewithal to receive it and take it for what are expensive new therapies. And I think that confounding ultimately undermines all these data.

There was a fair bit of conversation about streamlined or pragmatic cardiovascular outcomes

trials, so I'll try to touch on this in a couple different domains. First, I think it's critical obviously to always keep randomization. I think the prior examples really underscore that point. Certainly, I think we can make trials much more efficient, and we and other AROs are working to do that.

Well known to people is the PROBE design, prospective, randomized but open label and blinded endpoint assessment. Other approaches are showed in the prior slides. Sometimes we have thousands of sites. We really want to shrink that number down and instead use smaller numbers and very high-enrolling sites to really allow them to focus on the experiment at hand. And by doing so, not only do we save money, but we would hope to have then higher quality for that.

In many ways, obviously try to contact patients through mobile devices that might spare some of the burden of always bringing people physically back in for a visit. Risk-based monitoring; a lot of cost goes into sending

monitors out to go through documents at a site. Is all that absolutely necessary? There are certainly statistical ways to assess sites that appear to be outliers and then do targeted monitoring for that, r embed the trial in a healthcare delivery system.

Some of the notion, I think when people talk about pragmatic trials, at least in my mind -- and the terms are used differently by different people -- it's not to sacrifice randomization, nor is it to say it's okay if there is a ton of drop-in and drop-out.

Fundamental for the experiment is that there is a controlled difference between the two arms. The patients may look more like the patients you see in practice than you think they might in a more traditional trial, and we could talk about that point. But the issue is, to have a clean experiment, there should be a single difference between the two arms that you're then assessing when you look at the treatment effect.

Really, the pragmatic part is to say there's a lot of machinery put into place to gather the

data and to verify the data. Can we simplify that machinery to enable us to do it in a more economic way? So really, I think it comes down to data collection.

Obviously, you can try to get that from existing medical records, but it depends. Sometimes, there are requests for more detailed data, sometimes by the FDA itself wanting to have more detail, "So tell me about the size of the MI," tell me about the modified Rankin score for someone who's had a stroke. Those data are not going to be embedded typically in general medical records.

AEs will not be uniformly captured, and certainly some SAEs will be because typically, the patient will be hospitalized, but causality is unlikely to be assessed by an individual who's hospitalizing them for something maybe unaware of the trial and the drug that's being studied. And obviously, if there's safety laboratory testing, that would require dedicated visits.

There are some data that have looked at this if you try to use claims or national database data,

and I think it depends on the quality of the system. The U.S. is great for many things, but the healthcare system and having easy access to data for all is something that's still very much a work in progress.

There are questions regarding the fidelity of data, and one is really unable to do focused safety assessments. This is a nice paper that came out that looked at medical claims data versus physician-adjudicated events, and you can look at MI, and stroke, and bleeding events. And the kappas are kind of 0.5 to 0.6. They're not great, and for then bleeding, it's even worse at 0.2. I think the idea is certainly a good one, but we need to acknowledge that the fidelity is not going to be similar to what we see in the trials.

I will say much time and effort is spent on monitoring. I think much of this is spent on items that could not really meaningfully impact the internal validity of a large randomized double-blind controlled trial. But I think a lot of that is done for fear that, with inspections, if even

minor errors are discovered, it casts doubt on the integrity of the trial.

So while we want trials to be conducted to the highest standards, we have to think, are the resources being put into extensive monitoring for a trial looking for deviations that wouldn't affect the outcome. Is that worth the money being spent for that versus doing a different trial? Those are some of the hard questions for that.

Then the final, perhaps the most controversial, points I'll cover in the last 5 or 10 minutes, what about equipoise as data emerge in a drug class? Certainly, I think for safety, it's logical for initial trials for safety to have been placebo controlled.

Let's assume that we have shown safety, but not efficacy, with a drug in a class, in a given class. Is it ethical to continue to conduct placebo-controlled trials to study additional members of the same drug class? Sure, because there's no problem in doing that. Is it necessary, though? This is the question the group has been

discussing. Is it necessary to require such trials to study additional members of the same class? And that's a probably, I think.

There can be drug-specific adverse reactions. This example actually came up recently in the talk beforehand. This is work for one of the CPT inhibitors that we did in conjunction with the Oxford clinical trials group, looking at anacetrapib, and showed that it reduced the events -- not shown here, but it reduced it proportional to the amount of non-HDL cholesterol lowering. Fine.

The example brought up, which I think is a reasonable one, is for torcetrapib. So if one had said, look, let's imagine the trials were done in a different order, imagine REVEAL was done first, you'd say, "Great. We have a CTP inhibitor. It reduces LDL cholesterol or non-HDL cholesterol, and there's a risk reduction that's proportional to that, and it looks good. There really aren't any safety side effects."

Now, we have a new drug come along called

torcetrapib. Also, it raises HDL cholesterol, but it turns out that's neither here nor there, but it lowers LDL cholesterol by 25 percent. And you might say, "Well, that's great. Okay." And if they just wanted an indication for LDL cholesterol lowering, that might be fine. Obviously, if they want an indication for a risk reduction, they would need to do the dedicated trial

But you can see here that they never did it and just said I'm going to ride along with the notion that market forces will sort out what drugs they'll use. And we won't speak to this. We'll just speak to the LDL lowering that our drug achieves in the same class of another drug, that was shown to not only reduce events, but be safe, you'd be misled here, and there was a signal.

You couldn't have predicted that signal. The only way you know is by doing a big outcomes trial, but it shows what inevitably will be the case, that there will be drugs that can have off-target side effects that you won't know from the small trials, and only the large outcomes

trials will inform us as such. And I think the DPP-4 inhibitors are a good example. The MACE signal is entirely consistent between them, all essentially 1.0.

The heart failure signal is not entirely consistent between them. In SAVOR-TIMI 53, which we led, the hazard ratio for hospitalization for heart failure was significantly higher, 1.27; in EXAMINE, the weak trends towards that, 1.07; and TECOS, spot on, 1.0.

One could debate whether this is real, is it play of chance, and what about the heart failure events. We spend a lot of time investigating that, and they do appear to be bona fide heart failure events. But there are differences there, and you would never know that if you hadn't done the large trials to give you insight.

It's a topic that requires further exploration, but it may then lead to interesting insights into heart failure pathways. You would have never known that if you hadn't done these large trials.

You could then consider different thresholds of data as information for the class evolves. To use a double negative, it's not an unreasonable point. But I think one would need to be careful not to disincentivize a company to be first in class. There could be a problem that you set the bar here, and then the people who come later have a lower bar, at least initially. That could set up some odd forces there.

Certainly, you could consider different populations -- and this is a theme I'll get back to -- that would expand the overall knowledge base rather than studying it in exactly the same cookie-cutter population.

More challenging now is efficacy. Now, let's assume some drug has shown efficacy and decreased the risk of some cardiovascular outcome. Is it ethical to continue to conduct placebo-controlled trials to study additional members of the same drug class? Maybe.

So equipoise I think depends on several factors. First, what's the magnitude of the

clinical benefit? I don't mean magnitude in just a numerical sense. I mean, first of all, what type of events are being prevented, and then secondly, what is the magnitude of that relative risk reduction?

One might contrast, for example, a large mortality benefit, where then you might feel more uncomfortable for having a placebo-controlled trial whereby definition people could not as background therapy get another member of this class versus something that reduced the risk of coronary revascularization by a small percentage, where that's not a hard irreversible event.

Secondly, what's the certainty of benefit? What's seen in that first trial, is it plausible based on the data to date? Was that particular outcome the prespecified primary endpoint? Is there consistency with data from other trials if it's not the first, if it's the second one?

For example, pulling from the cardiology realm, the TAPAS, looking at thrombus aspiration and people coming in with an acute MI, the primary

efficacy endpoint was looking at an imaging parameter for the myocardium, the myocardial blush grade. There was also follow-up for cardiac mortality for a year, right and a fabulous reduction there with thrombus aspiration, essentially a 50 percent reduction? And you say, "Wow, well, that's cardiac mortality."

So it's a fatal outcome; it has to be true But that's of course not the case. So there was a 40 percent reduction in death in TAPAS, a 49 percent reduction in reinfarction, and then two subsequent trials done, TASTE and then the TOTAL trial. And ultimately, as more data accumulated, it turns out there really was no benefit.

Another point is, in terms of equipoise, what's the generalizability to the proposed study population? We heard a lot, appropriately so, about the types of patients being enrolled in these trials, and they have different proportions of patients with and without, for example, atherosclerotic cardiovascular disease.

So you might have an earlier trial focusing

on one population that certainly wouldn't preclude later trials that are primarily focusing on different populations.

Then this other issue with concomitant medications has come up. The issue there is you might say, fine, this drug has been shown to be beneficial in patients on certain background therapy, but now there might be a different class of medicines that have in the interim become standard of care, and you don't have data whether this drug in that same class will show additional benefit in patients already treated with a different drug, and that might be an opportunity to have equipoise.

Then the challenge is, is there really equipoise? Who defines if there's equipoise or not? When do we consider that equipoise has passed? Is it the announcement of the primary results in the press release? That seems premature. Maybe it's the publication of the primary results that the medical community has it. But if guideline committees haven't reviewed it,

then it's hard to say it's really standard of care yet. So you might say it's incorporation into guidelines, and that's not unreasonable.

I think from a regulatory point of view, the FDA might view it that until we've seen the data and analyzed it, and given an indication, it kind of doesn't really fully exist out there yet, and that's a very fair point as well.

Then is the drug available? And even if it is available, is there a payor willingness to reimburse? Is it accepted by practicing clinicians? So the issue is, let's say there is a drug that shows benefit. If it's being used in less than 5 percent of the population, then is there equipoise to have a placebo-controlled trial? I would say yes because clinicians aren't prescribing it and patients aren't taking it.

So the concern that you would deny someone access to it if 95 percent aren't on them seems a bit of an artificial one. Obviously, there would need to be in the consent process a thoughtful conversation not only with the patient, but with

their PCP to remind them what the standard of care is. But if the patient and PCP don't feel those data are compelling, then that really is where the equipoise lies.

Then this issue came up. What about an active control? I want to separate here the notion from background concomitant therapy versus an active control. By that I mean instead of placebo control and you have your experimental drug, the active control means people in the other arm have to get some other drug.

That's what I mean by an active control. That's different than saying one is doing a trial of an SGLT2 inhibitor; by the way, the GLP-1 analogs look good, and in future trials, yes, patients should probably be on GLP-1 analogs.

They will or they won't based on their clinicians. They'll be balanced between the arms of the start of the trial and maybe issues for drop-in later on. But that's different. That's background therapy versus what the actual control is.

Actually, Dr. de Lemos raised this point, and I'll hopefully address what he had raised. If you were to do that, what should the safety or the noninferiority boundary be? Assume the first drug in class that looked good had a hazard ratio of 0.8 for MACE versus placebo.

You could say, okay, I think you got to go against an active control. And I'm looking at this 2008 guidance. Okay, upper bound, 1.3. But you could argue that that's not quite fair because your comparator actually has a much lower risk. So you're actually being quite conservative there to insist on 1.3.

You could, if you follow the math there, say, well, actually, if our goal is just to make sure there's not harm, purely viewing it through a safety perspective, just to make sure there's not harm, then actually you'd be 1.3 times maybe 1.25, and you can set an upper boundary of 1.63 and make it easier and now enable a smaller trial.

Now, you might say you feel a little uncomfortable with that taking that point estimate,

so you might use the upper limit of the observed hazard ratios of 1.3 times 1.04, but set a boundary of 1.41. The point isn't so much the math as just the concept that if you insisted on an active control, recognize that a boundary of 1.3 is stricter than you would have set versus a placebo if you've acknowledged that active control as actually reducing events.

The other point that was brought up is somewhat related, but different. If you now say, "Look, there's an active drug and we want to show we're as good as the other drug out there, want to show that we're noninferior to that drug," similar in the common vernacular, then the typical rule of maintaining at least 50 percent of the benefit and applying the ratios there, the upper limit would need to be less than about 1.12.

That does, as Dr. Ratner had noted, necessitate a much larger trial, but that's a separate question. That's now looking for efficacy in this case. And it's hard because unlike, for example, the trials of NOACs versus warfarin, where

warfarin as the control arm had a two-thirds' reduction in events, here we're talking about risk reductions of the 15 to 20 percent margin. And therefore to be sure you maintain and that benefit, it then makes that upper limit quite tight.

