FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH

ENDOCRINOLOGIC AND METABOLIC DRUGS
ADVISORY COMMITTEE (EMDAC) MEETING

Tuesday, June 28, 2016
7:59 a.m. to 4:43 p.m.

Hilton Washington DC/Rockville
Hotel & Executive Meeting Center Plaza Ballroom
1750 Rockville Pike
Rockville, MD 20852
### 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)

**Daniel Budnitz, MD, MPH**  
Medical Officer, Division of Healthcare Quality Promotion  
Director, Medication Safety Program  
Centers for Disease Control and Prevention  
Atlanta, Georgia

**David W. Cooke, MD**  
Associate Professor of Pediatrics  
Interim Director, Pediatric Endocrinology  
Johns Hopkins University School of Medicine  
Baltimore, Maryland
Brendan M. Everett, MD, MPH
Assistant Professor of Medicine
Harvard Medical School
Director, General Cardiology Inpatient Service
Brigham and Women’s Hospital
Boston, Massachusetts

Diana Hallare, MPH
(Consumer Representative)
Visalia, California

Susan R. Heckbert, MD, PhD
Professor
Department of Epidemiology
University of Washington
Cardiovascular Health Research Unit
Seattle, Washington
William R. Hiatt, MD, FACP, FAHA
Professor of Medicine/Cardiology
University of Colorado School of Medicine
President CPC Clinical Research
Aurora, Colorado

James D. Neaton, PhD
Professor of Biostatistics
Division of Biostatistics
School of Public Health
University of Minnesota
Minneapolis, Minnesota

Robert J. Smith, MD
(Chairperson)
Professor of Medicine (Endocrinology)
The Warren Alpert School of Medicine
Professor of Health Services, Policy and Practice
Brown University
Providence, Rhode Island
Peter W. F. Wilson, MD

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
ENDOCRINOLOGIC AND METABOLIC DRUGS ADVISORY

COMMITTEE MEMBER (Nonvoting)

Reshma Kewalramani, MD, FASN
(Industry Representative)
Vice President, US Medical Organization
Thousand Oaks, California

TEMPORARY MEMBERS (Voting)

Leslie Cho, MD
Section Head, Preventive Cardiology and Rehabilitation
Director, Women’s Cardiovascular Center
Cleveland Clinic
Cleveland, Ohio

James De Lemos, MD
Professor of Medicine and Cardiology
University of Texas Southwestern Medical Center
Dallas, Texas
Judith Fradkin, MD
Director
Division of Diabetes, Endocrinology and Metabolic Diseases, NIDDK, NIH
Bethesda, Maryland

David Good, MD
Professor and Chair
Department of Neurology
Penn State Milton S. Hershey Medical Center
Hershey, Pennsylvania

Marvin A. Konstam, MD
Professor of Medicine, Tufts University
Chief Physician Executive
The Cardiovascular Center, Tufts Medical Center
Boston, Massachusetts
Melissa Li-Ng, MD
Consultant Endocrinologist
Cleveland Clinic Abu Dhabi
Abu Dhabi, United Arab Emirates
Staff Physician
Department of Endocrinology, Endocrinology and Metabolism Institute, Cleveland Clinic
Cleveland, Ohio

Richard Lumley, EdD
(Patient Representative)
Kansas City, Missouri

Kevin D. McBryde, MD
Medical Officer, Division of Extramural Research
National Institute of Dental and Craniofacial Research, NIH
Bethesda, Maryland
Paul Palevsky, MD
Chief, Renal Section
VA Pittsburgh Healthcare System
Professor of Medicine and Clinical & Translational Science, University of Pittsburgh School of Medicine
Pittsburgh, Pennsylvania

Michael Proschan, PhD
Mathematical Statistician
Biostatistics Research Branch
National Institute of Allergy and Infectious Diseases, NIH
Bethesda, Maryland

Yves Rosenberg, MD, MPH
Branch Chief, Atherothrombosis and Coronary Artery Disease Branch
Adult and Pediatric Cardiac Research Program
Division of Cardiovascular Services
National Heart, Lung, Blood Institute, NIH
Bethesda, Maryland
Morris Schambelan, MD
Professor Emeritus of Medicine
University of California San Francisco
San Francisco, California

Abraham Thomas, MD, MPH
Senior Vice President and Chair
Department of Medicine
New York University Lutheran
Brooklyn, New York

FDA PARTICIPANTS (Non-Voting)
Jean-Marc Guettier, MD
Director, Division of Metabolism and Endocrinology Products (DMEP)
Office of Drug Evaluation II (ODE-II)
Office of New Drugs (OND), CDER, FDA

Norman L. Stockbridge, MD, PhD
Director
Division of Cardiovascular and Renal Products
ODE-I, OND, CDER, FDA
William H. Chong, MD
Clinical Team Lead
DMEP, ODE-II, OND, CDER, FDA

Jennifer Clark, PhD
Mathematical Statistician
Division of Biometrics II, Office of Biostatistics
Office of Translational Sciences, CDER, FDA

Andreea Lungu, MD
Clinical Reviewer
DMEP, ODE II, OND, CDER, FDA
# CONTENTS

<table>
<thead>
<tr>
<th>AGENDA ITEM</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Call to Order and Introduction of Committee</td>
<td></td>
</tr>
<tr>
<td>Robert Smith, MD</td>
<td>14</td>
</tr>
<tr>
<td>Conflict of Interest Statement</td>
<td></td>
</tr>
<tr>
<td>LaToya Bonner, PharmD</td>
<td>20</td>
</tr>
<tr>
<td>FDA Introductory Remarks</td>
<td></td>
</tr>
<tr>
<td>Jean-Marc Guettier, MDCM</td>
<td>25</td>
</tr>
<tr>
<td><strong>Applicant Presentations – Boehringer Ingelheim</strong></td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td></td>
</tr>
<tr>
<td>Hans-Juergen Woerle, MD</td>
<td>37</td>
</tr>
<tr>
<td>Context and Background</td>
<td></td>
</tr>
<tr>
<td>Prof. Bernard Zinman</td>
<td>38</td>
</tr>
<tr>
<td>Design</td>
<td></td>
</tr>
<tr>
<td>Hans-Juergen Woerle, MD</td>
<td>43</td>
</tr>
<tr>
<td>Cardiovascular Outcomes</td>
<td></td>
</tr>
<tr>
<td>Uli Broedl, MD</td>
<td>53</td>
</tr>
<tr>
<td>Safety, Data Quality, and Integrity</td>
<td></td>
</tr>
<tr>
<td>Hans-Juergen Woerle, MD</td>
<td>66</td>
</tr>
<tr>
<td>Clinical Perspective</td>
<td></td>
</tr>
<tr>
<td>Prof. Bernard Zinman</td>
<td>75</td>
</tr>
<tr>
<td>AGENDA ITEM</td>
<td>PAGE</td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Summary</td>
<td></td>
</tr>
<tr>
<td>Hans-Juergen Woerle, MD</td>
<td>77</td>
</tr>
<tr>
<td>Clarifying Questions to Applicant</td>
<td>78</td>
</tr>
<tr>
<td><strong>FDA Presentations</strong></td>
<td></td>
</tr>
<tr>
<td>The EMPA-REG OUTCOME Study</td>
<td></td>
</tr>
<tr>
<td>Andreea Lungu, MD</td>
<td>115</td>
</tr>
<tr>
<td>Statistical Assessment</td>
<td></td>
</tr>
<tr>
<td>Jennifer Clark, PhD</td>
<td>135</td>
</tr>
<tr>
<td>Clarifying Questions to FDA</td>
<td>176</td>
</tr>
<tr>
<td><strong>Open Public Hearing</strong></td>
<td></td>
</tr>
<tr>
<td>Clarifying Questions (continued)</td>
<td>217</td>
</tr>
<tr>
<td>Questions to the Committee and Discussion</td>
<td>255</td>
</tr>
<tr>
<td>Adjournment</td>
<td>374</td>
</tr>
</tbody>
</table>
PROCEDINGS
(7:59 a.m.)

Call to Order

Introduction of Subcommittee

DR. SMITH: Good morning again. I would like to start out by reminding everyone to please silence your cell phones, smartphones, or any other devices if you have not already done so. And I would also like to identify the FDA press contact, Theresa Eisenman; if you are present, Theresa, in the back with her hand up.

My name is Robert Smith. I am the Chairperson of the Endocrinologic and Metabolic Drugs Advisory Committee, and I will be chairing this meeting. I will now call the Endocrinologic and Metabolic Drugs Advisory Committee Meeting to order. We'll start by going around the table and introduce ourselves. We will start with the FDA to my left and then proceed around the table.

DR. GUETTIER: My name is Jean-Marc Guettier. I'm the division director in the Division of Metabolism and Endocrinology Products
at the FDA.

DR. CHONG: William Chong. I'm the clinical team leader in the Division of Metabolism and Endocrine Products.

DR. LUNGU: Andreea Lungu, a clinical reviewer, Division of Metabolism and Endocrinology Products.

DR. STOCKBRIDGE: Good morning. I'm Norman Stockbridge from the Division of Cardiovascular and Renal Products at FDA.

DR. SCHAMBELAN: I'm Morrie Schambelan, professor of medicine, Division of Endocrinology, University of California San Francisco.

DR. BUDNITZ: Dan Budnitz with the medication safety program at the Centers for Disease Control and Prevention.

DR. COOKE: David Cooke. I'm an associate professor of pediatrics and the acting director of pediatric endocrinology at the Johns Hopkins University School of Medicine.

DR. NEATON: Good morning. Jim Neaton, professor of biostatistics, University of
Minnesota.

DR. CHO: Leslie Cho, director of women's cardiovascular center and section head of preventive cardiology at Cleveland Clinic.

DR. FRADKIN: Judy Fradkin, director of the Division of Diabetes, Endocrinology and Metabolic Diseases at the National Institute of Diabetes and Digestive and Kidney Diseases, NIH.

DR. ROSENBERG: Good morning. My name is Yves Rosenberg. I'm the chief of the atherothrombosis and coronary artery disease branch in the Division of Cardiovascular Sciences at the National Heart, Lung, and Blood Institute, the NIH.

DR. MCBRYDE: Good morning. My name is Kevin McBryde. I'm a pediatric nephrologist and a medical officer at the National Institute of Dental and Craniofacial Research at the NIH.

DR. EVERETT: Good morning. I'm Brendan Everett. I'm the director of the General Cardiology Inpatient Service at the Brigham and Women's Hospital and assistant professor of medicine at Harvard Medical School.
DR. BONNER: Good morning. LaToya Bonner, designated federal officer for this meeting.

DR. SMITH: I'm Robert Smith. I'm a diabetes and endocrinology clinician and researcher. I'm a professor in medicine, and also in the School of Public Health at Brown University.

DR. THOMAS: Abraham Thomas, chief of medicine, NYU Lutheran, and vice chair of medicine, NYU School of Medicine.

DR. KONSTAM: Marv Konstam. I direct the cardiovascular center at Tufts Medical Center in Boston.

DR. LI-NG: Melissa Li-Ng, consultant endocrinologist, Cleveland Clinic Abu Dhabi.

DR. GOOD: Good morning. I'm David Good, professor and chair of neurology at Penn State College of Medicine.

DR. DE LEMOS: James De Lemos, cardiologist and professor of medicine at UT Southwestern in Dallas.

MS. HALLARE: Good morning. Diana Hallare, consumer representative.
DR. PALEVSKY: I'm Paul Palevsky. I'm a nephrologist, professor of medicine at the University of Pittsburgh School of Medicine and chief of renal at the Pittsburgh VA.

DR. WILSON: Peter Wilson, professor of endocrinology and preventive cardiology and public health at Emory.

DR. HECKBERT: Susan Heckbert, professor in epidemiology, University of Washington, Seattle.

DR. YANOVSKI: Susan Yanovski. I co-direct the Office of Obesity Research at the National Institute of Diabetes and Digestive and Kidney Diseases.