I think the other issue is then to think we need to be more creative and look at different populations. And we heard very nicely about the different comorbidities that patients with diabetes have. We viewed it initially narrowly because of the rosiglitazone data in terms of MI, so thinking about those with atherosclerotic cardiovascular disease, which is so-called secondary prevention versus primary prevention for it.

But maybe you could have trials where you actually would have some patients with prediabetes and get insight into them. Maybe you enrich for patients, as we heard, with heart failure or renal disease, which are common and are associated with poor prognosis in patients with diabetes, or maybe with and without diabetes.

In some ways, I think this is analogous to

what happened with statins, which started with the obvious sweet spot of prior MI and high LDL cholesterol, and hit a home run with that. And then, as other statins wanted to test that, you couldn't really test that exact same population, so you said, okay, we're going to target lower levels of LDL cholesterol. Then we're going to shift from secondary to primary prevention, then look at other types of concomitant disease.

Finally, the last few minutes, what are the optimal safety and efficacy endpoints? What's the right safety endpoint? Well, as discussed, it really came from the concern for MI with rosiglitazone, and from that, a reasonable composite outcome of CV death, MI, or stroke. But is that necessarily the correct CV safety concern for all diabetes drugs?

Again, it's a very atherosclerosis-centered view for it, but we know for the TZDs and maybe for some DPP-4 inhibitors, that there's an increased risk for heart failures. Maybe that should be something that's monitored as well.

What's the right efficacy endpoint? The composite efficacy outcome, again, of that same triple, of CV death, MI, or stroke, is of course a natural extension of the safety analysis. If you're going to gather 600-some-odd events to show safety, why not gather 1200 events and then be powered for efficacy? And they're hard, irreversible outcomes.

There's no quibbling with them, but again, it's largely atherosclerosis-centered outcomes, and we've seen data for SGLT2 inhibitors, that they appear to be particularly good for decreasing the risk of heart failure and renal disease. And this is the public domain for DECLARE-TIMI 58.

That initially was the traditional, if you will, MACE for safety and MACE for efficacy, but with hospitalization for heart failure as a secondary outcome. Based on the data from the phase 2 and phase 3 work, that showed a very powerful effect for hospitalization for heart failure.

Then data from EMPA-REG outcomes came along,

so while we remained blinded and before the first data monitoring committee meeting, we then elevated CV death and hospitalization for heart failure, so that's now a dual primary endpoint. This I think fits with the notion of having adaptable trials, where again, it's not the watchmaker winding up the trial and letting it go. You have to pay attention to what's going on in the field.

In fact, for the press release that Astra Zeneca released, it indicated that dapagliflozin not only was safe, but then decreased CV death and hospitalization for heart failure. And while there were fewer MACE events, that wasn't statistically significant. And that actually fits with the magnitude of effects we've seen for other drugs, and more details will be presented by Dr. Stephen Wiviott at AHA.

What's the right population? Is it all patients with diabetes for trials going forward? Again, what about prediabetes? What about normal glycemia? Especially if safety has been shown for other members of the class, that might be an area

where you can then allow a little more latitude.

ASCVD is great, necessary to get the right MACE events, but what about patients with heart failure and chronic kidney disease and perhaps a program that combines them?

I'll finish this with a few notions. I think -- and this is from the cardiologist's perspective -- prior to the FDA guidance, the treatment of blood sugar is just, okay, your A1c is too high and we want to get that down. And that's from a knuckle-dragging cardiologist's view and not the elevated cerebral endocrinologist view for it.

Then after the trials -- and we would like it read into the Federal Register that I think it was the cardiology guidelines that first adopted the data from the diabetes outcome trials -- knowing there that empagliflozin should be considered in patients with type 2 diabetes to prevent or delay the onset of heart failure.

I mean, that's fabulous, that now we have those data to guide us. As beautifully outlined in the ADA guidelines, now that's part of the thought

process for which drugs we should pick, again, thinking about agents that reduce major adverse cardiovascular events. And this just shows -- not to read it -- but now that's part of the table of events.

In conclusion, about a decade out from the issuance of that guidance, cardiovascular outcomes trials are certainly large affairs. They require a large investment of time and money. And there are people pushing back against that. In theory, that might dissuade some companies to target resources for diabetes, although that hasn't appeared to be the case as of yet.

But I think we should be, as a community, very proud of the impact of that guidance. We now have trials that have a wealth of data that I think have tremendously advanced the care for patients with diabetes. I think, from my perspective, it shifted from battling over small differences in A1c to now looking at reduction in cardiovascular risk.

I think there's ultimately no substitute for a dedicated cardiovascular outcomes trial to

definitively answer these questions. Certainly, the trials can be similar with maybe less machinery needed to be brought to bear. The fundamental principles of randomization and careful ascertainment for the outcomes of interest, I think is paramount.

Then as robust validated treatment benefits consistently emerge for a class, then equipoise for ongoing trials will need to be considered. For thinking about designing a new trial, then one needs to think outside the box and think about the patients that need to be studied and perhaps be creative to explore different patient populations; think about the comparator; and then think about the endpoints that make the most sense.

Thank you very much for your attention.

DR. WILSON: Great. Thanks very much, Dr. Sabatine. We're going to have some time for questions now, and let me just say what we would like to do for the rest of the afternoon since we're at a juncture, so to speak. We'll take 15 minutes of questions for Dr. Sabatine. Then

Dr. Green will speak, and then she'll have a dedicated 15 minutes. And then if Dr. Sabatine's available after that --

DR. SABATINE: I will be on my way back to the CCU at the Brigham.

DR. WILSON: So we may cheat on the 15 minutes a little plus here, then. You have to leave immediately after this?

DR. SABATINE: I do.

Clarifying Questions for Dr. Sabatine

DR. WILSON: So let's see how we go. Any questions, Dr. Wang?

DR. WANG: Thanks for that really comprehensive and helpful review. Just three quick practical questions. One, just as a guesstimate, for these types of populations that are being tested in the CVOTs, if heart failure was added to the MACE endpoint, heart failure, hospitalization, how would that impact the sample size requirements? Would it be a dramatic change or would it be a modest one?

DR. SABATINE: No. It's a good point, and

there are a robust number of heart failure events. Heart failure is certainly an important outcome that can happen in patients with diabetes, both with and without having reduced ejection fractions, as you well know.

I think the one issue, I guess, in my mind is that the events will certainly accrue in these populations. We might have to think carefully about having a broad composite because that could be challenging in terms of the more you extend beyond the traditional athero, if it's driven just by one element, it's a little bit hard.

If you had a drug that decreased athero and there was a bit of benefit in MI, and a bit a benefit of benefit in stroke, and a bit of benefit in CV death, and it all lined up, you'd say, great. If you put in heart failure, and there was a 50 percent reduction in heart failure but nothing for MI and stroke, then you're sort of left in this quandary of saying, well, the primary endpoint was CV death, MI, stroke, heart failure. So that then gets to not fully reveal how the trial did.

But I think, as it emerges, that heart failure and renal disease are very important issues for patients with diabetes. I think those outcomes need to be front and center. And they're common enough that I don't think it's going to increase the sample size. I think we need to pay more attention to it.

As is being done for the SGLT2 inhibitors, you're now saying here's a class that looks wonderful for that composite of CV death, hospitalization for heart failure, so that's what we really need to focus on. And of course we'll pay attention to MACE and of course make sure it's safe.

But where we think the efficacy signal is most powerful is going to be in a CV death hospitalization for heart failure. But there are plenty of those events and, and obviously you could enrich, if you had your enrichment criteria, not on did you have 2 MIs, but did you ever have a hospitalization for heart failure? And then you would increase the event rate.

DR. WANG: Yes. And just to clarify for the record, I was getting to the point that maybe adding heart failure would actually lessen the burden of sample size because you're saying you care as much about heart failure as an atherosclerotic CV event, so it's part of your composite.

DR. SABATINE: Yes. I think you could decide as the field emerges, if you're first in class for MACE, you may want the same level of knowledge. But as MACE becomes more and more certain, how many more MACE events do you need to gather for DPP-4 inhibitors? Probably not a lot. But the heart failure signal could be different for them, and that you would want to get more events and have more granularity, I think.

DR. WANG: Just in the interests of time, I'll just jump to my last question, which is, there's been a lot of discussion about the different patient populations of primary versus secondary, but there's also the issue of the time course.

So going to a secondary prevention population, especially with some of these trials, with recent ACS, it strikes me that the things that determine cardiovascular events in that population over the first year and a half, which was a short period of follow-up, versus both safety and efficacy over 5 to 6 years, may differ.

DR. SABATINE: I think that's a fabulous point, and I couldn't agree with you more. Certainly, the ACS population has a very high event rate, but it is front loaded, and a lot of that risk is going to be based on the index event and other sort of coronary issues, which it's not clear that a drug to treat diabetes, a glucose-lowering agent is really going to necessarily impact that per se; whereas over the longer time course, you think of it as more of a disease modifier.

So I think it is different than an antithrombotic, where you start the antiplatelet, you start the anticoagulant, and the event curves diverge immediately. These are more of a disease-modifying drug, to kind of take a page out of

rheumatology, where the effect eventually kicks in, at least for the athero part. So I think in some ways, ACS trials are not well suited for this population.

The whole issue for primary versus secondary prevention is tough. We grapple with that, and even the primary prevention, we're not doing angiograms on all of them, so we're not saying their coronaries are whistle clean. But I think the point that was raised is absolutely true. There does appear to be a distinctly different response for reasons that really remain to be sorted out.

DR. WILSON: Dr. Budnitz?

CAPT BUDNITZ: Dan Budnitz, CDC. Thank you for a really important presentation. You kind of led us down a path maybe to consider streamlining cardiovascular outcome studies. I'm wondering if you could give any additional specifics about ways that you might streamline, based on your experience.

I think one of the other points you made is

the wealth of data that we have learned, both information we've learned from the CVOTs, and as we streamline things like collection of certain outcomes, maybe we wouldn't have learned quite so much. So any prioritizations of --

DR. SABATINE: I think that's a great question. I would say I think you hit the nail on the head. The value is in getting all the outcomes. So in having the outcomes and having confidence for what they are -- for example, for the heart failure outcomes, for SAVOR having adjudicated them, we said, okay, we think those are real outcomes. Had we not done that, that would have been problematic. And you'd say, well, I'm not really sure whether it was heart failure. Maybe it was some ankle swelling. I don't really know what was going on.

So I think of the money to put in, gathering data on important cardiovascular outcomes, renal outcomes, I think that's where the money is well spent. That's the information we want as clinicians.

The triple monitoring of every little data point I think isn't as useful. I'd rather try to use more remote monitoring and think that because it's a blinded trial, even if there are mistakes -- let's say a site fails to report an MI. Okay. There are 500 MIs reported in the trial. They're blinded to it. So if the CEC is blinded, it doesn't make a difference.

So obviously you want to have complete data, but that's not going to affect the integrity of the trial. So I'd much rather spend money on saying let's dive in deep for the heart failure outcomes, the renal outcomes, rather than triple-checking some eCRF form.

DR. WILSON: Dr. Yanovski?

DR. YANOVSKI: Yes. I think some of this has been brought up. But because I think heart failure's such an important outcome, it's my understanding that it is more difficult to adjudicate than some of the traditional cardiovascular MACE outcomes.

Can you talk a little bit about that? And

also, if we move to things like more pragmatic trials, things using registries or EMRs, what we can do about capturing heart failure, given that it's not quite as easy to discern.

DR. SABATINE: That's a great question. For clinicians in the group, patients who are admitted with questioned heart failure/pneumonia/COPD flare, that still exists every day in the hospital, and it can be hard to tease it out.

There are definitions, and Steve Wiviott, who actually chairs our CEC, is working closely with folks at the FDA, Karen Hicks and others, have definitions that we think will favor specificity. So ultimately, we're protected by specificity. We might miss some events, but we want to be sure the events that we have are really heart failure events. And there are ways to do that in terms of are they admitted, what data you have, what therapies we're instituted for, that we try to make sure the specificity is high.

You raise a good point. As you go to other data sources which have greater convenience,

there's appeal for that. But you could imagine, for example, from a safety point of view, the danger for that because if you start introducing noise, then noise will bias you, by and large, to the null.

So in that case, then -- for example, from a cardiology perspective, a patient's ability to distinguish whether they were admitted for a myocardial infarction versus for a rule-out myocardial-infarction is not great. If you just relied on what the patient said and there was no verification of that, you might introduce a whole bunch of atypical chest pains.

If you're looking for efficacy, that'll hurt you in the trial. If you're looking for safety, you're going to be biased towards 1.0. So I think that's a challenge to try to extract that.

As the healthcare records get better, there's more and more opportunity to say, okay, I want to extract from the record the relevant variables. I want to extract what the therapeutics were that were done for that. So you can start to

approximate. Again, to date, the kappas, the correlation coefficients have been kind of moderate, so I think it's tough. It's tough for that.

DR. WILSON: Dr. Grunberger?

DR. GRUNBERGER: Yes. Thank you very much for that excellent talk. Even though I might be a cerebral endocrinologist, I'm a very poor cardiologist. So maybe going back to the whole heart failure issue, when we start talking about endpoints, after SAVOR-TIMI 53, people talked about hospitalization for heart failure, then morph into heart failure.

So are we treating heart failure, are we preventing heart failure, or are we preventing hospitalization? And is the hospitalization itself a really good hard endpoint to a heart failure event?

DR. SABATINE: It's a very cerebral question, and it's a good one. You're right, and it dovetails a little bit to the prior question. Certainly, hospitalization for heart failure, it's

a little bit like in the cardiology world, admitted with unstable angina without biomarkers to show an MI, but you get cath. That just suggests there's enough going on there that a clinician has decided to bring it to the cath lab. The same thing for the hospitalization for heart failure; it adds a level of rigor automatically to the endpoint that you wouldn't have.

But it does bring up I think a very good point. You could imagine a variety of drugs could influence hospitalization for heart failure through a variety of ways, for preventing people from becoming symptomatic and coming in versus being a disease modifier. And all that is I think ripe for study, and both are valuable.

Certainly, I think what we have discovered or felt in our own practice is that certainly there are patients who die of pump failure for sure, but there are those patients who keep getting hospitalized for heart failure and die of the consequences of that hospitalization. And by that I mean they're brought in. They're short of

breath. Ultimately, they're intubated. They get a ventilator-associated pneumonia, and they die of an infectious disease there, technically. But that wouldn't have happened if they didn't have the heart failure hospitalization, or they had a pulmonary artery catheter put in, and they wind up getting bacteremic and septic, and then dying.

So I think those hospitalizations carry a big burden on the patient, so I think there's certainly value in preventing them. But I think scientifically your question's great, and you could imagine a variety of drugs that would affect the need for hospitalization as well as affecting the underlying substrate as well.

DR. WILSON: Ms. McCollister?

MS. McCOLLISTER-SLIPP: Again, I'm going to try to talk to you directly while using the microphone. I'm not a statistician. Nobody wants me to do statistics. But I did find your presentation really fascinating on a number of points, particularly the section where you compared the RCTs with the observational studies.

I think one of the questions that we need to consider is the cost-benefit of doing these massive, large RCTs, which produce really interesting, potentially very helpful data versus using other sources of data and other methods with observational data sources.

One question I have; given the general crisis of reproducibility with clinical trials, and since that's sort of like a backdrop issue, the fact that none of these trials -- all of which are well designed, none of them are going to be reproduced. We'll never going to be able to verify, in the exact kind of setting, whether or not these are reproducible in and of themselves.

Is it fair to do that -- I mean, I think it's appropriate to do that kind of analysis, but is it fair to conclude that because the observational analysis did not conform with the findings of the RCT, that is a less ideal form of data?

DR. SABATINE: No. That's a fair point. You could say, you know what? The observational

data has 10 times as many individuals. I'm going to believe those data.

I think, in my mind, the issue is a couple. One, I showed you in one example, you could see from observational data, as you change different analytic techniques, you got very different answers. That's typically not true for a randomized-controlled trial. You could do a variety of sensitivity analyses, and in general for robust trials, the answer remains almost always the same. So that's one worrisome.

Two, your point, though, would be well taken if you had just one RCT and one large observational trial where you might say, gee, RCTs can be wrong. So I think you have to view it in the totality, that if you said we've now had multiple trials of DPP-4 inhibitors and all of them have had a hazard ratio of 1.0 for MACE, now when you see observational data that suggests something that's showing benefit, you say, gee, that's running counter now to 4 trials. Now, you haven't had 4 trials with the same drug, but all the members in

the class have all shown a hazard ratio of 1.

But I think your point is fair that any one data set, you have to worry about. But you think back; that's why the FDA has required the RCTs, because the observational data were not viewed to be adequate.

So I think they're useful for certain bits of data, but at least in these examples, I think the totality of data is quite clear, that the observational data are off target.

MS. McCOLLISTER-SLIPP: Again, this is not what I do during the day, and I'm certainly no expert on it. But I do know that this is a broad issue. And I just pulled up a JAMA study saying that even randomized controlled trials, when they're reanalyzed by different statisticians using different methods, come up with different results, and that was actually published.

DR. SABATINE: Yes.

MS. McCOLLISTER-SLIPP: So I'm just wondering if it's an appropriate way to evaluate the efficacy of observational data points.

DR. SABATINE: I think it is. I think those analyses are tricky because oftentimes the other group doing isn't privy to the correct data dictionaries. At least for all these trials, we have the benefit, typically, and it's a good point for the analysis. The sponsor analyzes it. We as an ARO-TIMI independently analyze. And then our colleagues at the FDA get the raw database, and we all typically come up with exactly the same answer.

But I think, for any of these, you certainly could argue that there are differences in patient populations. So having multiple RCTs that point in one direction and observational data that point in another direction, that gets me concerned.

If you just had one RCT, then you'd say it's only one bit of data, and you could say I'm not quite clear which one will be right. But I think in these examples, there's a preponderance of data that show that the two are quite different.

DR. WILSON: Martha Nason, next?

DR. NASON: Thanks. I was actually going to make a similar point to what you made about the

noninferiority trials versus superiority trials about -- I keep flipping back and forth in my head, trying to keep those separate because the same thing can affect those very differently; for instance, adding heart failure or adding something that adds noise; or in a superiority trial, as you said, it might handicap you, but in a noninferiority trial can make the two groups look more similar.

There were several things throughout that -- not just your presentation, but broader that have been catching that aspect in my head, including you mentioned unblinding or having unblinded trials as one way to streamline, but that again would worry me that you might be adding variability. People would know, oh, I didn't get assigned to the new therapeutics, so they or their doctor might be more likely to start a different drug, for instance, which could then bring those two arms together.

Of course, unblinding would be a problem with either direction. And even if it's not

unblended -- maybe I'm rambling a little here, but even if it's not unblinded, seeing the effect on the A1c might sort of unblind in a blinded trial such that those were brought together.

I guess I was going to sort of push back a little bit on that idea of unblinding trials as a way to streamline it and still get the information you get from a randomized-controlled trial, but also to point out that through these discussions, even as we're talking about adding heart failure or whatever these issues are, to me we need to separate whether we're talking about the potential of showing a benefit or whether we're talking about looking at it as a noninferiority signal because those may be completely opposite effects that you have on your ability to get the right answer if you're adding something that's not quite the same, something like heart failure or something that adds some noise.

DR. SABATINE: No, no. I agree. I think, obviously, for the safety and the efficacy, you articulated exactly correctly, that if you're

looking for efficacy, you want to have specificity for what you think your drug is going to affect and not include endpoints that you think won't be modified by your drug. And if you introduce noise, that will hurt you for efficacy, but it could potentially help you for safety.

To be clear, I'm not advocating that we do unblinded trials. I just list it in ways that people have approached this issue. That is one approach. And you could say if my outcome is death or stroke confirmed by CT, well, if you're unblinded or not, it's going to be hard to shift that. But for other outcomes, for hospitalization, for heart failure, you could be swayed by that. You certainly could be swayed by that.

The issue's always existed for cardiology. If you think someone's on a drug that could prevent their progression of atherosclerosis, you tell them not to go in and be checked up for something, and if you know they're not on it, you worry about it, and they go in and they get a work-up for that.

So I'm totally in favor of blinded trials,

but for the sake of completeness, the PROBE design, people have advocated for that. But you'd have to be very careful what the outcomes were. They'd have to be very hard outcomes that would not be subjective in terms of a clinician deciding to do something.

DR. NASON: And I think it's more than just hard outcomes. I think it's also that you wouldn't want, especially in these longer-term trials, people to be added another medication if it's standard of care that allows the freedom for the physician to add other medications or to treat them differently, as well as the outcome, you could have an effect just by that.

That's related to some of the earlier comments I think about what the standard of care is and what the control group is.

DR. SABATINE: It's a good point. There's always the risk for drop-in for any of the diabetes trials. Usually unlike the tightly controlled phase 2 and phase 3 trials, for the outcomes trials, they can get pretty much whatever other

glucose-lowering agent they want as long as it's not in the same class as what you're studying.

So there can be people who stop study drug and drop in. As is often the case, usually those numbers are very small, but it is a theoretical concern. But practically, they tend to be small.

DR. WILSON: Every time I take a pause, Marc, I get more questions. Can we go another five? We're going to miss you.

DR. SABATINE: I'll miss you, too, but, yes, we can go another five minutes.

DR. WILSON: Can we go another 5 or 10 more minutes?

DR. SABATINE: Sure.

DR. WILSON: Because you're going to have to leave.

So next on the list, Marc, Dr. Ellenberg. Susan?

DR. ELLENBERG: So assuming that some consensus can be brought to bear on the endpoint issue, what's the right endpoint, what do you see as the minimum duration of trials of these new

agents? How long do we need to be able to study and follow up patients?

DR. SABATINE: Yes, that's a very good question. I think there's no magic answer for that, but I think, in general, as a scientist, I would say usually time is on your side. There are a couple of things, a couple of vectors pointing in different ways.

To the point that Tommy had raised earlier, in an ACS trial, you're going to have a high event rate early on. They're going to be typically, largely unmodifiable events for the sorts of drugs we're talking about here, so that's going to hurt you for that. If you think the drug is changing the underlying biology, then for lipid lowering, time is your friend and it needs time for that to kick in, so having a multiyear trial is good.

The vectors the other way are that the sponsors obviously want to get an answer in a reasonable time frame, so having a trial go on for 8, 9, 10 years becomes somewhat problematic. And there is some trial fatigue, that it's just hard to

get patients to stay on the drug.

The other point I should have mentioned is, for safety, obviously you do want longer exposure. So having more people exposed for a shorter period of time, same patient-years, isn't necessarily the same as fewer people for a longer duration.

So I would say that I think time is your friend in that to get a more robust answer, so I think on the order of 4 to 5 years; not that 3 years is wrong, but I think that time frame is good. Shorter than that, I'm not sure you're going to do yourself a favor to see the benefit. Longer than that, there's just hard experiments to conduct.

DR. WILSON: We're going to have three more questions, and then we're going to release you. And the three are Dr. Rosenberg, Dr. Newman, and Dr. Wasserman.

So Dr. Rosenberg, short questions, short answers.

DR. ROSENBERG: It'll be short, I think, for the first question because it's already been

answered regarding the PROBE design and how much it is dependent on the type of outcome you choose. The PROBE design doesn't really protect you if you have a refill [indiscernible] or another potential bias. You have to be very careful about that.

I think that's a major defense between also the streamlined trial and the pragmatic trial. All pragmatic trials are streamlined, but not all streamlined trials are pragmatic. There's a difference in approach.

A pragmatic trial, you want to reproduce what would happen, so you account in your design for drop-ins and costs, but you don't really control for them. So I don't know if we can use that in the context of a regulatory trial.

The second question is why don't we do more long-term follow-up of registry following clinical trials? Which will answer many of these questions related to what Dr. Ellenberg said about we need long-term follow-up. And Dr. Ratner just showed us that we have shown that when we follow patients included in trials, we show whether or not there's

really a long-term effect of this relatively short intervention.

DR. SABATINE: Yes, great points, and I agree with the PROBE issue, which I think we covered, so completely aligned for that. I think the long-term follow-up is great, and I think that comes in potentially two flavors. The simpler of the two is to say, fine; you finished your experiment and people come off study drug. But then is there some legacy or memory effect, and we've seen that for glycemic control and we've seen it for lipid lowering, so there is benefit.

At least in the lipid-lowering trials, that appears to, I think to my mind, convincingly exists for the first year, a little less clear beyond that, although different data sets have come up with different conclusions for that point. But I think that would be fabulous to have additional data.