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

DR. HIATT: William Hiatt, professor of medicine, University of Colorado, Division of Cardiology where I practice vascular medicine.

DR. PROSCHAN: I'm Michael Proschan. I'm a mathematical statistician at the National Institute of Allergy and Infectious Diseases.

DR. KEWALRAMANI: Reshma Kewalramani. I'm a
nephrologist and I head the U.S. medical organization at Amgen. I'm the industry representative.

DR. SMITH: Thank you. For topics such as those being discussed at today's meeting there are often a variety of opinions, some of which are quite strongly held. Our goal is that today's meeting will be a fair and open forum for discussion of these issues, and 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, FDA will refrain from
discussing the details of this meeting with the
media until its conclusion. Also, the committee is
reminded to please refrain from discussing the
meeting topic during breaks or lunch. Thank you.

Now, I'll pass the microphone to Commander
LaToya Bonner, who will read the conflict of
interest statement.

Conflict of Interest Statement

DR. BONNER: The Food and Drug
Administration is convening today's meeting of the
Endocrinologic and Metabolic Drug Advisory
Committee under the authority of the Federal
Advisory Committee 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
dependent committee's compliance with federal ethics and
collect 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 discussion of today's meeting, members and temporary voting members of the 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.

Today's agenda involves supplemental new drug application 204629 for empagliflozin tablets, and supplemental new drug application 206111 for empagliflozin and metformin hydrochloride tablets. Both sNDAs are sponsored by Boehringer Ingelheim Pharmaceuticals, Incorporated, for the proposed additional indication in adult patients with type 2 diabetes mellitus and high cardiovascular risk to reduce the risk of all-cause mortality by reducing the incidence of cardiovascular death and to reduce the risk of cardiovascular death or hospitalization for heart failure.

This is a particular matters meeting during which specific matters related to Boehringer Ingelheim's sNDAs will be discussed. Based on the agenda for today's meeting and all financial
interests reported by the committee members and temporary voting members, a conflict of interest waiver has been issued in accordance with 18 U.S.C. Section 208(b)(3) to Dr. Marvin Konstam.

Dr. Konstam's waiver addresses his employer's contract with a potentially competing firm regarding a product that potentially will compete with the products under review by the committee. The total funding is anticipated to be between zero and $50,000 per year. Dr. Konstam will not have any role in the actual conduct of the study.

The waiver allows this individual to participate fully in today's deliberations. FDA's reason for issuing the waiver are described in the waiver documents, which are posted on the FDA website at [www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/default.htm](http://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/default.htm).

Copies of the waiver may also be obtained by submitting a written request to the agency's Freedom of Information Division, on 5630 Fishers Lane, Room 1035, Rockville, Maryland 20857, or...
requests may be sent via fax to 301-827-9267.

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

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

We would like to remind members and temporary voting members that if the discussion involve any other products or firms not already on the agenda for which an FDA participant has a personal or imputed financial interest, the participants need to exclude themselves from such involvement and their exclusion will be noted for the record. FDA encourages all other participants to advise the committee of any financial
relationships that they may have with the firm at issue. Thank you.

DR. SMITH: Thank you. So, we'll now proceed with the FDA's introductory remarks from Dr. Jean-Marc Guettier.

**FDA Introductory Remarks**

DR. GUETTIER: Is it possible to project the slides on the screen? Good morning. My name is Jean-Marc Guettier, and I'm the director of the Division of Metabolism and Endocrinology Products at the Food and Drug Administration. I'd like to welcome all of the meeting participants and the panel members to today's advisory committee meeting, which was convened to review the results of the EMPA-REG OUTCOME study.

The EMPA-REG OUTCOME study was required by the Food and Drug Administration to evaluate the cardiovascular risk associated with the use of empagliflozin for the treatment of adults with type 2 diabetes. Recall that all anti-diabetic therapies are indicated as an adjunct to diet and exercise, and are used to improve glycemic control.
The regulatory context demonstrating an improvement in glycemic control over the medium term serves as a surrogate for microvascular disease risk reduction and for full approval of these agents. To date, no prospectively planned study has convincingly demonstrated a link between improvement in glycemic control and reduction in cardiovascular risk.

The EMPA-REG OUTCOME study is the first large, dedicated, randomized, prospective, controlled outcomes trial to report on a cardiovascular benefit of a specific anti-diabetic therapy.

Several other anti-diabetic agents have been evaluated in similarly designed studies, and until the publication of the EMPA-REG OUTCOME study, no anti-diabetic agent had been shown to reduce cardiovascular risk.

The results of the EMPA-REG OUTCOME study are thus a departure from previous scientific findings, and represent a potentially important advance in the treatment of adults with type 2
diabetes.

The committee was convened to discuss whether the results of the study established that empagliflozin is effective at reducing cardiovascular risk in adult patients with type 2 diabetes, and in so doing advise the agency with regards to adding a potentially new use to the product label.

The indication submitted when the supplemental application was filed in November is the first indication shown on this slide. This was the working indication FDA was asked to consider, and the indication published in the Federal Register notice for this meeting.

The indication reads, "Empagliflozin is indicated in adult patients with type 2 diabetes mellitus and high cardiovascular risk to reduce the risk of all-cause mortality by reducing the incidence of cardiovascular death, and to reduce the risk of cardiovascular death or hospitalization for heart failure."

In early May, the applicant submitted an
amendment to the application to have the proposed indication changed to the following.

"In adult patients with type 2 diabetes mellitus and established cardiovascular disease, empagliflozin is indicated to reduce the incidence of cardiovascular death." This new indication is the current working version and reflects the new use the applicant is seeking.

Since you will be asked to opine on whether the study results establish the efficacy of empagliflozin for a new use, let me briefly review FDA requirements regarding the evidence needed to form the basis for a new efficacy claim. The legal standard of effectiveness to support a new use of an approved product is described in the Federal Food, Drug, and Cosmetic Act as substantial evidence of effectiveness.

In the law, substantial evidence is defined as evidence consisting of adequate and well-controlled investigations, including clinical investigations, that the drug has the effect it purports or is represented to have under the
condition of use prescribed in the labeling or
proposed labeling thereof. The adequate and
well-controlled statement speaks to the quality of
the evidence required, and this will be reviewed in
a following slide.

With regards to the quantity of evidence
necessary, the law talks about investigations in
the plural form, and it has been FDA's position
that Congress generally intended to require at
least two adequate and well-controlled studies,
each convincing on its own, to establish
effectiveness and support approval of a new use.

The scientific basis for requiring two
studies is grounded in the fact that chance,
bias, biologic variability, generalizability
issues for other factors inherent to the single
positive trial, could lead to an erroneous
conclusion that the drug is effective when in fact
it is not.

Simply stated, a conclusion based on two
persuasive studies will always be more secure than
a conclusion based on a single comparatively
persuasive study. We will return to this concept in a later slide, but let me first briefly cover issues related to the quality of the evidence required.

With regards to the quality of the evidence, the applicant has to demonstrate that the studies were adequately designed and conducted. General attributes of an adequate and well-controlled trial are described in Title 21 of the Code of Federal Regulations and are summarized in this slide.

To demonstrate that the trial supporting an effectiveness claim is adequate and well-controlled, extensive documentation related to trial planning, trial conduct and data handling is needed. Records of this extensive documentation are submitted to the agency for review, and detailed patient records at the clinical sites participating in the study are made available for the purpose of auditing.

This morning Dr. Lungu will spend some time reviewing the complex regulatory history for the EMPA-REG study, which is in many ways unique, and
will summarize the major changes to the protocol, clinical endpoint charter, and analysis plan made during trial conduct. You will be asked to consider these changes in your discussion this afternoon when weighing the quality of the evidence generated by the EMPA-REG study.

Let me return to the issues surrounding the quantity of evidence necessary to support a new claim. Although, as stated earlier, two adequate and well-controlled studies are generally needed to form the basis of an effectiveness claim, the FDA has recognized that there are certain circumstances when a single study could constitute substantial evidence.

In 1997 the Federal Food, Drug, and Cosmetic Act was amended to make it clear that the agency may consider data from one adequate and well-controlled clinical investigation and confirmatory evidence to constitute substantial evidence if FDA determines that such data and confirmatory evidence are sufficient to establish effectiveness.
Indeed, FDA has at times relied on a single adequate and well-controlled efficacy study to support approval of a new drug or a new use. This has generally occurred only in cases in which a single multi-center study of excellent design provided highly reliable and statistically strong evidence of an important clinical benefit.

Whether to rely on a single adequate and well-controlled study to form the basis of a claim is inevitably a matter of judgment. And as stated previously, a conclusion based on two persuasive studies will always be more secure than a conclusion based on a single similarly persuasive study.

The agency's views on when a single trial could serve as the basis to conclude that substantial evidence has been met is described in a guidance document. In this document, the agency identifies some characteristics of a single adequate and well-controlled study that can contribute to a conclusion that the single study would be adequate to form the basis of a new
efficacy claim. These characteristics are shown on this slide.

Although the list is not exhaustive and no one single characteristic is necessarily determinative, the presence of one or more of these in a study may provide support to reach a conclusion that the study is adequate.

In your deliberations this afternoon, we will ask you to keep these characteristics in mind as you are weighing whether the single EMPA-REG study provides the necessary substantial evidence of effectiveness to support a new claim.

Having briefly reviewed the expectations regarding both the quality and the quantity of evidence needed to form the basis of a new claim, let me present the charge to the committee.

The committee was convened today to review the evidence generated in the EMPA-REG study and advise the agency on whether the study provides substantial evidence of effectiveness required to form the basis of a new claim for empagliflozin.
tackle this objective, so let me turn to the discussion points.

The first discussion point focuses on the quality of the evidence. The trial has had a complex regulatory history and several changes were implemented during trial conduct. This will be discussed in details this morning.

In addition, some sponsor unblinding occurred to support an interim analysis required for the purpose of regulatory filing, and this is unique to this trial. In this question, we ask whether specifics of the study conduct alter or do not alter your level of confidence in the main study findings.

The second discussion point focuses on the persuasiveness of the results for the primary analysis, and asks you to comment on several issues, including the persuasiveness of the statistical results, the findings across each individual component to the primary composite endpoint, and specifically on the topic of silent myocardial infarctions.
These issues will be covered in our presentation and we would like your opinion on how these factors influence your confidence in the results of the primary analysis.

The third discussion point focuses specifically on the persuasiveness of the mortality findings. While the discussion of mortality is certainly not limited to the issues listed here, we will be covering these issues in our presentation, and would like you to address these specifically in your discussion.

Discussion point number 4 relates to the heart failure finding in the EMPA-REG OUTCOME study. And discussion point number 5 relates to the renal endpoint findings in the EMPA-REG OUTCOME study. Specifically, it asks whether the changes in the endpoint establish a benefit of the drug on kidney disease related to diabetes.

The first voting question asks whether the EMPA-REG OUTCOME study results have fulfilled the recommendations laid out in the 2008 guidance for industry. Almost as important as your vote is the
rationale for your vote. We would like you to explain why you voted the way you did on this question.

The second and final voting question asks whether the results of the EMPA-REG OUTCOME study provides the substantial evidence of effectiveness necessary to establish that empagliflozin reduces cardiovascular mortality in the population studied. Again, we would like you to explain the rationale for your vote.

This concludes my introduction. Thank you and I look forward to a productive meeting.

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

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

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

We will now proceed with Boehringer Ingelheim Pharmaceuticals' presentations.

Applicant Presentation – Hans-Juergen Woerle

DR. WOERLE: Good morning. My name is Hans-Juergen Woerle. I'm an endocrinologist and vice president of therapeutic area metabolism at Boehringer Ingelheim. The EMPA-REG OUTCOME trial was conducted as a postmarketing commitment to the FDA following the approval of empagliflozin for glucose lowering.
As a sponsor, we would like to thank the FDA for this opportunity to present the data from this trial. Our presentation will follow this outline; study design, effects on MACE, effects on cardiovascular death, effects on heart failure outcome, and all-cause mortality, safety profile, data integrity, and finally the clinical relevance of the trial results.