There are other takes on that, where you could say after you finish the trial, then everyone's on an open-label extension. You won't

have a control group, but you'll say did some bad things suddenly crop up in a rate that seems to be higher than you would expect, and you just have more patient-years on the drug, and that's a plus.

Then there are even more provocative notions that would require careful conversation with regulators. Could you have for a composite outcome a certain endpoint that you look at, at a certain time, but you keep the trial going until you get more definitive data from mortality outcomes. And you have a conversation with regulators saying, you know what? We need the MACE data in a certain time frame because that data needs to come out. And that's important to a sponsor, and I think that's quite realistic.

As a scientific community, we would love to have long-term data on CV death. But to power the trial for that would be another 3 or 4 years. If we wait for that, that's a little difficult. Maybe we could do a two-step process where you have a cut point for that, but you keep the trial going in a blinded way, and patients could be reconsented

depending on the results. These are complicated conversations.

But to your high-level question, should we follow up patients more? Yes, and as we have more centralized and more accessible electronic health records, I think that's great to do, and we're trying to build that into our trials. So I agree with you a hundred percent.

DR. WILSON: Dr. Newman?

DR. NEWMAN: Thank you. Thank you for your presentation. I have a question about heart failure with the DPP-4 inhibitors. We know that saxagliptin had increased risk of heart failure, and there was a numerical increase in hospitalizations for heart failure with alogliptin, and other DPP-4 inhibitors did not show this.

If we had a new DPP-4 inhibitor, would you suggest a trial, a randomized-controlled outcome trial to evaluate for heart failure?

DR. SABATINE: Yes. I think that's an example where you do want that information. It's not clear to me why that signal would be apparent

with saxa, and less so for alo, and not at all for sita. Others might have other insights on that, but it's not obvious to me. I don't think it's necessarily obvious to the general community.

So I think if you had another drug in that class, you'd say, look, heart failure is something that we have a question about, there's sort of scientific uncertainty about that, and that's got to be established. Because I would think as a clinician, maybe it's noise, but maybe there really are differences.

So if I'm going to prescribe a DPP-4 inhibitor, I'd like to have a sense for how that particular compound does. So I think that's a perfect example of saying, for that particular class, the thinking has shifted. Now our focus has to be slightly different. And it's not just the MI with rosiglitazone that should be carried forward forever for every class. Now there's a question, at least for heart failure.

DR. WILSON: Dr. Wasserman, you get the last question.

DR. WASSERMAN: Thank you.

Marc, I was looking at how we've gotten here; it's the 10-year journey. You go back to the UKPDS, which was one of the studies that the FDA highlighted as what led us here. I just did a quick search on Google. It was about 40 events for diabetes-related death. I also just pulled the muraglitazar, which was about 34 deaths.

What I'm struggling with is when we have small numbers -- now we have a decade's worth of cardiovascular outcomes trials which have not replicated what we've seen in those small studies -- how do we make better decisions? How do we help the FDA? How do we help sponsors? How do we help the academic community to make better decisions, other than just running RCTs all the time? I mean, is there something that we can learn from this?

DR. SABATINE: There's nothing wrong with running RCTs all the time. But I think to your point, we do need to adapt for it. You brought up a couple of very good points. One is a lot of the

prior data is data now that we would say that's really quite scant data and we're not sure exactly what to make of that. But I think that's a sign that as a group, we've continued to raise the bar for it.

I think it is a matter really of adapting to the field and to individual classes to say, what's important to study for that? So I think the 2008 guidance was fabulous. There was a time when -- and I missed the morning presentation -- as the trials came out and they were all noninferior, people were saying, what are we doing with our time and our effort here? But now that we've had a series of superiority trials, I think that's yielded huge dividends. And I feel that the practice of medicine in that area has progressed tremendously.

So just like the FDA's doing now, I think you say, okay, what's the right guidance? And I think to be creative and say, as the field is evolving, let's think creatively. What are the outcomes we care about? How much certainty do we

need for this field?

So the answer may be that it can't be a decade in between. I raise this as a formal process, and you're always evaluating constantly, but there may need to be more adaptation, as the data come, to say we need to pivot so that the trials give us useful information.

But I think without those data, some of the effects on heart failure and on renal outcomes wouldn't be highlighted, and now we know those are really great areas for patients with diabetes to target.

DR. WILSON: Okay, you're off the hook.

DR. SABATINE: Thank you, sir.

DR. WILSON: Dr. Sabatine, thanks very much. Take care.

DR. CHONG: I'd like to welcome Jennifer Green, who's our last presenter of the day. Jennifer Green is an associate professor at Duke University, as well as being on the faculty at Duke Clinical Research Institute. Her career has been spent studying patients with diabetes and looking

at both glycemic control and also cardiovascular outcomes. So Dr. Green, thank you for your participation.

Speaker Presentation - Jennifer Green

DR. GREEN: Thank you very much for the opportunity to speak today. As you can see on the slide, I am affiliated with Duke University as well as the Durham VA Medical Center. However, I'm not here on behalf of those institutions today, and any of the opinions, perspectives, et cetera that I present are solely my own. These are my disclosures for the past three years.

I was asked to provide an endocrinologist's perspective on the impact and importance of the guidance in diabetes care. And I wasn't, for this presentation, given any particular assignments or questions to answer, so I've sort of roughly grouped them into the good, the bad, the ugly, and some thoughts about the future.

First, of course, I'll start with the good. And I'm sure I'll echo some of the themes that were covered during the earlier presentations. And I

think it's clear that the completed cardiovascular outcomes trials have provided reassurance for the fact that newer drugs currently available to treat type 2 diabetes don't increase the risk of major adverse cardiovascular events or MACE.

We've clearly established that there are in fact cardiovascular benefits of several drugs in patients at the highest cardiovascular risk and, interestingly, as a compliment to other cardiovascular risk reduction strategies that we commonly employ in this group of patients.

It's highlighted the fact that heart failure is an under-recognized, underappreciated, but still very, very important complication of type 2 diabetes. And although these drugs have not been tested against each other, for the most part, they've all been assessed with respect to safety and efficacy versus placebo, I think the results so far certainly do suggest that there's heterogeneity effects of these drugs on cardiovascular and other events, both between and within the drug classes.

Interestingly, as has already been touched

upon, there appear to be some unexpected other potential benefits of these treatments with respect to effects on heart failure and preservation of renal function in, again, this group of patients who largely either already have impaired renal function or are at risk for significant decline in renal function over time.

These trials have definitely provided additional information of importance regarding other safety outcomes, and these trials either addressed pre-existing or identified new issues of clinical interest to the care of these patients and shed a lot of light on the effects of the interventions or absence of effect on rates of thyroid malignancies, pancreatic safety, amputations, fractures, et cetera, and have also really been particularly helpful in identifying the frequency with which these particularly rare complications might occur in this patient population.

Some if not all of this safety data potentially might not have been available

elsewhere. So these are of course some of the benefits of having performed these trials that satisfy the guidance requirements.

Maybe as an off-shoot, but certainly related are other benefits of generally enhanced safety expectation regarding drugs to treat type 2 diabetes. And as you can see listed on the slide, there are three agents whose development was discontinued due to the discovery of potentially rare but still potentially also very serious side effects, including allergic reactions, for example, worsening of kidney function and GI bleeds for aleglitazar, or drug-induced liver injury. And you can see listed here, essentially, the number of patient-years of exposure that were required to determine that the drugs increase the risk of those side effects.

I think perhaps most importantly, and again as has been alluded to by the other speakers, the evidence generated by these trials has contributed to a truly remarkable evolution in refinement of diabetes care guidelines for the highest-risk

patients.

If we take a step back in time, again, to the diabetes care guidelines that were in existence around the time or just shortly prior to the issuance of the 2008 guidance, you can see that at the time, there was a very heavy emphasis upon intensive glycemic control strategies to reduce the risk of complications.

For example, the 2006 ACE targets for glycemic control were for everybody, an A1c target of less than or equal to 6.5 percent. ADA was a little bit more relaxed, so it was generally less than 7 percent. But clearly commentary that more stringent glycemic goals and in fact treatment to a normal A1c could be strongly considered to reduce the risk of complications. But of course, there were caveats that in order to get this range that early use of insulin was going to be necessary, and there would be a trade-off regarding an increased risk of hypoglycemia with such a strategy.

I think it's also important to remember that the justification for these recommendations largely

came either from extrapolation of findings from other trials, which in fact did not test this degree of intensity of glycemic control, or based on observational data suggesting that people with diabetes with lower A1cs were generally healthier than people with higher A1cs.

If we jump to mid-way between 2008 and now, you can see in 2012, again, there was an evolution in what we recommended as far as the approach to diabetes care. And in fact, this major change at this midway point came after we in fact tested the effects of very intense glycemic control as a cardiovascular risk reduction strategy in high-risk patients, and found that, in fact, did not reduce the risk of those complications, and in fact increased risk of harm in at least one trial.

So in 2012, the ADA relaxed or modified their guidelines to permit more relaxed and perhaps individualized glycemic targets where less stringent A1c goals could be implemented for people with established complications, et cetera.

However, interestingly, if you could look at

how we get to those targets. If you look at your recommendations for use of medications for type 2 diabetes, you can see that after metformin, it's in fact largely just a list of the other classes that are available. And rather than focusing on demonstrated individual benefits of those classes, instead the choices of drug after metformin were largely guided by the type and risk of the side effects of the class that you were trying to avoid.

Again, changing to the current day, the current guideline -- and this is just one example from the recent ADA EASD set of guidelines that were published earlier this month -- incorporate evidence from the CVOTs in a robust and important fashion. And this is just a snapshot from this very extensive guidance that recommends that for people with type 2 diabetes and established cardiovascular disease, that their therapy include use of an agent that has a demonstrated cardiovascular benefit when used in patients of that type.

So it's a remarkable evolution from

recommendations that weren't particularly evidence based but seemed like they might be helpful, to midway, sort of a listing of what's available, to currently being able to make recommendations of drugs that will in fact benefit our patients when incorporated into their care.

Now, of course, it can't all be good, and the guidance requirements undoubtedly increase the cost of drug development. And it's clear that the large cardiovascular outcome trials that have traditionally been conducted can cost upwards of $500 million, and those costs are almost certainly conveyed to patients in increased total care expenditures.

So we're all in a sense funding this, and certainly, it may serve as a disincentive to future diabetes drug development. But that's a little bit hard for me personally to assess. I'm going to counter that just a bit.

Depending on which CVOTs you consider as falling within this group of trials, somewhere between 10 and 15 CVOTs of agents, and 3 new drug

classes, and one of insulin therapy has been completed, and there are more ongoing. So there's certainly been interest in conducting these trials to satisfy the guidance.

Since 2008, there have been 14 new agents for type 2 diabetes approved, and that doesn't count insulins or drugs that are essentially combinations of multiple drugs. And of course, some of these were probably already in the pipeline before the guidance requirements kicked in, but that's a non-trivial number of new therapies available to our patients. And you can't deny that the market for these drugs is steadily increasing as the prevalence of diabetes increases in the U.S. and elsewhere.

I think as a potential incentive for diabetes drug development would be some of the findings that some of these drugs may in fact have multiple avenues of physiologic benefit, and some of these drugs may end up having multiple indications for treatment other than their original intent. So for example, there could be benefits to

be explored in heart failure and chronic kidney disease.

Again, it's an imperfect way to get a handle on what's happening within a certain area of research, but if you do a clinical trial search as I did at clinicaltrials.gov search earlier this month, and you look for active trials, interventional studies in diabetes that include an investigational drug, there are over 800 registered in clinicaltrials.gov.

So I wouldn't say that there's an absence of interest or any sort of a suppressive effect of the guidance that I can see within this research space. And in fact, I think it may stimulate additional investigation.

When I look at data that is available to me in the public domain -- for example this is something that was available online from just over a year ago; and this resource lists the number of drugs and research projects that are ongoing in various clinical spaces at that time -- you can see that, of course, the diabetes space pales in

comparison to the number of drugs and clinical trials that are being conducted for cancers. But the numbers aren't terribly different from those listed in the cardiovascular space, and there might even be more than are currently categorized under gastrointestinal.

So again, I can't make a comparison as to how this relates to the space in 2008. I don't think anyone can say, well, if there had not been a guidance in place, we'd have X many more. But there seems to be continued interest, I think, in developing new drugs for diabetes. And it would be, I think, helpful to have a fuller conversation about this.

Now, there may be additional disincentives that are either directly or indirectly related to the guidance requirements that are separate from the costs associated with conducting clinical trials. What we may find is that once there are 5 or 6 drugs available in a given class that have already been found safe or potentially beneficial, yes, there may be fewer "me too" drugs that are

developed, but I'm not sure that's actually a problem for future drug development.

I think there's been a lot of uncertainty about the path forward for drugs that are studied in this CVOT and found to be safe from a cardiovascular perspective, but don't necessarily reduce the risk of cardiovascular events. But in my opinion, these agents remain clinically relevant.

I think, being someone who's managed patients with diabetes for decades now, I would say we rarely have the luxury of managing diabetes adequately over the long term with one or two drugs. It just doesn't really work that way, so people end up requiring very complex medication regimens, multiple drugs over time, and changing drugs over time depending on their other health issues.

It's important to understand the safety, as best we can, of all of the agents that are being used. It's probably not adequate to say, well, you're on this one drug that's good for you. The

others are irrelevant. So in fact, we really need to understand as best we can the entire platform or portfolio of drugs available to us that we're using in combinations.

I think that the current guidelines are doing a better job of outlining the role of these agents in the care of people with type 2 diabetes. And certainly, I think it's fair to assume that if a drug appears to be safe from a cardiovascular perspective in a high-risk patient group, then it's probably safe for people at lower risk and could be used with confidence there. And also, these drugs could certainly serve as a component of anti-hyperglycemic care for the higher-risk patient, but wouldn't substitute for perhaps a beneficial drug.

Again, it's a bit hard for me individually to understand the numbers of patients in the U.S., for example, on a given drug at this time and whether or not they're being prescribed to appropriate patient populations. But I am concerned that a key disincentive exists in the fact that there must be underutilization of this

data and of beneficial drugs in clinical practice, and I think that is an area that deserves our attention.

Some information that's relevant to this comes from the Harmony Outcomes trial, which again was a CVOT of albiglutide versus placebo, but was a fairly contemporary trial, so patients were enrolled between late 2015 and early 2018. And every person in this trial had type 2 diabetes and established atherosclerotic cardiovascular disease.

You can see in this trial -- this is data from the placebo arm -- that at baseline, the use of SGLT2 inhibitors in this patient population was miniscule. And although it increased a bit during the study, still very, very low percentages of patients were on drugs for which, in theory, they had an indication.

Finally, the ugly. This is a recent tweet from a colleague whom I've de-identified for today's purposes, but I think there has been a tendency with this increased emphasis on cardiovascular risk reduction, which is incredibly

important, but there is a tendency to devalue the importance and benefits of glycemic control in diabetes management.

This is just a reminder, since I have the floor today, that this is not an area that should be forgotten in the mix of diabetes care. And although as I alluded to previously, it's certainly clear that with the drugs that we had traditionally available, that a very intensive glycemic control strategy did not reduce the risk of cardiovascular complications in high-risk patients.

Across the board, we see in high-CV-risk patients, in patients with newly diagnosed type 2 diabetes, in patients with type 1 diabetes, tighter glycemic control, consistency reduces the risk of microvascular complications, appears to reduce the long-term risk of cardiovascular complications as well in people with newly diagnosed type 2 diabetes. So these are important findings.

Because the benefits of tighter glycemic control on microvascular outcomes has been demonstrated so consistently across trials, of

course the initial approval of diabetes drugs has been based on A1c lowering, as that is anticipated to clearly demonstrate into a microvascular benefit.

The other question that comes up a lot is if there's a certain point where glycemic control contributes less or becomes unimportant to cardiovascular risk, and we certainly don't know when that switch is flipped, if it truly is. And in fact, the older trials that suggest that really might be findings that are only true of older drugs and may not be applicable to the current medications that we have available.

Now, the drugs that have been studied so far that have demonstrated a cardiovascular benefit clearly do not reduce cardiovascular risk through glucose lowering. That's not the primary mechanism of benefit. But there's still been some degree of better glycemic control in the patients who received active therapies compared to placebo in all these trials.

People have debated to what extent the

better glycemic control has contributed to between-group differences in macrovascular outcomes. I don't think anyone has those answers, but it's probably a non-zero effect.

The other thing that is really important to consider, too, is the issue of competing risks. And as we become better at preventing major adverse cardiovascular events and deaths in people with type 2 diabetes, and they live longer with diabetes, then in fact these traditional microvascular complications that we know are glycemia dependent may in fact be of increasing importance to the health and quality of life of people with type 2 diabetes.

We know that we've been really quite effective in reducing the rates of various important events in people with type 2 diabetes over time. To some extent, these reductions in risk of MI, stroke, and amputation are probably, to some extent, dependent upon the increasing recognition of the benefits of tighter glycemic control. You can see the greatest rate of change

occurred after availability of the initial UKPDS results.

Obviously, there were other things that changed in clinical care with respect to guidelines for blood pressure and lipid management that also contributed, but we don't want to negate the importance of glycemic control in assuring our patients' long-term health.

In fact, the idea that cardiovascular outcomes should supersede any importance of glycemic control is a false construct. And in fact what we have are multiple overlapping outcomes of interest that in fact are largely complementary. And it's just not important for us to only think about one aspect of this care. We can think about and treat more than one thing at a time.

In fact, such an approach, again, all with older drugs in the Steno-2 trial, has demonstrated that pursuing and treating multiple risk factors for micro and macrovascular complications -- so treating glycemia, blood pressure, lipids more aggressively -- does result in reduced risk of

micro and macrovascular complications over time. So again, these are complementary strategies.

Moving to the future, I just want to echo some of what has been expressed earlier today by the prior speakers. And in my opinion, I would suggest that adequately powered and randomized CVOTs of individual and/or hyperglycemic agents should continue.

I think there really is no reliable substitute for this information, and it allows patients and providers to understand the effects of available drugs and make informed decisions regarding their care. But work is still needed with respect to implementation, and again, more efficient but still robust trials and methodologies.

Thinking about the implementation, I think, again, this is my area of concern at present because I don't think, although the guidelines are recommending the use of these drugs preferentially in people with established cardiovascular complications, that locally, it's really an

exception for an institution or a healthcare system to expect that this is going to happen and be part of standard of care for these high-risk patients.

There are lots of reasons for that, and I think there may be unawareness or generally limited awareness of the trial findings. People are confused about how to apply this data in clinical practice. But there are a number of unanswered clinical questions that probably contribute and need to be addressed by other bodies.

I think there's also a real factor of inability to appropriately assess and weigh the balance between benefits and risks of given drugs. And I hate to say overblown, but disproportionate concern about the risks of very rare but serious side effects, I think, serves as a barrier to implementation of effective therapies.

We obviously have the eternal cost and access issues and the amount of time that it just takes for people to learn about, discuss, and implement new features of care, but these are not a justification.

I think that there are a couple of ways that this can be addressed and that we can ensure that all of the work in this space designed to ensure patient safety is actually being translated into clinical care. And again, I think we need to establish relevant local care expectations and quality measures to ensure that this is translated into clinical action.

I think that guidelines also should be very readily understandable, summarizable, applicable, and need to consider the audience. In the United States, it's estimated that only 15 percent of people with diabetes see a specialist for their diabetes care, so the audience is primary care. And if we are writing guidelines so that they're relevant, or memorable, or translatable only by specialists, it's going to be a problem.

I think we should definitely avoid unnecessary complexity and cost, particularly if it's not evidence based. And I provide this block of text from the 2018 ADA guidelines because I think this is a great example of an understandable

guidance. And again, in patients with type 2 diabetes and established atherosclerotic cardiovascular disease, that after metformin, their anti-hyperglycemic therapy should incorporate an agent to reduce MACE and cardiovascular mortality. And I think that's understandable and readily useable by the majority of physicians who are treating patients with diabetes.

I think it's going to require revision of traditional roles. And this is a little dramatic, and I'm sure that diabetologists and cardiologists are not this different and ignorant of their areas of responsibility, but I think traditionally, yes, diabetes care providers focused on blood sugar, focused on reducing risk of microvascular complications, might defer to a cardiologist to implement other cardiovascular risk reduction strategies. And the cardiologists, on the other hand, focuses on management of hypertension, lipids, established atherosclerotic complications, and might defer to the diabetologist regarding any changes that might be indicated with their diabetes

medications.

There has been a novel paradigm for care. This is an example of one for the management of patients with type 2 diabetes and cardiovascular disease that we should all consider. In particular, cardiologists at minimum could screen their patients with cardiovascular complications for diabetes if they don't already have a diagnosis.

If they have patients with cardiovascular disease and a previous diagnosis of type 2 diabetes, we really need to start discussing more actively whether or not they have indications for specific anti-hyperglycemic therapies. And this will require collaboration between specialists and a multitude of diabetes care providers to implement successfully.

I think some of these other barriers, again, are related to the fact that there is still existing questions and missing pieces with respect to the data that's been accumulated so far. And as has already been expressed earlier today, we'd like

to know more about the effects of these drugs in lower risk or underrepresented populations, people who didn't participate in the CVOTs to date in large numbers.

Again, I think we need to better define high risk. What's the threshold? Do we wait until someone's actually had an MI before we changed their therapy, or can we be a bit more sophisticated about prediction? As has, again, already been mentioned, I think there's a lot of interest in doing trials with active head-to-head comparisons in combinations of drugs and a desire to better understand longer-term effects and the best place in therapy for use of these drugs. But this is going to require the engagement of other agencies, institutions, societies, interested groups, other stakeholders, I think, to be fully implemented.

Again, this is going to be a bit redundant, given the prior discussions, but I think, again, that the guidance and the fundamental principles expressed in the 2008 guidance remain incredibly

important. However, the previously used CVOT model designed to satisfy these requirements, although very effective so far, shouldn't be the only acceptable path forward.

Again, we need to think about new ways to design and operationalize these trials so that they're less costly, we can get answers more quickly. And again, as was given in an example of the dual primary endpoint approach, maybe we can maximize the ability to identify potential benefits of a given drug within a single trial if it's based appropriately on the expected physiologic benefits of a drug or the experience of other drugs within the class.

Finally, we can always think about a possible new paradigm for diabetes drug approval in the future, and there have been some who have suggested that if a drug demonstrates a meaningful benefit with respect to other important outcomes, perhaps a drug could be approved for that reason rather than based on its effect on A1c. And you may in turn then reduce the requirements for the

battery of glycemia trials that accompanies that. So perhaps we can think about other areas of determining benefit and efficacy of new drugs.

That's the end of my slide presentation. I'd be happy to answer any questions that you have in the time remaining.

Clarifying Questions for Dr. Green

DR. WILSON: Thanks very much. I have a question. So the TIMI study group led the SAVOR trial, and then we had EXAMINE, which was another class and gliptin, and you were the one who reported on TECOS. This committee met several years ago and reviewed the first two, and then waited for TECOS to see what happened, and you had a null result for heart failure.

In the future, how could we get there quicker? And that's trying to assemble some of what we heard from Dr. Ratner, some of what we heard from Dr. Sabatine, and some of what you have.

Could we use consideration, for instance, of standardized heart failure criteria for those studies that was just emerging? And I can see, in

the future, what will emerge for yet other new definitions of perhaps even heart failure.

Could we abuse Bayesian analyses? Could we have immediately, as the data was available, posted the data on, for instance, an FDA special closed intranet, so to speak?

The reason I'm saying this is -- and this is what's happening in other fields. Genetics is one of the great examples. As soon as data are coming out in major, major studies, like U.K. Biobank, they're immediately collaborating and accelerating the path of scientific discovery, and care, and decision-making.

You must have thought about this. Are there other ways; for instance, DCRI could collaborate and foster this, the same way the TIMI group might, to move this field faster?

DR. GREEN: I'm sure that there are ways to help better understand the effects of these drugs as a class. I'm not sure that, right off-hand, I can think of a great alternative for answering the question about the effects of a particular drug on

hospitalization for heart failure as an outcome. And I think that this particular class is particularly difficult to understand and come up with creative ideas around, because we don't know why these differences were seen.

If in fact some of these drugs increase the risk of hospitalization for heart failure, we really don't know the mechanism through which that occurs. So there aren't any good surrogate outcomes or pieces of information that we might be able to use as an earlier signal.

Some of what you've mentioned I think lends itself to better understanding and reconciling differences in outcomes that are noted across trials. And I think it's incredibly important that, not just within the cardiology space, but within other aspects of drug safety, that we really strive to standardize the way that we collect information about patients and the types of events that occur to them during a trial so that what we're actually comparing is apples and oranges.