Professor Zinman, the chair of the EMPA-REG OUTCOME trial steering committee, will now discuss the context and background.

Applicant Presentation – Bernard Zinman

DR. ZINMAN: Thank you, Dr. Woerle. My duality of interests have been submitted. And just to repeat, I do receive consulting and grant support from Boehringer Ingelheim.

Good morning. As shown on this slide, the prevalence of diabetes has been steadily increasing in the United States. Of note, in the age groups studied in the EMPA-REG OUTCOME trial, the prevalence of diabetes has actually doubled. This comes at a considerable human and healthcare cost.
Compared with those without diabetes, patients with diabetes have a 70 percent higher risk of cardiovascular mortality.

Having a diagnosis of diabetes decreases life expectancy by six years, driven largely by premature cardiovascular death. A prior history of myocardial infarction or stroke doubles this loss of life to 12 years. Thus, developing strategies to modify this unacceptable outcome remains an important goal in the management of diabetes.

A frequently less-recognized cardiac comorbidity of type 2 diabetes is heart failure. Heart failure is highly prevalent in patients with diabetes, occurring in more than one in five patients above the age of 65.

Patients with diabetes and heart failure have a very poor prognosis, with a median survival of less than five years. It is noteworthy that no glucose-lowering medication has been shown to improve heart failure outcomes, while some, namely thiazolidinediones, are associated with an increased heart failure risk.
It is now well accepted, as you heard, that hemoglobin A1c lowering reduces microvascular complications. A 1 percent decrease in hemoglobin A1c is associated with a 37 percent reduction in microvascular complications. However, there is limited evidence that reduction in hemoglobin A1c results in cardiovascular outcome benefits.

Previous randomized trials comparing more intensive with less intensive glucose lowering showed no general reduction in cardiovascular events or mortality in type 2 diabetes. In fact, one of the trials, the ACCORD trial, showed an increase in mortality with intensive glycemic therapy.

Data from these four trials have been pooled together in a meta-analysis. This forest plot demonstrates the absence of a general benefit of glycemic control on cardiovascular outcomes. Stroke, shown in the first row, was not impacted. There was a modest reduction in non-fatal myocardial infarction. However, the hazard ratio for heart failure outcomes was one, and no
reduction was observed in all-cause mortality or cardiovascular death.

The uncertainty around the cardiovascular risks and benefits of glucose-lowering therapies led to the development of the 2008 cardiovascular guidance. This FDA guidance set the scene for recent cardiovascular outcome trials, and the strategy of investigating drug-specific effects through placebo-controlled trials.

In these trials, study drug is added to standard of care, aiming for glycemic equipoise in patients at high cardiovascular risk, with at least two years of follow-up and with prospective independent and blinded cardiovascular adjudication.

Non-inferiority can be concluded if the upper bound of the confidence interval is less than 1.3. Superiority can be concluded if the upper bound is less than 1.0. In the rest of the presentation today, you will see results from our trial that follows this guidance.

Empagliflozin has been hypothesized to
reduce cardiovascular risk through multiple mechanisms. The primary target of empagliflozin is the kidney where it increases sodium and glucose excretion resulting in an osmotic diuresis. This leads to a reduction in hemoglobin A1c and weight. The hemodynamic effects of decreasing intravascular volume results in decreased systolic blood pressure, and an increase in hematocrit.

Importantly, these hemodynamic effects are not associated with a corresponding increase in sympathetic tone. In turn, these effects could potentially decrease cardiac wall stress, minimize the risk of new or worsening heart failure, and reduce the risk of cardiovascular death.

In this context, results from the EMPA-REG OUTCOME trial represents a new paradigm in the treatment of diabetes. For the first time, physicians and patients have a drug that not only controls metabolic parameters, but also decreases the risk of cardiovascular mortality.

Dr. Woerle will now take you through the design of our trial.
DR. WOERLE: Empagliflozin is a highly selective inhibitor of sodium glucose co-transporter 2 in the kidney. In the United States, it’s indicated as an adjunct to diet and exercise to improve glycemic control in adults with type 2 diabetes mellitus. Two doses are approved, 10 and 25 milligram per day.

In our clinical development program for empagliflozin we studied more than 13,000 patients with type 2 diabetes. In these trials, we saw reductions in HbA1c, weight and blood pressure. The EMPA-REG OUTCOME trial was designed to meet the postmarketing requirement to conduct a randomized, double-blind, placebo-controlled trial, evaluating the effect on major cardiovascular events. Key results of this trial are summarized in the next slide.

Empagliflozin demonstrated cardiovascular safety with the upper bound of the confidence interval being below 1.3, thus fulfilling FDA postmarketing requirements. We had a statistically
significant reduction in the primary endpoint.

MACE, the composite of cardiovascular death, non-fatal MI, and non-fatal stroke was reduced by 14 percent. There was a 38 percent risk reduction in cardiovascular death. This cardiovascular mortality benefit drove the reduction in MACE with no significant changes in the risk of myocardial infarction or stroke.

The reduction in cardiovascular death was meaningful enough to be translated into a reduction, even in all-cause mortality, by 32 percent. Heart failure hospitalization was reduced by 35 percent. And the composite of heart failure hospitalization or CV death was reduced by 34 percent.

In addition, the established safety profile of empagliflozin was confirmed. Our fulfillment of the prespecified statistical plan of the EMPA-REG OUTCOME trial not only allowed us to demonstrate cardiovascular safety, but also allowed us to propose the following indication for empagliflozin. In adult patients with type 2 diabetes mellitus and
established cardiovascular disease, empagliflozin is indicated to reduce the incidence of cardiovascular death since this was the component that drove the success of the primary endpoint.

Let us now take a look at the design of the EMPA-REG OUTCOME trial, followed by results in more detail. The EMPA-REG OUTCOME trial examined the long-term effects of empagliflozin on cardiovascular morbidity and mortality, as well as general safety in patients with type 2 diabetes and established cardiovascular disease. The study was run by a steering committee, including representatives from academia and the sponsor.

Our blinded adjudication panel and the data monitoring committee consisted of independent medical and statistical experts. In addition, CV outcome analysis were independently validated by statisticians at the University of Freiburg. Membership of the independent study panels are included on page 49 of our briefing book.

EMPA-REG OUTCOME is a randomized, double-blind, placebo-controlled trial.
Empagliflozin, 10 and 25 milligram, once daily were tested. There were three arms, and patients were randomized on a 1 to 1 to 1 basis. This was an event-driven trial targeting at least 691 events, patients with an adjudicated primary outcome event. Study treatment was added to local standards of care for patients with type 2 diabetes and established cardiovascular disease.

Prespecified and adjudicated endpoints in this trial included the following: cardiovascular death, myocardial infarction, stroke, hospitalization for unstable angina, coronary revascularization, transient ischemic attacks, hospitalization for heart failure, hospitalization for heart failure or CV death, and all-cause mortality.

Of these, 3-point MACE, which is the composite of CV death, nonfatal MI and nonfatal stroke was our primary outcome. 4-point MACE, which includes hospitalization for unstable angina in addition to the components of 3-point MACE, was a key secondary outcome. The prespecified and FDA
approved statistical analysis plan was to compare pooled empagliflozin groups with placebo. Type 1 error was controlled by hierarchical testing.

First, we tested for non-inferiority for 3-point MACE. If statistical significance was achieved, non-inferiority for 4-point MACE was tested. If non-inferiority was established, superiority for 3-point MACE was tested, and if successful, superiority for 4-point MACE was tested. These confirmatory endpoints were tested on an alpha level of 0.0249, allocating a 0.001 alpha for the interim data extraction.

This interim analysis was done to support the original new drug application of empagliflozin to the FDA in 2013. The 1-sided alpha level of 0.0249 equates to a 2-sided alpha of 0.0498. Throughout the presentation, we will be reporting 2-sided alpha, consistent with our publications.

Other endpoints were tested at a nominal alpha level of 0.05. Non-inferiority was to be concluded if the 2-sided upper bound of 95.02 percent confidence interval was lower than
1.3.

Superiority was to be concluded if the upper bound of the confidence interval was lower than 1.0. Analysis of cardiovascular outcome were based on Cox proportional hazard models. This prespecified model includes treatment, age, gender, baseline categories of BMI, HbA1c, eGFR, and geographical region.

Patients who did not have an event were censored on the last day they were known to be free outcome. For cumulative incidence functions for all cardiovascular outcome events, except for all-cause mortality, we censored patients at risk by correcting for the competing risk of mortality.

The primary analysis was conducted in patients treated with at least one dose of study drug. We referred to this population as intent-to-treat population hereafter. Additionally, on-treatment and per protocol analysis were conducted. The three treatment groups are shown in columns.

Discontinuation of study medication was 29
percent in the placebo group. These rates were lower in both empagliflozin groups, 24 and 23 percent. Ninety-seven percent of patients completed the study. These are the patients who had an event contributing to the primary outcome, or were known to be event free at the end of the trial. Final vital status was available for more than 99 percent of patients. The median observation time was 3.1 years.

EMPA-REG OUTCOME trial was a global effort, conducted at 590 sites, from 42 countries. Countries were grouped into region as shown on this map. The next few slides show baseline characteristics of these patients.

Study groups were balanced. Mean age was 63. Approximately one-third were women. Patients from the United States, Canada, Australia, and New Zealand were pooled together as a prespecified region. Twenty percent of the total population was from this pooled region, and 90 percent of these were recruited from the United States.

Approximately 72 percent of patients were
white; 350 patients, 5 percent of the overall population, identified themselves as black or African-American. Around a fifth were Asian; 18 percent identified themselves as Hispanic or Latino.

Ninety-nine percent of the population had established cardiovascular disease. Approximately three-quarters had coronary artery disease and almost half had a history of myocardial infarction. About 23 percent had a history of stroke; 10 percent had investigator-reported history of heart failure at baseline.

Baseline HbA1c was around 8 percent. The mean body mass index was just over 30. Patients were well-controlled for blood pressure and cholesterol. Patients with an eGFR less than 30 at screening were excluded per protocol.

Consistent with clinical practice, metformin was the most commonly used oral glucose-lowering therapy. Approximately half the patients were treated with insulin at baseline. Ninety-five percent of the population was on anti-hypertensive
therapy with ACE, ARBs the most common, used in about 80 percent. Beta blockers and diuretics were used by approximately 65 and 43 percent of the patient population.

Within the class of diuretics, the use of loop diuretics was also balanced. Around 80 percent of patients were on lipid-lowering drugs at baseline. And around 90 percent of patients were treated with anti-coagulants or anti-platelet therapies. In this population, with established cardiovascular disease, therapies to address CV risk were widely used.

We also largely succeeded in keeping patients in the trial and on study medication. Discontinuation rates for empagliflozin was lower than that of placebo. We ensured optimal ascertainment of outcomes by following up patients irrespective of drug discontinuation. This is evidenced by the greater than 99 percent obtainment of vital status. Furthermore, we implemented a process to ensure 100 percent source document verification.
Let us now look at changes in cardio metabolic parameters over the course of the study. In these analyses, all patients, including those who discontinued study drug or initiated new therapies, were included. This followed the intent-to-treat principle.

During the first 12 weeks, glucose-lowering therapy was meant to be kept stable. Thereafter, investigators were encouraged to treat patients according to local guidelines for HbA1c targets, adding additional glucose-lowering therapy if necessary.

HbA1c reduction during the first 12 weeks was consistent with the reduction previously observed in the empagliflozin clinical development program. While glycemic equipoise was the objective in the trial, a modest difference between active drug and placebo persisted throughout, as we have seen with other cardiovascular outcome trials in type 2 diabetes.

Some weight loss was observed early, and these findings were largely persistent throughout
the trial. Both systolic and diastolic blood pressure showed reductions with empagliflozin. Heart rate remained largely unchanged. Small increases were observed in LDL cholesterol with empagliflozin. HDL cholesterol also showed small increases.

Dr. Uli Broedl, head of clinical development metabolism at BI, will now take you through the cardiovascular outcome data.