I think, with respect to hospitalization for

heart failure, if you look at the definitions that have been used, they're more alike than they are different, but they're not identical. And whether or not that, to some degree, has affected or contributed to differences between trials, I can't answer.

So I think standardizing and enhancing the way that we capture information about patients, both at baseline and throughout the trial, with respect to these particular areas of interesting, is an easy, easy next step, that should be helpful.

DR. WILSON: Dr. Grunberger?

DR. GRUNBERGER: Thank you very much, Dr. Green, for an excellent presentation. I certainly share with you the good, bad, and ugly. But as long as endocrinology is on the stage, I'd like to have a little bit of a different twist, because you talk about evolution of the guidance and guidelines with organizations.

You might recall in 2013, ACE decided to get away from the simplistic glucose-centric view. And we decided to focus on management of diabetes, a

comprehensive disease with a cardiovascular burden. And at that time, we changed it, and now we have these annual algorithms for comprehensive management of type 2 diabetes, which look first at lifestyle medication, treatment of obesity, prevention of diabetes, and controlling cardiovascular risk factors.

Then we talk about glycemia, but in 2013, a decision criticized by many, we decided to rank the drugs approved for glycemic control in order of preference. So this is not evidence based. This is imminence [indiscernible] based because we said so.

But my point is that we put GLP receptor agonist as the top choice. In 2014, in the update, when SGLTs became available, we put it on top also. The reason I'm saying that, as you're thinking about a bad, the cost of clinical trials, they unfortunately cost. We put GLP receptor agonists and SGLT2 inhibitors as the two top choices well before the first CVOT was finished.

So I was asked -- I remember when the LEADER

results were announced, I was asked by a journalist, "When will ACE change its guidelines?" And I said, "Are you near a computer? If so, look at ACE.com, and you tell me how we can put GLP receptor agonists and SGLT2 inhibitors higher than number 1 and 2." But the point is we didn't have to change the guidance once the CVOTs were announced because they were listed, from that perspective, as the two top choices already.

So I'm wondering, with all the wonderful things we've learned, the good stuff you've presented, has that really altered the practice of the management of diabetes among experts? Because obviously, the bad is that we have not had much impact, as you showed. So I'm just wondering how to position that in the knowledge we gain from these trials. Can we justify that expense?

DR. GREEN: That's a very complicated question. It's a great question. And I do want to make sure, before I attempt to answer as many pieces of that as possible, that I'm not here intending to criticize those who write guidelines.

That's an incredibly complicated process and very thoughtful process. And the intent was not to suggest that they were not adequate.

I guess the concern is, if you start recommending more expensive therapies without demonstrated outcomes benefit -- and I know that some of what factored in here was the reduced likelihood that these drugs would cause hypoglycemia, weight gain, et cetera -- if that's what you focus upon, that's almost undoubtedly going to increase the cost of diabetes care overall as those recommendations are implemented broadly across the population of individuals with type 2 diabetes.

So those are non-trivial decisions if you decide that you're going to adopt the use of those expensive newer drugs early on.

Now, that may not be wrong, particularly if you have particular areas of concern that are readily addressed by the information available about these drugs. However, I think we need to be very careful not to assume that drugs that convey,

in general, weight loss, lower blood pressure, or lipids a little bit across the board are necessarily going to have that translate into a cardiovascular benefit or cardiovascular risk reduction.

In fact, when you look at the metabolic effects of the drugs that have been shown to convey a cardiovascular benefit in trials, it looks like the effects of these drugs are largely not explainable by those metabolic benefits. And I don't think there's any solid evidence to suggest that the reduction in cardiovascular risk was attributed to fewer hypoglycemic events.

So you do go out on a limb, I think, when you decide to priorities use of drugs based upon their effects on surrogate or intermediate outcomes. Now, those are non-trivial, and I think patient preferences and lifestyle and how it fits best into their overall care management, those are all important features of care. But I think with prioritizing expensive drugs, drugs that haven't been prescribed to a lot of people early in the

algorithm, does expose you to some risk because you don't have all the information available, and it increases cost.

So you were right, but you might have also been a little lucky.

DR. WILSON: Dr. Blaha?

DR. BLAHA: Yes. I have a question, just a clarifying question, about your last slide, which I thought was intriguing, where you said a possible new paradigm for diabetes drug approval, is A1c required, lowering required, if benefit demonstrated other meaningful outcomes.

I guess I'm trying to imagine what those outcomes would be to maybe reconcile with what you just said. As far as other outcomes, I don't know if that would be triglyceride lowering, or obesity lowering, or what you're picturing there.

DR. GREEN: Right. I think there are a lot of drugs in development that can -- go ahead.

DR. BLAHA: Just to finish that thought, I guess what I come to is what constitutes a diabetes drug, then? And if you could comment on that.

DR. GREEN: I think what constitutes a diabetes drug would be a drug that when administered to people with diabetes meaningfully reduces their risk of important complications, whether that be -- or risk of death, risk of major adverse cardiovascular events, risk of, essentially, progression to end-stage renal disease.

A drug that has been tested and demonstrated beneficial in that patient population, I would consider to be a diabetes drug, but one that may work through mechanisms other than A1c lowering. I know that's not the current standard, but perhaps something that could be considered in the future.

If you have a drug that demonstrate that you've really focused on that as the benefit and the primary outcome, maybe there just won't be so much work, and expense, and time required to do all the little shorter-term trials looking at how this drug compares to insulin, how this drug compares to every other oral agent on the market; how this drug works in combination with another glycemic-lowering

drug. That just may be a way to enhance and increase efficiencies.

DR. WILSON: Dr. Newman, did you have a question? No? Dr. de Lemos.

DR. DE LEMOS: I know I'm not supposed to talk about this, but you just brought us back to questioning A1c as a surrogate. Your last part said there had to be some meaningful impact on a patient-related outcome.

DR. GREEN: No, no, no. I am absolutely not questioning A1c as a surrogate, but it --

DR. DE LEMOS: Well, help me. How do I know that a drug that is perfectly safe from a cardiac standpoint, that lowers A1c by half a percent, and shows no demonstrable benefit on, say, weight or any patient-related quality of life outcome, how do I know that drug is helping my patient?

How does that drug merit approval in a space where we have so many drugs that offer so many benefits either on measurable cardiovascular outcomes or on other patient-related outcomes like obesity that we know are meaningful even if the

cardiovascular outcome hasn't been demonstrated?

It just seems like such an incredibly low bar to a cardiologist. It's just mind boggling to me. And help me out. Is there no way to get an insight into microvascular complications? I mean, we saw with semaglutide, right, where the A1c goes down and the microvascular complications go up in a short-term trial. Is there really no way to understand this in the context of a trial so skeptics like me can be confident that we're not giving patients useless drugs?

DR. GREEN: Again, there are many aspects to that question. First of all, none of these more recent trials were designed to assess the effect of the intervention on microvascular outcomes, and as available, insight or snapshot into the various parameters that were collected. In some trials, there essentially was nothing other than looking for major safety signals with respect to complications like retinopathy, looking at eGFR over time.

So we have very little really detailed

information regarding the effects of the newer drugs on microvascular outcomes per se. However, it's very clear that tighter glycemic controls, so somebody with a lower A1c over time, is going to have a reduced risk of microvascular complications compared to someone with a higher A1c. And that's been demonstrated consistently in trials designed to answer that question.

Now, you may ask or you may think, in general, this drug on average reduces A1c by 0.5 percent compared to placebo. How can that be important? I think it's important to remember that that's the average. Right? And so for a given person who starts a drug, they may have a more or less pronounced response to that.

It's important that we don't just put people on drugs and fail to assess their response to it. If they have a meaningful improvement individually in A1c and it's a drug that's approved for glycemic control, then I think it's a reasonable extrapolation to assume that they're going to derive a longer-term microvascular benefit.

I think this is another argument to follow patients for longer periods of time in some way, shape, or form after the CVOTs are completed so that we can gain greater insight into the effects of the interventions on microvascular complications because they may take a very long time to really declare themselves as being either increased or decreased compared to the treatment arm to which they are assigned.

So I think we just have very little information from the recent trials, and they haven't gone on long enough for us to have really reliable insights into the effect of the drug on glycemic-mediated microvascular complications.

DR. WILSON: What we are going to do now is we have two more questioners, Dr. Robbins and Dr. Ellenberg. Then we're going to call Dr. Ratner back up, give her a little bit of a break for a couple questions. And it's really going to be for both the questions -- we've held off from Dr. Ratner -- as well as Dr. Green if she's available for the rest of this time.

DR. GREEN: Sure.

DR. WILSON: We're front-loading those, and then we had some carryover questions for FDA from this morning. And hopefully, we'll do all that in a timeline, but the carryover for the FDA, I believe we could continue tomorrow.

Is that fair to say, Dr. Chong? If we don't get to those carryover questions, we could take those tomorrow?

DR. CHONG: You are the chair. You are the boss.

Additional Clarifying Questions

DR. WILSON: Yes. We'll do that then. I'm concerned about we're supposed to end at 3:30 according to the schedule here, and I'm not sure we'll get all that in, in the next half-hour. But Dr. Green and Dr. Ratner take the highest priority because they're here specially just for this.

Just to remind you, our guidance question is evaluating cardiovascular risk and antidiabetic therapies in type 2 diabetes treatment. They're reminding me of what the mission is of this

meeting.

Dr. Robbins?

DR. ROBBINS: Thank you.

Just briefly, about the increase in the small vessel disease, we've seen that before, the Kroc study, going back to the 1970s, where they treated patients with insulin with retinopathy. There was worsening of it, and it gets better. I'm not saying that's the explanation for this, but I would just be calm right now and wait.

More importantly, I want to talk about a signal that I think we're missing. There is a common denominator in all of these trials that have reduced heart disease that has been somewhat discrepant from the change in hemoglobin A1c. And I'm going to make a big fat postulate here that, someone, later on you may ride me out of town.

All of these studies have an inordinate impact on post-prandial hyperglycemia. The Honolulu heart study 20 years ago showed that men with normal glucose tolerance that have high post-prandials have almost the same heart disease

rate as diabetic patients. The STOP-NIDDM trial, which was an Acarbose study, not designed to look at heart disease, but was given to people with prediabetes, had a 50 percent reduction in cardiovascular death with a hemoglobin A1c drop of 0.3 percent. And there was just another Acarbose study that showed a similar result.

Likewise, the GLP-1 agonists and the SGLT2s have a very high impact on post-prandial glucose, really an area of diabetic science that has been understudied because most of our work has been done in the fasting state.

The reason I'm harping on this is that we do have the tools now to look at this. With continuous glucose monitoring, and time, and range, we can begin to look at volatility and these peaks and valleys, if you will, as being cardiotoxic. I think this could be very helpful to us in identifying the drugs that are up and coming that have the drop in hemoglobin A1c, which might be modest, but a strong signal for cardiovascular safety.

DR. WILSON: I'm not sure. That was a comment. Do you want to make a further comment on that?

DR. ROBBINS: The question is, what do you think about that?

DR. GREEN: I think these trials are not designed to readily assess the contribution of post-prandial glycemic control to overall outcomes. Certainly, you could do CGM testing in a subset, but I think it would be unlikely that you could correlate those findings with the actual events that occurred within the trial, but it certainly might be of some interest.

I would also note, too, that there tended to be fairly broad inclusion criteria for many of these trials, and they had very wide ranges of A1c and eligibility at baseline. The effect of the contribution of post-prandial hyperglycemia to the A1c is really going to vary quite a bit depending on where you start out.

So I think it's an interesting question, I think difficult to answer in the context of the

trials designed to satisfy the guidance.

DR. WILSON: Dr. Ellenberg?

DR. ELLENBERG: So in response to a question positing a drug with a half a percent drop in A1c, you made a reference to a meaningful improvement in hemoglobin A1c. And I wondered if there is some understanding of what a meaningful improvement on an individual basis is?

DR. GREEN: I think that the FDA has established criteria for what a meaningful hemoglobin A1c reduction needs to be for a drug to be marketed for those purposes. And that may be different from what is considered clinically meaningful.

Certainly, we'd like any intervention to lower A1c to the range that is desirable for our patient, and often that's considerably higher than a hemoglobin A1c difference of 0.3 percent, for example. But that doesn't mean that that's an inappropriate definition of meaningful change. And again, we're talking about either mean or median differences in A1cs achieved, and there can be

within that group of responders quite a bit of heterogeneity as far as how they respond to a given intervention.