Applicant Presentation – Uli Broedl

DR. BROEDL: Good morning. My presentation will cover outcome related to MACE, CV death, heart failure, and all-cause mortality. The following figure shows the cumulative incidence of 3-point MACE. The red line represents patients on empagliflozin and the grey line placebo.

The vertical axis displays the percentage of patients with events. The horizontal axis shows the number of patients at risk at a given point in time. Incidence rates per 100 patient years are in parentheses. Next to these we have provided the number of events accrued, as well as frequencies of
events. We will use this scheme throughout our presentation.

This slide also shows the hazard ratio with 95.02 percent confidence interval and p-value. Empagliflozin reduced the primary endpoint by 14 percent. The hazard ratio was 0.86. The upper bound of the confidence interval was 0.99 with a p-value of 0.0382. This figure displays patients on the two doses of empagliflozin separately, 10 milligrams in purple and 25 milligrams in blue. Hazard ratios were virtually identical, 0.85 and 0.86.

We also looked at additional predefined analyses populations. In the intent-to-treat population, the hazard ratio is 0.86. In the on-treatment population, defined as patients with events occurring up to 30 days after last intake of trial medication, the hazard ratio is 0.87. And in the per-protocol population, the hazard ratio is 0.86.

We have drawn a vertical dashed red line to indicate the hazard ratio in the intent-to-treat
population. The hazard ratios were virtually identical in the three analyses populations. We will follow this scheme throughout today's presentation.

We analyzed prespecified subgroups according to baseline characteristics, including age, sex, race, ethnicity, HbA1c, BMI, and eGFR. This forest plot shows the overall hazard ratio with 95 percent confidence intervals for the intent-to-treat population in the first row. The subgroups are listed in rows below.

The magnitude of the treatment effect was generally consistent across predefined subgroups, although, for this composite endpoint, some heterogeneity was observed. But please note that when subgroups have a small number of events, they are subject to play of chance.

We also tested a number of subgroups based on concomitant therapies at baseline. We observed no meaningful heterogeneity in subgroups defined by the use of glucose-lowering therapies, such as metformin and insulin, lipid lowering agents such
as statins or ezetimibe, and anti-hypertensive therapies, including RAS blockers, beta blockers, and diuretics.

Let us now look at the individual components of 3-point MACE; cardiovascular death, non-fatal myocardial infarction, and non-fatal stroke. The effect on the primary endpoint was driven by a 38 percent reduction in the risk of cardiovascular death, with little effect on the atherosclerotic endpoints of non-myocardial infarction and non-fatal stroke.

Looking beyond 3-point MACE to 4-point MACE, which includes hospitalization for unstable angina, non-inferiority was also achieved. Similarly, the hazard ratio for coronary re-vascularization was 0.86, with the upper bound of the confidence interval crossing unity. Of note, the point estimate for non-fatal stroke was numerically larger than 1.

So we will now present more detailed analyses. The next slide shows the cumulative incidence of all stroke, including fatal and
non-fatal stroke in the intent-to-treat population. The analysis includes all events irrespective of whether patients were on-treatment or not.

Here, the hazard ratio was 1.18 with confidence intervals crossing unity. The curves overlap during periods of the trial where we followed most of the patients and observed most of the events.

We repeated this analysis for all strokes that occurred up to 90 days after drug discontinuation. In this analysis, the hazard ratio comes closer to unity. We further analyzed recurrent stroke, fatal stroke, or transient ischemic attack. Recurrent strokes, although infrequent, were balanced between placebo and the two empagliflozin groups. While the hazard ratio for non-fatal stroke was numerically greater than 1, the hazard ratio for fatal stroke was 0.72.

Transient ischemic attacks share the same pathophysiology as stroke. While the hazard ratio for stroke was numerically greater than 1, we see the opposite picture for TIA. The hazard ratio for
this adjudicated outcome is 0.85. And if we look at all stroke-like events, including fatal stroke, non-fatal stroke, and TIA, the hazard ratio is 1.05.

We also looked at changes in hematocrit or blood pressure and stroke. No associations were observed between stroke and changes in hematocrit or blood pressure. No associations were observed between the occurrence of volume depletion and stroke either. Data supporting these conclusions are provided in our briefing book.

So far, we have looked at components of the primary endpoint, which were not significant. Now we will look more closely at cardiovascular death, which was the principle driver of the primary outcome.

Overall, 309 cases of cardiovascular death occurred. Empagliflozin showed a 38 percent reduction in the risk of cardiovascular death. The hazard ratio was 0.62. The 95 percent confidence interval was 0.49 to 0.77, with a highly significant nominal p-value. These effects are
evident early in the trial and were sustained until the end. Both doses of empagliflozin yielded comparable results for cardiovascular death. Hazard ratios were 0.65 for the 10-milligram dose, and 0.59 for the 25-milligram dose.

We also compared intent-to-treat and on-treatment populations, as we did for stroke. The two hazard ratios in the intent-to-treat and on-treatment populations were comparable at 0.62 and 0.59.

This figure shows subgroups that we have previously looked at. There is no meaningful heterogeneity across these subgroups. And the picture is similar when we look at concomitant therapies used at baseline. We observed no meaningful heterogeneity in the effect of empagliflozin on cardiovascular death.

What are the modes of cardiovascular death that contributed? All deaths were adjudicated by a central, blinded, independent adjudication committee based on prespecified criteria that are consistent with guidelines. The committee
categorized deaths into cardiovascular death and non-cardiovascular death.

Cardiovascular deaths were further classified into fatal MI, fatal stroke, heart failure death, and sudden death. Cardiovascular deaths with full documentation, but due to causes other than those four categories, were classified under other CV cause. These other causes included dysrhythmia, pulmonary embolism, or coronary intervention.

In addition, there were deaths where the amount of information available was not sufficient for the adjudication committee to definitely assign into one of these categories. As prespecified, in this population with established cardiovascular disease, all deaths not attributed to cardiovascular death or non-cardiovascular death were presumed cardiovascular death.

The next figure shows a hazard ratio and 95 percent confidence interval for the various types of cardiovascular death. Empagliflozin reduced the risk of cardiovascular death by
38 percent. The hazard ratio is 0.62. Essentially all categories of CV death, fatal MI, fatal stroke, heart failure death, sudden death, and presumed cardiovascular death showed a favorable point estimate.

The most frequent causes of cardiovascular death were heart failure death, sudden death, and presumed cardiovascular death. These modes of death are commonly observed in patients with heart failure.

To better understand the category of presumed cardiovascular death, we looked at the investigator-reported adverse events leading to death in these patients. As one would expect in this population with established cardiovascular disease, these investigator reports were largely consistent with cardiovascular death. Importantly, even if one excluded presumed cardiovascular death in a sensitivity analysis, reduction in cardiovascular death remained highly significant.

Empagliflozin reduced cardiovascular death, excluding cases of presumed cardiovascular death,
by 41 percent. The hazard ratio was 0.59 with an upper bound of the 95 percent confidence interval of 0.79 and a highly significant nominal p-value.

To conclude this section on MACE and cardiovascular death, a 38 percent risk reduction in cardiovascular death drove the reduction in MACE with no significant change in the risk of myocardial infarction or stroke. There was some heterogeneity for subgroups of 3-point MACE. The findings of cardiovascular death however, were consistent across the two doses, analyses populations, and subgroups.

Modes of cardiovascular death associated with heart failure were most frequent. Therefore, let us now look at heart failure outcomes in our trial. Heart failure hospitalization was a prespecified and adjudicated endpoint in this trial. A composite endpoint of heart failure hospitalization or cardiovascular death was also prespecified. This is an endpoint that is commonly used in heart failure trials.

Empagliflozin reduced the risk of
hospitalization for heart failure by 35 percent. The hazard ratio is 0.65. The 95 percent confidence interval was 0.5 to 0.85, with a highly significant nominal p-value. The effect was evident early and persisted throughout the trial.

This figure shows heart failure hospitalization within the two doses of empagliflozin. Hazard ratios for the 10- and 25-milligram doses are 0.62 and 0.68 respectively. Results for both doses were comparable. We also compared the consistency across adjudicated hospitalization for heart failure, investigator-reported heart failure, and initiation of loop diuretics, a commonly used surrogate for heart failure.

For adjudicated heart failure hospitalization, the hazard ratio is 0.65. For investigator-reported heart failure, 0.70. And for the initiation of loop diuretics, it is 0.62. This consistency provides internal validity for the hospitalization for heart-failure endpoint.

Empagliflozin also reduced the composite of
hospitalization for heart failure or cardiovascular death. Virtually identical results were seen for this composite endpoint. The hazard ratio was 0.66. The 95 percent confidence interval was 0.55 to 0.79, with a highly significant nominal p-value. Ten percent of the EMPA-REG OUTCOME population had investigator-reported heart failure at baseline. For patients with or without history of heart failure, hazard ratios are comparable.

A similar result is seen for cardiovascular death, suggesting that empagliflozin may reduce heart failure hospitalization and CV death in patients with and without preexisting heart failure. Empagliflozin reduced cardiovascular death largely by reducing the risk of deaths that are common in patients with heart failure.

The drug also prevented heart failure hospitalization in patients with established atherosclerotic cardiovascular disease. The composite endpoint of heart failure hospitalization and CV death was reduced by 34 percent. Results were consistent for both doses and across
When a drug reduces cardiovascular death, it is important that this benefit is not counterbalanced by an increase in non-cardiovascular death. In EMPA-REG OUTCOME trial, all-cause mortality was therefore included as a predefined endpoint. In this trial, vital status was available for more than 99 percent of patients, making the ascertainment of this important endpoint methodologically robust.

As shown in this Kaplan-Meier plot, empagliflozin reduced all-cause mortality by 32 percent. The hazard ratio was 0.68. The confidence interval was 0.57 to 0.82, with a highly significant nominal p-value.

Virtually identical results were observed for empagliflozin 10 and 25 milligrams. Hazard ratios were 0.70 and 0.67 respectively. We saw similar results when we looked at intent-to-treat and on-treatment populations. The hazard ratios are 0.68 and 0.69 respectively.

We also analyzed the contribution of
cardiovascular death and non-cardiovascular death to the reduction in all-cause mortality. For cardiovascular death, the hazard ratio is 0.62. For non-cardiovascular death, the hazard ratio was 0.84.

The reduction in cardiovascular death with empagliflozin was enough to be translated into a reduction in all-cause mortality by 32 percent. Results were based on a large number of events, 463 cases of all-cause mortality, consistent across doses, across analysis populations, and across subgroups.

We have seen the benefits of empagliflozin on cardiovascular outcomes. Now I will hand over to Dr. Woerle to discuss safety results.

**Applicant Presentation - Hans-Juergen Woerle**

DR. WOERLE: The next few slides will depict the number of adverse events and their frequency. We will show data for placebo, empagliflozin, 10- and 25-milligram groups. Both doses of empagliflozin were well tolerated. The safety profile was consistent with previous trials with
empagliflozin.

Adverse events were comparable across three groups. Occurrence of any adverse events leading to discontinuation and serious adverse events were all balanced with numerically fewer patients having these adverse events on empagliflozin. Drug-related adverse events occurred more frequently on empagliflozin.

In addition to general safety, we examined safety topics of particular relevance to empagliflozin or the class of SLGT-2 inhibitors. As previously observed in our registration trials, genital infections were increased in patients treated with empagliflozin. This was the case both in men and women.

Volume depletion and venothrombotic events were balanced. Urinary tract infections were balanced between placebo and empagliflozin for male and female patients. Urosepsis was infrequent, but numerically higher in the empagliflozin groups. On the other hand, pyelonephritis was comparable with numerically fewer events in patients treated with
Diabetic ketoacidosis is included as a labeled adverse event for all SLGT-2 inhibitors. In the EMPA-REG OUTCOME trial, diabetic ketoacidosis cases were infrequent, 1, 3, and 1 case in placebo, empagliflozin 10-, and 25-milligram. There was no imbalance in bone fracture. Cases of cancer were balanced. Hepatic injuries was also balanced with numerically fewer events on empagliflozin.

Similarly, the risk of decreased renal function was not increased. Further reassurance can be drawn from analysis of the time to first occurrence of decreased renal function. In this analysis, the risk of decreased renal function is significantly lower on empagliflozin in comparison to placebo.

Furthermore, we looked at eGFR over time. Patients on empagliflozin showed an initial dip in eGFR, consistent with our registration trials, that lasted up to 52 weeks. These previous observations led to a concern that empagliflozin may cause a
decline in eGFR, which is reflected in our current label. However, this picture appears to be different over the entire course of the study.

eGFR in the placebo group showed the expected decline in renal function of approximately 2 milliliters per year. In contrast, eGFR in patients on empagliflozin remained largely unchanged. These data suggest that the initial worsening of renal function is transient, and over time empagliflozin appears to slow the deterioration of renal function. These data are reminiscent of the effects we see with the initiation of ACE inhibitors.

The figure shows baseline characteristics at last eGFR value on-treatment. Patients were enrolled thereafter for a median of 34 days. In the placebo group, the eGFR remained largely unchanged during this follow-up period. However, in both the empagliflozin groups, eGFR almost returned to baseline.

In addition, our trial included the prespecified renal endpoint, new or worsening of
nephropathy. The hazard ratio for this
prespecified composite outcome was 0.61.
Confidence interval was 0.53 to 0.70, with a
p-value of less than 0.0001. New onset
macroalbuminuria, doubling of serum creatinine, and
initiation of renal replacement, all clinically
relevant components, contributed to the results
with hazard ratios of 0.62, 0.56, and 0.45
respectively; 3 renal deaths occurred on
empagliflozin, 10-milligram, none on placebo or
empagliflozin, 25-milligram.

In this analysis, only the time to first
event was considered. Nevertheless, we assessed
the same endpoints for sustainable changes, and the
findings are largely consistent.

To conclude this section on safety, no new
safety signals were identified in the EMPA-REG
OUTCOME trial. The rates of volume depletion, bone
fractures, venothrombotic events, hepatic events,
hypoglycemia, cancer, and DKA were comparable
between empagliflozin and placebo. General
infections were increased with empagliflozin, as
known from our previous trials.

The data that we have just seen suggests that empagliflozin has no adverse impact on renal outcome. Empagliflozin appears to slow the progressive decline in renal function in peril to its favorable effects on cardiovascular death and heart failure hospitalization.

Now, interpretation of the safety and efficacy need to be made in the context of quality and data integrity. It is important to note that the EMPA-REG OUTCOME trial was designed and executed according to the contemporary clinical trial standards and are now analyzed according to prespecified statistical analysis plan.

Our plans for a CV outcome trial for empagliflozin were driven by the 2008 FDA cardiovascular guidance. The initial objective of this trial was to demonstrate cardiovascular safety to fulfill the guidance by meeting non-inferiority margin at the time of the NDA submission in 2013. In 2011, based on emergent science and regulatory feedback, we decided to expand the trial. This
expansion to 7,000 patients aimed to accrue at least 691 events, providing adequate power to test not only for non-inferiority but also for superiority.

In this trial, we prespecified an interim data extraction to support empagliflozin's original new drug application. It is important to note that we performed the interim analysis after major amendments to the protocol were made. The only amendment that followed the interim focused primarily on adjudication of hepatic and cancer events.

An independent firewalled team performed the interim analysis and regulatory submission, and all trial personnel, investigators, and our participants, were kept blinded throughout the trial. Neither the sponsor nor any regulatory authority published EMPA-REG OUTCOME interim results. No major protocol changes were implemented after the interim analysis.

Finally, we performed sensitivity analysis to assess any potential impact of this interim
analysis. When you look at trial population recruited before and after the interim analysis, results are comparable and consistent with the overall trial population. The hazard ratios for cardiovascular death were also virtually identical. Similarly, there were no differences in all-cause mortality finding between the populations of patients recruited before and after the interim analysis.

Central adjudication of endpoints is another key principle of trial quality. Adjudication criteria for endpoints were predefined indicia and are shared across our CV outcome trials in diabetes. Endpoint definitions were reviewed regularly by adjudicators for changes in clinical practice and guidelines.

Whenever changes were made to endpoint definitions, we re-adjudicated all events to ensure consistency. Importantly, the definition of cardiovascular death, including presumed CV death, remained unchanged throughout the trial in our adjudication charter.
Changes to the endpoint definitions before or after the interim analysis did not affect the outcome. Only events adjudicated as outcome events were included in the MACE endpoint. Silent MIs occur in patients with type 2 diabetes, but its ascertainment is challenging. In previous cardiovascular outcome trials in diabetes, the occurrence of silent MI has been included in some trials but not in others.

When a trial prospectively includes silent MI in the primary endpoint, a formal process of confirmation is put in place. Now, in the EMPA-REG OUTCOME trial, we decided not to include silent MI in the primary endpoint. Therefore, we did not implement procedures to confirm such events.

Let us now look at missingness, another important aspect of trial quality. More than 97 percent of patients completed the study and we had vital status in more than 99 percent. In the most conservative sensitivity analysis, in which all patients on empagliflozin with missing vital status were assumed to have died from
cardiovascular cause while all patients on placebo to be alive, empagliflozin significantly reduced cardiovascular death by 25 percent. Hazard ratio 0.75, upper bound of the confidence interval 0.93, with a highly significant nominal p-value.

We did similar analysis for all-cause mortality presuming all patients on empagliflozin without vital status had died, and that all patients on placebo with vital status were alive. Even in this worst-case analysis, empagliflozin significantly reduced all-cause mortality by 23 percent. The strength of these data provides a firm foundation to make a complete assessment of risk and benefits.

Professor Zinman will now provide his clinical perspective.

**Applicant Presentation – Bernard Zinman**

DR. ZINMAN: Thank you. Since cardiovascular death, as you've seen, is the number one cause of mortality in diabetes, the results of the EMPA-REG OUTCOME trial become particularly relevant for our patients. We've come a long way
in the prevention of macrovascular events in patients with diabetes. In the 1990s, major trials established the importance of statin therapy as a foundation for cardiovascular risk prevention in people with diabetes. ACE inhibitors then became part of the standard of care, with demonstrated beneficial effects on both renal function and cardiovascular outcomes.

Now we have seen the added value of empagliflozin, which in our trial was added to standard of care, including statins and ACE inhibitors. We can now offer our patients a new therapy that further reduces cardiovascular death.

What have we learned this morning from the EMPA-REG OUTCOME trial? First, and perhaps most importantly, we now have additional confidence with respect to the long-term safety of empagliflozin. Secondly, in the context of patient care, treating a thousand patients with type 2 diabetes and established cardiovascular disease for 3 years with empagliflozin would prevent 27 deaths, the vast majority of which would be cardiovascular deaths.
Clearly, these remarkable and clinically relevant benefits exceed the known risks associated with empagliflozin.

I will now hand the microphone back to Dr. Woerle to close the presentation.

Applicant Presentation – Hans-Juergen Woerle

DR. WOERLE: We would like to thank our patients, our investigators, our coordinators, and members of the steering committee, adjudication, and data monitoring committees. Without their contribution this large multinational, long-term outcome trial would not have been possible. We are proud of the work that we have done together and look forward translating these CV mortality benefits to patients.

We have with us today a number of renowned clinical trialists with expertise across several disciplines: Richard Bernstein, neurologist, stroke expert; Jared Butler, heart failure cardiologist; James Januzzi, cardiologist and member of the adjudication committee; Robert Makuch, statistical expert; Marc Pfeffer,
cardiologist; Stuart Pocock, statistical expert and
senior advisor to our data monitoring committee;
Christoph Wanner, nephrologist; Bernie Zinman,
endocrinologist and our trial chairman. Thank you.

Clarifying Questions to Applicant

DR. SMITH: Thank you.

We now have time for some clarifying
questions from the committee. And the focus here
is really on data clarification. We'll have time
for discussion later. I see a few hands up
already. If you would particularly signal to
Commander Bonner, she will list your names and I'll
try to get to everyone in order. Dr. Hiatt?

DR. HIATT: Thank you for a most informative
presentation, and I also found the background
material to be quite accessible. My primary
question at this stage is to help us understand how
you got from a MACE primary endpoint to a
cardiovascular-labeled indication.

The question I think will probably be to the
FDA later in the day, but specifically the trial
was designed as a safety trial; 3-point MACE was
the primary endpoint. In the hierarchal testing, an adjustment of alpha was based on testing for non-inferiority to demonstrate cardiac safety, and then to superiority.

I specifically noted in your materials and in the FDA materials that there was no hierarchal testing on components of the primary endpoint or other secondary endpoints. So I don't understand how we're going from MACE and the totality of the cardiovascular events to a component of MACE that seems selected to maximize demonstrating the benefit of the drug. And I guess I'd like you to help me understand where that strategy came from and how you can defend it.

DR. WOERLE: Okay. Empagliflozin reduced 3-point MACE, the primary outcome, by 14 percent on top of standard of care, independent of glycemic control. This was entirely driven by a 38 percent reduction -- can I have slide 3 up please -- reduction in cardiovascular death, without significant effects on the atherosclerotic endpoints, non-fatal MI, and non-fatal stroke.
Now there are medical and methodological reasons to support the CV indication, and I have with us Dr. Pocock, who will give his statistical perspective on the finding and this, if we are allowed, will take a few minutes.

DR. POCOCK: All right, thanks very much. Besides being a statistical expert, I sort of specialize in critiquing evidence of cardiovascular outcome trials in general. So what I'd like to give with my first slide please, the next slide, is my take on the strength of evidence and why it merits concentrating on cardiovascular death.

My next slide, please. As you have already seen, the overview of the primary endpoint findings are shown on this slide, showing a 10.5 percent primary endpoint rate on both doses of empagliflozin combined versus 12.1 percent on placebo, which gives you a primary endpoint p-value of 0.04, statistical significance in the conventional sense.

Now, I think with all primary endpoints, they are a mixed bag, and therefore you'd like to
understand what's driving it, because from a clinical patient point of view, that's what really matters. Where does the benefit really lie?

Well, you see on the rest of this slide, broken down by the three components, cardiovascular death, myocardial infarction, and stroke, you see that it's really all driven by cardiovascular death, 3.7 percent versus 5.9 percent, with an astonishing p-value, which you don't see very often, of 0.0003, whereas for myocardial infarction and stroke, there was no suggestion either way of a really firm benefit or harm. So then therefore you concentrate on the component where the action is, because that's what matters to patients.

Let's move to the next slide, please. Then you look for a consistency across doses, and this is confirmed very clearly on this slide. You can see both for the 10-milligram and the 25-milligram doses, you see a cumulatively lower rate of cardiovascular death over time, substantially below the grey placebo curve.

Above you see the hazard ratios for the two
separate doses, and that's very convincing. You see hazard ratio of 0.65 for the lower dose, 10 milligrams, and 0.59 for the higher dose, 25 milligrams. And you see the high level of statistical significance is still there when you split by dose, p of 0.0016 and p of 0.0001 respectively.

Before I go to the next slides, I'd just like to say it's important to see a consistency across subgroups -- those were slides 54 and 55, to be precise, in the earlier presentation -- and showed that across subgroups you saw a consistency of benefit of empagliflozin versus placebo for cardiovascular death. That was subgroups by baseline characteristics for the patients and also subgroups by different baseline treatments that the patients were on. So consistency is important.

My next slide, please. Then we come into multiplicity considerations. Where does cardiovascular death sit alongside all the other endpoints that we could have considered? So there were 10 prespecified adjudicated cardiovascular
outcomes. So are we cherry-picking the one that's most significant, and if we are, how do we correct for that?

Well, that's where Bonferroni correction comes in. It's a traditional way of multiplying any p-value you see by the number of endpoints that you could have studied. And you see the p-value for cardiovascular death then goes up to 0.003, so still very, very convincing. Multiplicity issues aren't really of concern.