So I think that's a difficult line to draw, and I think I'd have to leave it to the experts assembled here today to decide what's meaningful from the FDA's perspective.

DR. WILSON: I'm sure we'll have time to discuss some more of that. Follow-up questions for Dr. Ratner, if they still are in your memory?

Dr. Yanovski? She has been answered, before Dr. Ratner comes to the podium.

Dr. Lumley, did you have a follow-up you wanted?

DR. LUMLEY: As a patient, I was moved by some of his comments, especially the idea of kids. And I'm not sure if this is in the realm of what the FDA can do. But one very short story, this is kind of an ain't it awful story.

I was a school administrator, a district school administrator, for a fairly large district, Kansas City. And they wanted me to do the health

team, to chair the health team, so I did. So I brought in three or four docs from the Kansas City area to talk to my teachers and board members and administrators. And they talked about how there's so many kids now on type 2 diabetes, and they're heavy, and they're not exercising. They were wonderful, and everybody walked away and we said, "Boy, that's the way to go, prevention looking through the lens of prevention."

So I was all thrilled and happy and thought everything was going to go well. Then I ran into parents, and parents wanted the kids to bring cookies to birthdays and to have candy. I ran into politicians who thought we shouldn't take soda out of the machines in the schools. And I ran into kids who would throw away all their vegetables, literally. They would walk over -- we were required to put vegetables on the plate, and they'd throw them away.

So in short, in Kansas City, it was an epidemic of kids. We cut down on P.E. and increased with no child left untested. All we did

was science, math, and language arts, and we cut out history and P.E. So I guess my question to Dr. Ratner; what the hell should I do?

(Laughter.)

DR. RATNER: Where do I begin? I wish I had the cure for childhood obesity. I think that we've gone from an era where food was expensive and exercise was mandatory for work, to an era where food is cheap and you now have to pay in order to get your exercise in a health club.

The genetics haven't changed in the last 30 years. The environment has. The availability of inexpensive, high-calorie, high-density foods has been disastrous. The fact that individuals are working two and three jobs to make do, and run by KFC to pick up dinner for the kids is unavoidable when it's that cheap. I wonder what would happen if there were a $10 surcharge on every Big Mac.

DR. WILSON: We have another follow-up question for Dr. Green. Dr. Everett?

DR. EVERETT: Thank you. Brendan Everett from cardiology.

Dr. Green, I just wonder, one of the things that we're tasked with thinking about over the course of today and tomorrow is whether or not, by mandate, every single new diabetes drug should go through the process of a cardiovascular outcome trial.

I just was wondering if you could give us your insight as a trialist who's done a lot of really important work in this area as to whether or not there might be a method of ascertaining a particular cardiovascular risk, for example, in a smaller trial that would then lead you to believe that a larger trial, probably predicated with the focus on safety, so noninferiority, versus a trial where the hypothesis was that there was actually a cardiovascular benefit.

Is there something you can think of in your experience that might allow you to downsize what has become a large, as we heard from you and from Dr. Sabatine, operation, to something that is perhaps more selective in terms of what has to then go through the 10, 15, 20,000 patient multi-center

randomized trial?

DR. GREEN: Yes. That's the million dollar question, isn't it, or $500 million dollar question. I think it all comes down to numbers of events that are available for the analysis. And when you have 75 events, for example, it's very, very difficult to draw from conclusions.

I think in the preceding presentation, there were really good examples of why even using meta-analyses of events collected in smaller trials did not provide results that were confirmed in a dedicated cardiovascular outcomes trial.

So at present, I'm not sure that I can think of a terrific substitute. There are ways that we could potentially think about, again, doing the large-scale cardiovascular outcome trials more efficiently and perhaps trying to assess the effects of these drugs really as the patients being randomized to their therapy, embedded as fully as possible into usual care and perhaps thinking very carefully about the patient population being enrolled.

There's going to need to be some, I think, inherent element of risk in the patient populations studied so that these trials will accumulate enough events in a reasonable period of time that we can answer the question. So I think maybe if you could identify a lot of people within, for example, a health system and systematically assign them to one drug in a class versus another, and follow that large number of people for maybe not a terribly long time, you might be able to answer these questions a bit more rapidly, more efficiently, and maybe in a patient population that people would find more readily akin to those in their clinic practice.

But I don't know how to answer this question reliably without that. I'm not sure that I've heard a great proposal for doing the trial or answering these questions earlier on in the process that has really stood up to criticism. So not that I'm aware of, but certainly a great topic for further discussion.

DR. WILSON: So if I could follow up on that

one, Dr. Green. We often think of trials as being 3 to 5 years with a certain end. And reflecting on your experience in TECOS and others, I know you know very well that the TIMI trial results could have a lot bigger denominator and do these studies in 2 years. The simple question, sort of rephrasing Dr. Everett's.

DR. GREEN: Sure. I think, absolutely. You could have a lot of people in. I think you'd still want to build in a finite minimum period of follow-up to really be able to understand the effect of your drug other than in the short term. And people will be taking these drugs for years, and years, and years. So we need to accumulate some modicum or some semblance of what will happen in the clinical use of these drugs.

But I think very, very large trials of select patient populations could allow you to answer the question with a minimum length of follow-up.

DR. WILSON: Dr. Ratner, you wanted to add something?

DR. RATNER: Yes, if I might make a comment to both of the questions. I think that the critical issue for this committee is really whether or not the cardiovascular outcome trials, as provided in the 2008 guidance, should remain mandatory. I don't think anyone would deny the fact that we've learned an enormous amount of information from these studies. But if a sponsor wants superiority, clearly, it's required.

If all we're doing is ensuring safety, then the question that you asked, Dr. Wilson, I think is critically important. What is the minimal amount necessary to feel safe? Clearly, the FDA does that with every other safety indication within the diabetes field. They do it for all of the other fields as well, whether you're talking about liver disease, or lung disease, or rheumatoid arthritis. But diabetes is the only one where there's a requisite study to show cardiovascular safety, and I think we fall back on the issue of what's the signal.

DR. WILSON: Ms. McCollister-Slipp, you had

a question for either of them?

MS. McCOLLISTER-SLIPP: Yes. I guess this is most likely for Dr. Ratner, but Dr. Green, if you have an answer, and maybe neither of you do.

Has anyone done any economic analysis of the impact of the requirement for CVOTs on the cost of newer drugs?

DR. RATNER: We know the cost of the CVOTs is somewhere in the vicinity of $5 [billion] to $6 billion total. The impact on drug development, the impact on cost of drugs is much, much more complicated. I'm not sure that you can draw a straight line between the cost of the studies as Dr. Green showed.

DR. GREEN: I struggled to find that information myself, and I don't think anybody's really got a handle on it. You can access various pieces of information, but no, we certainly, I don't think, understand the full scope.

DR. WILSON: Dr. Wang had a question for either of our speakers.

DR. WANG: Yes, a question for Dr. Green,

again, a trialist question. One of the things we're asked to comment on is the requirement, if we're continuing this mandate for adjudication committees. In the context of all the discussion we've had regarding, on the one hand, the importance of standardizing the endpoints, our endpoints are clean.

But on the other, embedding these trials perhaps in a medical system so we can make them simpler, which at its extreme is basically just extracting diagnoses out of a medical record or administrative code, which doesn't, per se, require adjudication, what is your view on the part of the mandate that requires adjudication committees?

DR. GREEN: I think some form of adjudication will need to be a component of the trials, but I think the way that adjudication is performed is ripe for refinement, evolution, et cetera.

I think it really, again, sort of boils down the way that you collect information to ensure that you're collecting a standardized and complete set

of information and the greatest level of detail regarding the event of interest as possible.

If you're successful and these pieces of information are directly applicable to the definitions that you're using during your adjudication process, there are people who are really interested in looking at machine learning as a potential means of performing the adjudication with some sample or some percentage of the events reviewed by humans to make sure that it makes sense.

But I think there are a lot of people interested in doing this more efficiently, more consistently, and better, but I don't think it's going to be satisfactory or reliable to just collect -- for example, events based on ICD codes that pop up in people's health record, We know that's an area that's fraught with errors and omissions.

So I think that, yes, adjudication needs to be part of it, but there's no one right way to do adjudication, and I think we can probably do it

better.

DR. WILSON: Dr. Rosenberg, you had a question?

DR. ROSENBERG: Yes, a question or comment following Dr. Ratner's comment, the need for studies for safety and not efficacy? I have a problem differentiating the goals when the criteria for approval is a surrogate like we have here.

When we know and we have the example of the CPT inhibitors that were mentioned earlier, it's only if you have long enough and large enough studies, which are mostly designed for efficacy, your long-term goal, that you can detect those safety signals. So if you really lower the requirement too much, one, you will never detect those signals. They're completely unexpected, like in those studies.

If you don't have the motivation to do a long-time study with an efficacy outcome, you will never be able to do that, or you may risk missing a lot. And if there's not a motivation also to have a long-term efficacy study, why would we bother

having 5-6, whatever drugs in the same class, "me too" drugs, continuing to be approved on HB A1c if they don't have further benefit.

I'm struggling with that. I think that will be further discussed tomorrow.

DR. WILSON: That was mostly a comment, I believe. Dr. Budnitz?

CAPT BUDNITZ: I think this is for either one of you based on your experience with these CVOTs. We talk about streamlined studies and different ways to do that from maybe the -- it was trivial for triple-checking kind of forms, maybe we don't need to do so much of that, to maybe new analytic methods that might shorten the time it would take.

Are we just kind of nipping around the edge and reducing the costs on the order of 10 percent, or is there fat, so to speak, to significantly cut out so that it could significantly reduce the cost of these trials in a meaningful way, but still have the outcomes?

DR. GREEN: Yes. It's my

understanding -- and I don't have the figures in front of me -- that, in fact, the expenditure on site monitoring is tremendous and actually represents a very significant chunk of the overall cost of performing one of these trials. So yes, I think there is opportunity there.

DR. WILSON: Dr. Kushner?

DR. KUSHNER: Just curious. Besides using Bayesian techniques and possibly using healthcare data -- and I'm not sure how we could adjudicate healthcare systems data with all of the problems with the data quality -- is there a way to get more events from the earlier trials so that we could be confident that a safety signal wouldn't be passed up?

DR. RATNER: Yes. I think there is, and I made the comment towards the end of my presentation that the very simple issue would be loosening the inclusion/exclusion criteria for the regulatory trials. And this is something that the agency can actually mandate, to a certain degree, that they have to have individuals up to the age of X, that a

certain percentage has to be older than Y, or that a certain percentage needs to have risk factors or other confounders.

What you're doing, then, is you're actually getting a more generalizable population for the regulatory trials. Now, clearly, what's happened in the past is that without that requirement, then every company wants to have as clean a study as possible with no other confounders and have a minimal number of adverse events that they have to report, and that becomes the trade-off.

DR. KUSHNER: The follow-up to that is, is there a way to reanalyze some of the data from some of these trials to look at whether changing -- as an intellectual process, changing those inclusion/exclusion criteria would have changed a signal prior to the onset of the CVOT?

DR. RATNER: I don't know how you would do that as an intellectual exercise. I don't deny the fact that it might be possible. I think that one of the observations from the CVOTs is this very dichotomous response in those individuals who are

over the age of 60 with risk factors, but no CVD, versus over the age of 50 with CVD. You really have a quite substantial differential response there.

I think that, moving forward, if companies want to really enrich their population with events, you just have 100 percent with established CVD, and that clearly increases your event rate. But as I point out, that really restricts the generalizability of the findings.

We need to keep in mind that unlike an acute myocardial infarction, or an acute stroke, or death, we're dealing with a chronic disease. So we're talking about decades of experience with the disease as opposed to months or even years.

DR. WILSON: Dr. Wang, did you have a question?

DR. WANG: I'm sorry. I just had a quick follow-up question, actually, to Dr. Rosenberg's question, but I wanted to get actually either of your thoughts as an endocrinologist.

If there's another GLP medication coming to

market that has, based on the non-CVOT data, similar hemoglobin A1c lowering and similar safety profile to all the other drugs on the market, yet I don't have a CVOT, I don't have an outcomes trial to tell me whether it has any impact on cardiovascular risk, as an endocrinologist, would there ever be a reason to prescribe that drug, given that there are other GLP-1 agonists that have been shown to lower cardiovascular?