If you want to be particularly fussy you could say what about all the endpoints, microvascular and others, that we could have looked at, 42 in all? Even applying Bonferroni correction in that circumstance, we could still get a corrected p-value of 0.0012. So there remains strong evidence of benefit after corrections for multiplicity.

Can I go to my next slide, please? Now, when we come to interpreting a remarkable benefit, there is in the back of one's mind, is it too good to be true. And therefore I'm going to use -- I'm
not a natural Bayesian. I often argue against
Bayesians in other messier settings. But I think
it's an appropriate methodology to look at how we
can integrate the evidence we have in the trial
with the prior skepticism people might have had
about not expecting a cardiovascular death benefit
for empagliflozin.

The first slide here of this Bayesian
approach shows the observed data. You say on a
hazard ratio scale in green, we see the peak of
what we're inclined to believe from the data alone
is an observed benefit of 0.62, 38 percent
reduction in cardiovascular death. But we accept
the uncertainty of that estimate given the
finiteness of any given trial. But do remember,
we're dealing here with 309 cardiovascular deaths
in the whole trial, so that's a considerably larger
amount of data than one normally gets in
cardiovascular outcome trials.

Therefore you see the spread of evidence
around that, so the probabilistic interval could go
from a benefit as high as a hazard ratio of 0.49,
51 percent reduction. On the other hand, the data themselves say, well it might go down on a 95 percent probabilistic basis, down to a hazard ratio of 0.77, i.e. 23 percent reduction.

Now, my next slide, please. It's interesting to say, how can we express prior skepticism? You can just vaguely say it's all too good to be true, but let's try and be a bit more quantitative than that and say prior belief, in the red curve, could have been that we really didn't think empagliflozin would do anything to cardiovascular death. We have an aura of pessimism from previous trials of previous drugs in diabetes.

But we could say, well, you never know. So let's propose that it could be beneficial, but we think that benefit is unlikely to be more than 25 percent. Or it could even be harmful, up to a 33 percent harm. We just don't know before we do the trial. So let's see. How can we put that prior skepticism alongside the observed data?

The next slide, please. That's where Bayes' theorem comes in, and you end up combining the
evidence with the prior skepticism to say what
should we believe now. In that spirit of too good
to be true, the prior skepticism does shrink the
estimate somewhat. So you end up now with a
posterior belief, the most likely result is that
there's a true 26 percent reduction.

I think the data were perhaps a little bit
lucky in contrast to our prior belief, but there's
still a very positive outcome. And that now ranges
from a possible 38 percent benefit down to just an
11 percent benefit, but with a strong tight
posterior belief. In other words here, the data
are so strong they overwhelm prior skepticism
preserving an underlying positive belief now.

Move to my conclusions, please, the last
slide. In general, evidence from secondary
endpoints need cautious interpretation in any
trial. But here, mortality is clearly of huge
importance, far more important than non-fatal
outcomes. As Dr. Zinman has pointed out, it merits
really special attention. It really needs
improvements to patients.
But cardiovascular death is actually the key component of the primary endpoint, so it's just helping one interpret the primary endpoint result where the action really is, which is of course cardiovascular death. A p-value of 0.0003 for cardiovascular death is overwhelming evidence of benefit. You don't see that sort of strength of evidence very often. And sometimes, we therefore adopt the phrase, proof beyond reasonable doubt of great clinical importance to future patients.

Adding in non-cardiovascular deaths preserves that survival benefit. You've seen data on that already. And therefore this cardiovascular death evidence establishes a strong foundation for benefit-risk assessment of empagliflozin. Thanks very much.

DR. HIATT: Thank you, but it didn't answer the question, but it did provide more really interesting statistical analysis, which I appreciate. The question is not why you get to CV death. The question is how you get there. So typically in a cardiovascular outcome trial, you
prespecify the primary and then you don't look at
the secondaries or components until you
statistically hit the primary, which you did on
superiority. But because it's a safety trial, you
focused on the appropriate kinds of hierarchal
testing, which all makes sense.

What doesn't make sense to me is how -- and
since you're asking us to ignore the
fatal/non-fatal MI, fatal/non-fatal stroke, the
other cardiovascular components are sort of put to
the side because there's one compelling finding.
And I understand why we're there, and obviously
that's a mortality benefit, which is incredibly
important. I just don't understand how you got
there, because legitimately this trial wasn't
designed to test a component of the primary in the
way you're doing it.

DR. WOERLE: Let's review briefly the
statistical hierarchy of testing, because you say
this is a safety trial. We started off with this
trial as a safety trial, this is correct, and the
trial was powered to demonstrate safety to pass the
1.8 hurdle. And through the conduct of the trial, we changed the scope of the trial. We enlarged the trial and we fully powered the trial with more than 690 events to demonstrate safety and potentially efficacy.

May I have slide 3 up, please? This is the hierarchal testing procedure we have prespecified in our statistical analysis plan and protocol.

When we started the trial, already in 2010, it foresaw testing for non-inferiority of 3-point MACE followed by testing for non-inferiority for 4-point MACE.

Now, here comes the important part. Both of that has been fulfilled. However now, superiority for 3-point MACE was tested. So up to this stage superiority for 3-point MACE, we were successful. However, we did not reach superiority for 4-point MACE.

Now, superiority for 3-point MACE has been established. We describe in our current indication statement the most important finding of the trial, which is this large finding on cardiovascular
mortality, overall mortality and CV mortality being reduced remarkably in a remarkable high number of events of more than 300 or 400 events, whereas the atherosclerotic components have not contributed.

We are of the belief that our data interpretation is best reflected in a descriptive statement that, while it has reduced significantly 3-point MACE, the most relevant finding to be described is the finding on cardiovascular mortality.

DR. SMITH: Thank you. This has been very helpful, but I do want to stress that as we go forward I would like the sponsor to be as brief and focused as you can in answers so we have time to get around to all the questions here. So it's really useful information, but as a precedent going forward, I'd like to be as focused and brief as we can. Dr. Konstam, you had a question.

DR. KONSTAM: Yes. Thanks very much.

If Dr. Pocock wouldn't go too far away.

Maybe I can do this. I'd like to just park my request, which is that I'd like to see the impact
of missing data on the primary MACE endpoint, success. That's the gateway to anything else we get and I haven't seen an analysis of missing data effect on that.

So if you could, hold that. But I'd like to ask Stuart to follow up on his comments because, first of all, I want to just say you know as well as anybody that there have been a number of observations of what look like fairly substantial mortality benefits out of secondary endpoints, subgroup analyses, components of the endpoint. We've seen those before and in every case that I'm aware of, when we went on and did the definitive trial on that endpoint as the primary, it didn't hold up.

In light of that, I think the issue of the p-value to me is really very important. And as you know, the agency generally requires two trials, and one of the ways they get around that is to say, maybe if the p-value is real small, it's sort of like two trials. So we have this small p-value in this secondary endpoint that is not part of the
hierarchal testing.

I get the Bonferroni correction. I think that the hierarchy that was set up is actually about adjusting for multiplicity. I mean that's what it's there for. It's prespecified. It's agreed to. I think that you can have strength that you know what those p-values mean.

Now you take this other p-value that was not part of the hierarchy and you say I can adjust that by Bonferroni correction, not prespecified. And I just want you to say, how can we really trust that? What is the mathematical meaning of that p-value? Is it sufficient to drive the two-trial requirement?

DR. POCOCK: Well, I think it's partly mathematical or statistical and partly what matters to patients. So I'm a trialist. I'm more than just a statistician. I'm a trialist. I collaborate all my life in cardiovascular outcome trials.

So I think we need to take here what matters to patients in the context of the statistical
evidence. And we have success in a formal statistical sense on 3-point MACE. You try saying to the patient, we have a benefit on 3-point MACE. He's kind of losing interest. His cardiovascular death is the driver.

DR. HIATT: I get all that. I get all that, but I'm challenging your view that you can know what that p-value really is, even though it was not part of the hierarchy and it's a post hoc Bonferroni adjustment. How confident are you that you know what that real p-value is?

DR. POCOCK: Well, a p-value of 0.00003 is --

DR. HIATT: Is small. Yes, I get it.

DR. POCOCK: -- pretty staggering. I threw Bayes at it to try and say can prior skepticism overcome that. It is a component of the primary endpoint, which I think is important, but it's the driver of the primary endpoint.

I have, on many occasions, talked about composite endpoints in general. And when one looks at composite endpoints, always I say, in general,
not just about this trial, look at the details. What's the driver within the composite? And pretty obviously the driver is cardiovascular death.

So people can nitpick about type 1 errors at anything left, but cardiovascular death is there. I think it's the most staggering result we have seen in cardiovascular outcome diabetes trials probably ever. And therefore to stay within the strict hierarchy of 3-point MACE, 4-point MACE, I think is missing the point about the, importantly, strength of evidence for cardiovascular death, a component of the 3-point MACE, and obviously deaths are more important than non-fatal events as well.

DR. SMITH: Thank you. Dr. Proschan, you had a clarifying question?

DR. PROSCHAN: Yes. On CQ90, you talked about the fact that the amendments were made before looking at any data. So what made you think that you were going to be able to show superiority without looking at the data? That increase to 7,000 -- why did that thought even cross your mind?

DR. WOERLE: The last part of your sentence,
I didn't hear.

DR. PROSCHAN: Pardon?

DR. WOERLE: The last part of your sentence I didn't hear.

DR. PROSCHAN: Well, basically, what made you think that you were going to be able to show superiority if you didn't look at the data?

DR. WOERLE: Three things; medical considerations, effect on blood pressure, of which we know is very important when it comes to outcome. We knew that the drug has a diuretic effect. We knew it reduces glucose and it reduces body weight. From earlier trials, we saw that all relevant biomarkers are going in the right direction.

We already speculated in 2010 about a potential beneficial effect in heart failure, which is very important when it comes to mortality in type 2 diabetic patients. That's why we prespecified the endpoint heart failure and adjudicated the events. All those considerations, together with some also preclinical evidence, led us to enter this journey to test here for
full-blown superiority.

DR. PROSCHAN: But weren't all those things that you just mentioned -- weren't they available before the trial? So what made you decide that you needed to amend it?

DR. WOERLE: See, the guidance came out in 2008. At the time we put together this trial, together with our phase 3 program, there was a large amount of uncertainty, how to best address this issue. There were proposals that you could do two trials. One, you do pre-registration. You could do another one post-registration and try to pool. Or you do a small trial and you see if you clear the 1.8 hurdle, and then you do another trial.

We had some consultation also with the FDA as part of our end of phase 2 meeting. Together with statistical consideration, we said the best approach is to go with one large trial, which allows you to test for non-inferiority and superiority. And this thinking process took some time until we took the decision.
DR. SMITH: Dr. Cooke, you had a question?

DR. COOKE: With regard to the cardiovascular outcomes, including 3-point MACE and the others, when you look at the efficacy of the two different doses, it seems like they basically overlap. So I'm curious, your interpretation of why there's a lack of a dose-response effect.

DR. WOERLE: When we knew from early trials in phase 2b -- remember, when we designed this trial at a point we had no information on larger scale trials, but we knew from our phase 2b trials that there is a small dose response in terms of HbA1c and very subtle also in terms of blood pressure lowering and BMI.

We knew from earlier trials in diabetes that the small differences in glycemic control, which we most likely will see in phase 3, will not contribute to hard outcome, because this takes, if any, 10 years or longer to establish. So our underlying assumption was that most likely we will see very similar results with, if any, very small dose response effects in terms of hard outcome.
And that's exactly what we saw.

But what is remarkable? I wanted to make that point about the assurance of the data, that we see this consistency with two doses. And even though it's not entirely proper, but you almost could see this as two studies in one, with this remarkable consistency on CV death.

DR. SMITH: So I have a question. And what I'm looking for is some more clarification on the non-assessable death category being assigned to presumed cardiovascular death. I still don't fully understand. What I'm looking for is to understand what you actually did and some of the specifics of that. Because I understand the words, but I actually don't have a good clinical understanding of what was done and what those patients were.

DR. WOERLE: Cases of CV death that could not be attributed to CV death or non-CV death were prespecified in our adjudication charter as presumed CV death. Now, fortunately, we have today Dr. Januzzi with us, who was part of the adjudication panel who adjudicated all those
events, who is best equipped to address your
question.

DR. JANUZZI: Good morning, everyone. It's
a pleasure to represent the clinical endpoints
adjudication committee. We're an independent group
of cardiovascular specialists and clinical
trialists that review the cases through the several
years of the study. And specifically, what I'd
like to do is answer the question that Mr. Chairman
posed, but address on a methodological level this
question of the higher number of presumed CV death
events.

The first point, as any clinical trialist
will know and people that have performed clinical
endpoints adjudication, the heart and soul of the
CEC is the charter by which we judge cases. We
review cases from the beginning to the end of the
trial using the charter as our gold standard.

So in order to provide high level and
consistent high quality data, it's necessary for
the CEC to follow specific definitions in evidence
requirements to meet the definition of
cardiovascular death, or non-cardiovascular death as prespecified in the adjudication charter.

Stated in a more simple way, if patients were submitted for event review that did not meet specifically the expectations outlined in the charter, by definition, they were presumed CV death because of the lack of ability to assess.

The point here is it's done to ensure consistent evidence-based and reproducible work from beginning to end of trial out of the adjudication committee. So therefore, the high number of presumed CV death, in the opinion of the reviewers, reflects the fact that we adhere very carefully to the definitions set forth in the charter for meeting CV death.

Now, Mr. Chairman, to answer your specific question, this was a broad range of possible scenarios, as you probably know through your review. In many trials, for example, patients found dead in bed would be classified in some studies as sudden cardiac death, whereas in the EMPA-REG OUTCOME study, if we lacked the specific
variables to judge sudden cardiac death, we would be compelled to call it presumed CV death. So that's just one example of how those numbers might be affected somewhat through our processes.

DR. SMITH: Okay, I think I'm getting more clear about it, but I see -- Dr. De Lemos, yes?

DR. DE LEMOS: Pull up CO-57.

DR. WOERLE: Slide up, please.

DR. DE LEMOS: Just to clarify, three-quarters of the cardiovascular deaths in this trial were either sudden death or presumed CV death. And so that's really the issue if we're going for a CV death claim here. And the question -- in the charter, the individuals with sudden death include a last category on page 124 of your briefing document, where the last category of sudden death is presumed CV death.

Now, is this category at the bottom that category? Or in fact some of the sudden death is actually unwitnessed so that the proportion of individuals in whom we really don't know how they died is actually larger? Does that make sense, my
So the sudden death definition in the charter includes both witnessed sudden death and unwitnessed death. And the question is, is that bottom category the unwitnessed sudden death or are many of those sudden deaths actually unwitnessed?

DR. JANUZZI: That's a good question, James. So basically, in order to be termed sudden death, they had to meet specific charter definitions. Therefore, in the presumed CV death, there were patients that were found dead in bed, is the term that's used, who, given their background cardiovascular risk.

These patients were enrolled in a cardiovascular safety trial. The mode of death in many trials would have been called sudden death. But in this study, because they did not meet the charter definitions, we called them presumed CV death. Maybe I'm not being clear. To be simple --

DR. DE LEMOS: Well, there's a bullet. The question is how many of the sudden -- there are five bullets under the sudden death definition.
Four of them require some witnessed event, and the fifth says -- the fifth bullet under sudden death reads, "Unwitnessed death," and there's no conclusive evidence of another non-cardiac cause of death.

The question is, did you pull that group out as presumed CV death? Or in fact are the sudden deaths mostly unwitnessed as well, in which case the conclusion is that we basically don't know why the large majority of these people died, which is understandable? It's hard to figure out why people die in an outpatient trial, but if it's a CV death claim, that's a different question.

DR. BROEDL: Uli Broedl, Medicine. To answer your question, indeed in addition to the requirement of witnessed death, there was one bullet point, as depicted in our briefing book, that refers to an unwitnessed death. But it also says that the patient should have been seen alive within the last seven days.

So in essence, if someone has not been seen for more than a week or so and then was found dead
in bed or if, for instance, at the end of the day we have only very limited information what happened to this patient, this patient could then either be categorized, if seen within the last seven days, as sudden death.

Or if nothing was available and the patient has been seen for quite some time or limited information, or only at the end of the day a death certificate was available, those patients would have been categorized according to the prespecified charter as presumed CV death.

DR. DE LEMOS: Would it be possible at the break to pull them out? I think when we're thinking about an indication in a single trial to meet a definition of CV death, it's important to see what the impact of these presumed deaths and these unwitnessed deaths are on the strength of the association for a single trial for CV death, I think. And that goes to Marv's point, I think about --

DR. BROEDL: Then maybe Dr. Januzzi can comment on this. But what has been reported on the
CEC voting form, at the end of the day, is CV death and then the type of CV death, but not the reason why presumed CV death versus sudden death was chosen. So we cannot just break down the five bullet points.

But what we can do to give you some reassurance is review slide 1, please again, which is the sensitivity of the cardiovascular analysis that I have presented earlier today. Given the uncertainty around presumed CV death, we conducted this sensitivity analysis where we excluded presumed CV death, 124 cases. And the results for empagliflozin on CV death remained significant, virtually identical.

DR. SMITH: So if I understood, you're not able to produce the requested breakdown in data -- did I understand that -- for later today, as Dr. De Lemos was requesting?

DR. BROEDL: That is correct.

DR. SMITH: All right. I think we should move on now. We've got just a few more minutes here before the break. Dr. Good had a question.
DR. GOOD: Thank you. As a neurologist, I do have a question about the important secondary outcome measure of non-fatal stroke. Of course, as you pointed out, the HR is higher in the treatment group compared to placebo, although it's not statistically significant, I understand that. So, you noted, when you added TIA, that the HR approached 1.

So a comment -- the diagnosis of TIA has some clinical noise in it. Some things that are thought to be TIAs are not TIAs, and it might falsely suggest that it isn't as close to 1 as you suggest.

The other thing I'd be interested in is that I'd be interested in the number of disabling strokes, which is quite important to patients, and how you determine the disability. You may want to talk about this, this afternoon.

The final thing I want to comment on is the disparity between stroke incidence among different regions. In fact, there's some suggestion that the incidence in Europe was much, much higher than
other regions, and that puzzles me. It suggests there's a variability perhaps in diagnosis in different regions, and Latin America is very low. So I'd just like you to comment on that and it raises some question in my mind about the accuracy of diagnosis across the regions.

DR. BROEDL: First we have to say, this trial was set up as a cardiovascular outcome trial. We did not set it up as a stroke trial. So there is some certain limitation on the degree of information we have from this trial.

From our point of view, the most important hazard ratio is actually fatal and non-fatal stroke together, because that should go along with each other. And there, the hazard ratio is 1.18, with largely overlapping confidence interval, indicating a non-significant finding.

We have a team here of Sven Koehler and a stroke expert who reviewed extensively the data. And I would ask Richard Bernstein, who is a stroke expert, to give his assessment on when he looked at the stroke data, in particular on disabling stroke
and also the regional question.

DR. SMITH: Again, what I'm going to ask for is to as specifically and briefly as possible answer the question that Dr. Good has raised, rather than just give us some more background on this.

DR. BERNSTEIN: Thank you. My name is Richard Bernstein. I'm a stroke neurologist and stroke clinical trialist at Northwestern University. Thank you, Dr. Good, for your question. Let me talk first about the stroke severity issue relating to fatal and non-fatal. Slide 2 up, please.

When we reviewed stroke in some detail, when I did with the team, we wanted to look at the range of severity of stroke endpoints because of this 1.18 hazard ratio. And if this finding were a real finding, for example, I would have expected to see an increase across the range of severity from fatal, non-fatal, non-disabling, and TIA, which are all really the same disease.

In fact, if you look at this forest plot,
what is see is the point estimate bouncing back and forth across unity, so that there are fewer fatal strokes and disabling strokes, a few more non-fatal strokes, fewer TIAs. And if you put stroke plus TIA together as a total cerebral vascular event count, we're at about unity. So to me, this oscillation back and forth suggests that it's all sort of a random finding.

Regarding region, if we could get the slide up by different regions, yes, here we are. Slide 2, please. Thank you. So looking at stroke by region, what's most striking to me is what is likely anomalously low incidence of stroke in the placebo group, which as a clinical trialist is the thing you never really want to see. But we were unable, in going through all of the risk factors and data in detail, to find any explanation. And so my conclusion is this is a chance finding.

DR. SMITH: Okay, Dr. Neaton, you had a question?

DR. NEATON: I did. Thank you. So actually, I had a question that will ultimately
lead to the same one Dr. Konstam had. So you had
an extreme result for cardiovascular mortality, but
the component for which it's part, 3-point MACE,
the p-value is more borderline. And there's also
more unknown information about those events than
for mortality, which makes sense.

So two questions. One, precisely, I could
not determine how you did the censoring when you
did the 3-point MACE because there were deaths,
obviously, which must have occurred after you
became unknown about the non-fatal status, for
example, and how those were handled. And also what
other sensitivity analyses were done around that
3-point MACE outcome that kind of prompts us, if
you will, to think very carefully about the
contributing components?

DR. WOERLE: We counted. This was a time to
first event analysis, so patient could have had a
stroke, time to first event.

DR. NEATON: No, I understand that, but I'm
particularly interested in understanding precisely
when you censored follow-up for people when they
had an unknown event status because the unknown event status could have happened earlier for non-fatal events than fatal events.

    DR. WOERLE: I see.

    DR. NEATON: So presumably a person could have died from cardiovascular disease after you lost track of them for non-fatal MI and stroke.

    DR. WOERLE: Our project statistician is best equipped to answer this question, Stefan Hantel, may I ask you?

    DR. HANTEL: Stefan Hantel, biostatistics. When we looked at MACE 3, we censored the patient as presented in the presentation at last time when we were aware that the patient was event free. This means vital status and knowing that the patient is alive at study end was not sufficient. Therefore, we have only 97 percent of patients followed up for MACE 3 at the end of the trial, but 99 percent for vital status.

    So the difference is that we censored patients with a positive vital finding in the sense that the patient was still alive at the last time
we were aware that the patient was event free for non-fatal stroke, non-fatal myocardial infarction.

DR. NEATON: Presumably a person could have died from cardiovascular disease after that time point and they would have been excluded from the analysis?

DR. HANTEL: These patients were included because we tried to capture the event. So you are right. In this patient, it can happen that we missed a non-fatal event. But if you focused on --

DR. NEATON: More of my question is, what happens if you've lost track of their non-fatal endpoint status and then later on, you determine that they've died, say, from cardiovascular disease? Is that event counted?

DR. HANTEL: Yes.

DR. NEATON: It is counted?

DR. HANTEL: Yes.

DR. NEATON: So what sensitivity analyses have you done around that censoring mechanism, and more generally the missing non-fatal endpoint, to kind of convince us that 0.038 is really robust?
DR. HANTEL: We had prespecified sensitivity analysis focusing on events happening up to 30 days after treatment stop. This was presented earlier today.

DR. NEATON: But we're interested in the missing data. You had 144 people, for example, with unknown non-fatal endpoint status for MACE 3.

DR. HANTEL: What we did for MACE 3, when we assume, looking at vital status, patients are that in EMPA and alive in placebo. The hazard ratio is almost unchanged. When we went through this analysis for MACE 3, assuming that those patients we didn't follow up to the end, we would lose significance for superiority but still maintain non-inferiority.

DR. NEATON: That's a very extreme kind of sensitivity analysis and so I accept that that's what you would find, but have you done any kind of analyses to understand the reasons for missingness that would give us kind of something that might be closer to the truth?

DR. HANTEL: What we did not perform is like
multiple implementation to address missingness.
What we can show you is time to censoring for
MACE 3. Slide 2 up, please. And here we see that
the time to censoring is not really different
between empagliflozin and placebo.