DR. GREEN: Sure. I can think of two really good examples and common examples. One is, let's say that that drug was the one that was available to your patient on the formulary of their insurance company and they didn't have established cardiovascular disease, in which case you can prescribe for them any available drug with confidence provided they have no contraindications.

The other instance in which you might well use that medication, if it provided some other benefit or was easier for the patient to get or tolerate, would be in a high-risk patient, so a patient with type 2 diabetes and established

cardiovascular disease, who for example was on an SGLT2 inhibitor with a demonstrated cardiovascular outcome benefit. And there's no proven additive benefit from use of drugs from the various classes that's been demonstrated in clinical trials at this time.

So it could be a component of such an individual's care, but not a substitute for a drug with a proven CV benefit, so absolutely, absolutely.

DR. RATNER: There's also a practical example of that with semaglutide. So semaglutide got its approval. It met the requirement for noninferiority. But in fact, if you look at the data, though it was not a pre-defined outcome, they even met not superiority, and yet there's not a label indication for semaglutide for CVD protection in the approved drug.

If they want that in the label, then it looks as though they're going to have to do a CVOT, and that's a decision that is a business decision. It's not one of, do we know or how are we going to

make the choice. The data are there. But in fact, the label does not say that it's approved for CVD benefit.

DR. WILSON: Dr. Fradkin?

DR. FRADKIN: So I think Dr. Ratner made his position very clear with regard to the need for safety studies, but I was less clear about your position.

DR. GREEN: I missed it when he expressed his position. I guess it boils down to, are you a lumper or a splitter, really, with respect to the data that's available from these trials. And I tend to be a splitter. I don't think it's unreasonable to expect to understand the safety and effects of each individual drug. And I understand the costs and potential problems that such a position might convey. But I'm not using a class of drugs. I have to prescribe a drug to someone. And I think it's good if I can, as accurately as possible, explain the risk-benefit ratio to the person who will be taking it.

So I'm personally in the splitter group, and

I'd like to know as much as possible about individual drugs.

Now, in the future, let's say there is a class where the effects are so reproducible from one drug to the next, with respect to all outcomes of interest, it might be a different conversation. But then how do you decide, oh, there have been -- what would it be, 3 trials, 4 trials? How many CVOTs showing that in a given class would be necessary for you to decide you didn't need to know as much about all the others that might come in that class?

I think these are difficult questions to answer, and at the present time, there's enough heterogeneity of outcomes in these trials that we really can't substantiate such an approach.

DR. FRADKIN: That actually was not my question, though. I wasn't asking whether you were a lumper or a splitter. I was asking you about the major question that the FDA is asking us.

In your talk, in the good, you really made a very compelling case for the fact that the decision

in 2008 has led to a paradigm shift in the way that we take care of diabetes, but that was really due to what was not at all expected in terms of the positive benefits that emerged from these studies.

So my question is, given that none of these cardiovascular outcome trials have shown harm, should there be a requirement for approval to demonstrate the safety? I'm just asking you what your opinion is on the basic question that the FDA is asking us.

DR. GREEN: Yes. The answer is, yes, I do think that it should continue. I think there certainly should be modifications and perhaps appendices, for example, to the guidance that can allow new approaches to the provision of this information that would be considered satisfactory. And I think that's an ongoing and evolving negotiation as to how you satisfied requirements.

But in my opinion, the fundamental principles and really the key elements of the requirements of the guidance remain critically important. Just because we have a better handle on

what 10 or so drugs do to people, I don't know that we want to make an enduring decision to do otherwise.

DR. WILSON: Thank you.

We have three last questions. We're closed for questions now. So let's have short questions and short responses, and respect where our speakers have been standing, especially Jennifer, for quite a while now.

So we have Dr. Ellenberg, Dr. Yanovski, and Dr. McCollister-Slipp, and then we're going to hear final comments from the FDA before we're --

DR. ELLENBERG: I'll be very brief because Dr. Wang asked my question.

DR. WILSON: Thank you. Susan Yanovski?

DR. YANOVSKI: Yes. My question is just to what degree -- let's say we start enrolling higher-risk patients in the phase 2/3 trials for safety. To what degree, then, can we use -- if a drug is in the same class, can we use those data to alleviate our need to do a dedicated CV outcome trials, if we're not getting any signals of anything new?

DR. RATNER: I personally don't have a numerical number for you, Dr. Yanovski. I think that the FDA certainly has the biostatistical support systems to come up with a better answer than they did with the goalposts of 1.3 and 1.8.

DR. WILSON: I'm not sure she's satisfied, but let's go on with it. Ms. McCollister-Slipp?

MS. McCOLLISTER-SLIPP: I'd like to ask both of you the question that I posed to Dr. Sabatine, about the big question that the FDA has asked us to evaluate, is should they continue to require these studies.

I was intrigued by his comparison with the RCTs with the observational studies. I mean, sure, there was a difference in the outcomes, but when you get down to statistics, what kind of numbers are we really talking about?

We've come a really long way in terms of being able to analyze, and being able to curate, and normalize sources for observational data and to develop statistical methods based off of that data.

So given the fact that you can pull from a

much larger data set many more people that don't have so tightly defined inclusion/exclusion criteria, and it's more indicative of what these drugs are like in the real world; and given the inconsistencies and the lack of reproducibility that we're seeing broadly speaking with clinical trials, what is the relative merit of continuing to do these versus relying on the latest version of observational studies?

DR. GREEN: I think as far as the observational analyses go, I think absent any form of randomization to therapy that would equalize utilization of various medications and minimize the likelihood of major and perhaps underappreciated differences between treatment groups, I just don't think that observational data, irrespective of the pool of patients to whom you have access, is a substitute for a randomized trial.

Aside from that -- and that's been discussed earlier today -- there are very serious issues regarding the type of data and completeness of data available through EHRs that have not been fully

resolved. There is significant miscoding/ undercoding, and there can be quite a bit of missingness of data, particularly with respect to a patient's vital status that really isn't available in the EHRs at all.

There can be a tendency in these analyses. If the fact that a person has died is missing from the data set, the assumption is made statistically that the patient is alive and doing well. So in fact, these are imperfect ways of answering the questions.

MS. McCOLLISTER-SLIPP: I would also argue that the RCTs are very imperfect ways of doing -- I mean, they're very beneficial, but they answer the very specific question that they were asked. They don't address the complexity that exists in the real world, and we're getting much better with statistical methods and evaluation of EHR data, at figuring out what some of those holes are.

So again, I haven't decided one way or the other, but I would think that -- I don't know. I'm just not convinced that the benefits of RCTs

outweigh the cost, and I'd love to be argued out of that position if that's possible.

DR. RATNER: I can't argue you out of that position. I think that there are pros and cons to RCTs. They're still considered to be the gold standard, but they're limited to the question asked and the populations studied. So the generalizability is significantly impaired.

I'll remind everyone there is no randomized-controlled trial to show that cigarette smoking causes lung cancer. It can't be done. It's not feasible and it's not ethical. So you look for other ways of answering the question if in fact the RCT, the gold standard, which I continue to accept, is problematic.

RCTs are absolutely necessary in my mind if you want to show superiority. If you're looking for safety signals, what better way of finding a safety signal than looking at the total population that's exposed. And I think that there are clearly potential advantages there.

I was very reassured when Dr. Sabatine

actually showed the data that I just alluded to with the CVD-REAL data, which shows that the real-world study looks really good compared to the CVOT. The example that he used with the Nurses Health Study and the Women's Health Initiative is really not entirely on point because the people who were in the Women's Health Initiative were randomized to estrogen therapy, where no one would ever give an 85-year-old, who's been off therapy for 30 years, to start.

So these are some of the limitations. I think one has to view the value of each approach. There are benefits to either approach, but ultimately, if you have a large population-based study, you're really generalizing the observation, and all of the founders actually begin to be modified. So there are benefits to both, and I think we need to utilize both.

DR. WILSON: I think it's time to stop at this point. I thank our speakers and thanks for all your efforts in answering the questions.

FDA has some final comments.

DR. CHONG: I have one request and a few comments. My understanding is that there are still some outstanding questions to the FDA that we're going to try to cover tomorrow. I would appreciate if we could find out if it's anything that would require additional data so that we have a chance to dig into that.

I wanted to thank our speakers as well. I think they did very good presentations, very thought-provoking presentations, and I want to thank the committee members for all their excellent questions. And if we could find out if we need to do any more work overnight, that would be helpful.

DR. WILSON: Any response to that?

(No response.)

DR. WILSON: Not right now. I'll ask one because it was brought up. Has there been -- for instance, for the gliptin trials that Dr. Green -- SAVOR, EXAMINE, and TECOS, has there been a post hoc analysis done by propensities -- and it might not be in the FDA's wheelhouse to have such data -- that would give us an indication for

the heart failure signal?

It's a rephrasing of the question that I asked her to start out. Is there a way to get there more efficiently with patient-specific data as these trials potentially go forward in the future? I don't think the FDA's going to have that data, and I don't think they've undertaken those analyses, though.

DR. CHONG: I think you're correct. I don't know that we've undertaken that sort of analysis. I think it's unlikely that we would be able to perform that sort of analysis before tomorrow morning.

DR. WILSON: For sure, but it was a question just in general because that topic was brought up several times today as part of their methods that might accelerate our ability to really understand safety at a higher level.

DR. ARCHDEACON: I was just trying to parse out what you were saying about, could we apply some propensities to that. And I was just trying to understand what that question meant.

I think what I could say is that when we looked at the individual trials, that randomization appeared to have worked, and so the things were equally distributed across the group. So I'm not sure what applying a propensity-score approach would add to that.

DR. WILSON: Could you use the individual data within the trial participants to predict who was especially going to develop heart failure? And it may have been related to their underlying burden of risk, not to which arm of the trial they were assigned to.

DR. ARCHDEACON: Right. I understand the question now. I don't have that --

DR. WILSON: And I think it's a challenge to the field. That may be one of the issues to help the field move forward, but I don't think we've seen those analyses at this point.

Yes? Martha?

DR. NASON: I don't know that this would actually be something you guys may need to get more data on or not, but I'll ask just in case. I am

still a little confused for myself in the trials you guys presented about what standard of care means as far as what other drugs they were allowed to take, regardless, assuming they were blinded and didn't know what they were taking, and whether they were just restricted from taking anything in the same class, or whether they could be prescribed by their provider other drugs if their A1c looked like it was high, for instance.

So I was wondering if you guys have data on that just to talk about how many people took how many drugs and how many different types during those trials.

DR. YANOFF: The standard of care was any drug other than what was in the class, including insulin. The data for the trials that have been FDA reviewed are in the original AC packages for each of those trials. And they're online, and we could pull them up if you wanted to look at any specific one. And then the ones that are not FDA -- finished being FDA reviewed and published, you might be able to find that information in the

papers.

We did look at medications added in the placebo versus the treatment group, and generally, there were more antidiabetic therapies added in the placebo group to try to get that A1c up. As I said, it generally did not reach the level of A1c control of the actual treatment group, but there was also no imbalance. There was no pattern in the drugs added that we felt was contributing to any risks or any findings found.

DR. WILSON: Any other comments from FDA? Mary?

DR. THANH HAI: Actually, at the risk of reopening Pandora's box, Dr. Green's last slide, I just feel like I need to provide some clarity to actually answer Dr. Green's question, but also to the panel members.

LaToya, if you don't mind putting up that last one. So it's that last question in terms of possible new paradigm. Is hemoglobin A1c lowering required of benefit demonstrated in other meaningful outcomes?

I just want to remind people, it goes back to the indication. What a company asks for in the indication is what they have to establish, and what they have to establish is based on the endpoint. So if the indication is as an adjunct to diet and exercise to improve glycemic control, then it's hemoglobin A1c. But if it happens to be diabetic retinopathy, that's the indication they're seeking, it's not hemoglobin A1c. It's the endpoint for diabetic retinopathy, diabetic nephropathy, diabetic neuropathy, and those certainly are conditions in which we have approved without it.

If a company comes in with an established antidiabetic that is a glucose-lowering drug, or the mechanism of action we understand is glucose lowering, but they don't want a diabetes claim with respect to glycemic control, that's their prerogative.

If they don't want that, but they want to go and explore another indication, we may still require that they assess glycemic measures from a safety perspective, but we wouldn't make them if

they have no interest in being a glycemic control agent.

I hope that provides some clarity.

Adjournment

DR. WILSON: We are now adjourned. We'll meet again tomorrow at 8:30.

(Whereupon, at 3:47 p.m., the meeting was adjourned.)