DR. NEATON: I only see one line --

(Laughter.)

DR. HANTEL: I'm close enough, so there are
some small differences, but it's virtually
identical.

(Laughter.)

DR. NEATON: We'll come back later and talk
to you about your next presentation about the
censoring. Thanks.

DR. WOERLE: Yes. It tells you about the
accuracy of our statisticians.

(Laughter.)

DR. SMITH: We need to take a break. We're
going to take about a 15-minute break. I'm going
to start just a little bit before 10:30, and we'll
return to those panel members who have questions
that we didn't get to. We'll get to those later
Panel members, please remember there should be no discussion of the meeting topic during the break among yourselves or with any member of the audience. So we'll start about 28 past the hour here.

(Whereupon, at 10:14 a.m., a recess was taken.)

DR. SMITH: I would like people to start taking your seats. We're going to start in just another minute or so here. So we'll now proceed with the FDA presentations, and we will come back to more clarifying questions later on today.

**FDA Presentation – Andreea Lungu**

DR. LUNGU: Good morning. My name is Andreea Lungu, and I'm a clinical reviewer in the Division of Metabolism and Endocrinology Products. On behalf of the review team, I would like to thank the committee for being here today.

Today, I will be presenting the FDA's review of the findings from the EMPA-REG OUTCOME study. I will refer to this study as the EMPA-REG study today.
throughout my presentation. I will --

DR. SMITH: Could we have the volume up just a little bit?

DR. LUNGU: I will begin by briefly reviewing the regulatory history for this trial, which is complex. I will follow with a description of the EMPA-REG OUTCOME trial, including the objective, design, trial conduct, and study population. Dr. Jennifer Clark will then present the FDA's statistical assessment.

Following her presentation, I will discuss the factors affecting the interpretation of the primary endpoint, selected exploratory endpoints, followed by some of the non-cardiovascular safety findings. I will then conclude with a summary of the FDA presentation.

The EMPA-REG OUTCOME trial was conducted to address the 2008 guidance for industry on evaluation of cardiovascular risk for new anti-diabetic drugs, and was primarily designed to address a safety question. As a result, the trial design may differ from trials conducted to address
an efficacy question.

As a reminder, the guidance recommends that applicants exclude 80 percent excess ischemic cardiovascular risk pre-approval, and 30 percent excess ischemic cardiovascular risk post-approval. This can be achieved through a meta-analysis or a dedicated trial. I will now discuss how the EMPA-REG study was utilized to address this guidance.

The role of the EMPA-REG study in the cardiovascular risk assessment evolved over time, and changes to the trial design occurred while the trial was ongoing. These changes, to some extent, reflected changes in the applicant's strategy to address various aspects of the guidance, and were in part informed by advice received from the FDA.

The initial proposed cardiovascular risk assessment is shown here. The EMPA-REG study is shown in red and was to provide cardiovascular events for the purpose of evaluating the pre-approval cardiovascular risk margin only. There were three planned looks to exclude the
1.8 risk margin, denoted by the three arrows. The first was to occur at completion of phase 3 diabetes trials. The second was to occur when 60 events had accrued, and the last after accrual of 152 events.

The role of the EMPA-REG study in the assessment of the post-approval cardiovascular risk margin was unclear. The agency communicated that this proposal lacked clarity, did not sufficiently control for type 1 error, and was not powered to address the 1.3 risk margin. The agency asked that these issues be addressed before agreeing to the plan.

The second proposal for the cardiovascular risk assessment is shown here. The EMPA-REG study is again shown in red. In this proposal, the EMPA-REG study was to provide events for the purpose of evaluating both the pre-approval and post-approval cardiovascular risk margins.

The applicant proposed to evaluate the 1.8 cardiovascular risk margin using a meta-analytic approach and to perform the first analysis when 118
MACE-plus events had accrued. The applicant proposed to then initiate a second cardiovascular outcomes trial, shown in green, which would be combined with the EMPA-REG study to evaluate the 1.3 cardiovascular risk margin using a meta-analytic approach when 711 MACE-plus events had accrued.

The applicant had plans to test both noninferiority and superiority hypotheses after accruing all 711 events. The FDA commented that increasing enrollment in the EMPA-REG test both non-inferiority and superiority study would be more efficient than initiating a second study for the purpose of excluding the 1.3 cardiovascular risk margin.

The FDA also expressed reservations with the applicant's plan to use a meta-analytic approach to test for superiority for the purpose of obtaining a cardiovascular risk reduction claim. The FDA recommended that the applicant consider using a dedicated outcomes trial if their intent was indeed to pursue a potential cardiovascular risk reduction
The final cardiovascular risk assessment plan is shown here. The plan to exclude the 1.8 risk margin remained the same. The plan no longer sought to rely on a second cardiovascular outcomes trial to accrue sufficient number of events, but added an additional 3,000 patients to the ongoing EMPA-REG study instead. Changes to the EMPA-REG study protocol were made to reflect this change in strategy and submitted with the final cardiovascular risk assessment plan. The FDA agreed with this plan.

With this final plan, the protocol was changed and the trial size was adequate from this point onward to independently exclude the 1.3 risk margin for MACE. The protocol specified that the trial would stop once 691 MACE events accrued. The protocol also specified that the planned interim analysis would occur after 118 events had accrued.

The plan for the interim appropriately adjusted the type 1 error to account for this interim look. Superiority testing was allowed, and
was to be conducted only after the 1.3 risk margin was excluded for both MACE and MACE-plus.

The empagliflozin new drug application was submitted in March 2013. The database for the EMPA-REG OUTCOME study interim analysis supporting the pre-marketing cardiovascular risk assessment had been locked on August 21, 2012. The analysis for the pre-marketing cardiovascular risk margin was based on a total of 196 MACE events, with the majority of these events being contributed by the EMPA-REG study.

Interim data from the EMPA-REG study was used to test both the 1.8 and 1.3 risk margins at that time. To give you an idea of the ongoing trial, approximately 5,000 patients had been randomized into the EMPA-REG study at the date of the database lock.

Finally, to support regulatory filing of the application in the U.S. and worldwide, some level of unblinding at the applicant level occurred. Indeed, approximately 230 individuals were given access to patient-level data and interim results.
for the ongoing trial. The unblinded individuals were required to sign a confidentiality agreement stating that the results would be kept confidential.

The results of the prespecified pre-marketing cardiovascular risk analysis revealed no signal of increased cardiovascular risk associated with the use of empagliflozin. The hazard ratio for 1.8 and for the prespecified interim analysis for 1.3 are shown here.

Empagliflozin was approved in August 2014. At the time of approval, the FDA required that the applicant continue evaluating the cardiovascular risk associated with empagliflozin use by issuing a postmarketing requirement.

The requirement specified that the EMPA-REG study alone be used to evaluate the postmarketing risk margin, and that the assessment be based on 3-point MACE. Adverse events of interest, unrelated to cardiovascular risk, were to be collected as part of that requirement.

Having covered the regulatory history, I
want to now spend some time reviewing the study design and study conduct. There were some general features of the design that did not change over the course of the study, and these are illustrated here.

The study was always a randomized, double-blind, placebo-controlled trial. The study population evaluated in the study were adults with type 2 diabetes who were also at high risk for atherosclerotic cardiovascular disease.

The duration of the study was always to be event driven. Subjects were randomly assigned to 2 doses of empagliflozin or placebo in a 1 to 1 to 1 fashion. The investigational drug or placebo were added to local standard-of-care drugs for both diabetes and atherosclerotic cardiovascular disease.

Therapies for diabetes and atherosclerotic cardiovascular disease were to be adjusted throughout the trial by treating physicians to achieve therapeutic targets set by local professional guidelines. These instructions were
expected to result in minimal differences in
glycemic control and cardiovascular risk factors
susceptible to confining interpretability of the
final study results.

The primary objective of the study was to
evaluate the ischemic cardiovascular risk of
empagliflozin, and the primary endpoint variable
was always the time to a first major adverse
cardiovascular event consisting of a 3 part
composite comprised of cardiovascular death,
non-fatal myocardial infarction, and non-fatal
stroke.

Key participants involved in study conduct
and operation also did not change. The applicant
was responsible for the conduct of the trial,
communication between investigators and various
committees and regulatory bodies, analyzing the
interim and final results, and reporting these
results to the various worldwide regulatory
agencies.

The steering committee provided scientific
leadership for the design and conduct of the study.
An independent data monitoring committee was established to monitor the progress, safety, and efficacy of the empagliflozin phase 3 clinical trials, including the EMPA-REG.

The clinical event committee was responsible for central blinded adjudication of all major adverse cardiovascular events. This committee was also to adjudicate MACE events for another anti-diabetic program.

Finally, the general process for event adjudication did not change drastically over the course of the study. All fatal events and any events suspected of stroke, transient ischemic attack, myocardial ischemia, hospitalization for unstable angina or heart failure, stent thrombosis, and revascularization procedures were to be referred for adjudication. The adjudication process, briefly outlined on this slide, was a standard adjudication process for outcomes trials.

While the general features of the trial did not change drastically, there were some important changes made over the course of the trial to the
clinical trial protocol, the CEC charter, and
statistical analysis plan that I will now discuss.
To frame that discussion, it may be helpful to
orient yourselves with some of the major milestones
of the trial, which are shown in the figure.

The date of first consent occurred in July
2010, and the first patient was randomized in
Canada on September 15, 2010. At this point, the
FDA had not seen the protocol. The date for
interim data cutoff was June 22, 2012, and the date
of interim database lock was August 2012. The last
visit for the last subject occurred April 13, 2015,
and the date of final database lock occurred on
June 22, 2015.

I will use this timeline as a frame of
reference to describe other changes. I will
discuss some of the major changes, starting with
the changes to the clinical trial protocol, in the
next few slides.

The original version of the protocol was
finalized on May 10, 2010. This version was not
submitted to the FDA as the study was not yet
initiated in the U.S. at this time. Version 2 of the protocol, dated September 22, 2010, was submitted to the FDA to support enrollment of patients from U.S. sites. Comments on the protocol were not requested at that time.

In April of 2011, the applicant submitted version 3 of the protocol, which included changes to the inclusion and exclusion criteria, mainly to liberalize the criteria to allow for expanded enrollment of patients with single-vessel disease in a major coronary artery. Previously it only allowed patients with multi-vessel disease.

At that time, approximately 1200 individuals had already been randomized. As illustrated here, this change was made prior to the interim analysis. The applicant reported that these changes were made based on advice received from the steering committee.

Version 4 of the protocol was finalized December 2011, and submitted to the FDA in January of 2012. The changes made in this version were in response to the discussions that occurred between
the applicant and the FDA with respect to version 2 of the protocol. At that time, 3400 individuals had been randomized into the trial.

Many important changes were made in this version of the protocol. Some of these changes reflected advice received from the FDA related to the overall cardiovascular risk assessment strategy for empagliflozin, and included changes to the sample size and anticipated trial duration.

This change was made to ensure that at least 691 MACE events would be captured in the EMPA-REG study. From this point onward, the trial was powered to independently exclude the post-approval 1.3 cardiovascular risk margin.

Additional changes included specifying that silent MI would be excluded from the primary endpoint. This version also specified that the within-trial interim analysis to assess for 1.3 would be carried out, and that type 1 error would be controlled to account for this analysis.

All these changes to the protocol occurred prior to the interim analysis, as you can see from
the timeline. While the FDA agreed with the
general changes, the issue of silent MI was not
specifically discussed.

Finally, version 5 made some specific
changes to require that ECG be reviewed by
investigator, change the follow-up period from 4
weeks to 30 days, and added adjudication for
malignancies and hepatic events to address
potential safety concerns.

While I have limited my discussion of the
protocol changes to the major revisions, it is also
worth noting that there were changes to secondary
and tertiary exploratory endpoints.

I will now discuss the changes to the CEC
charter. The CEC charter contained the endpoint
definitions to be used in the adjudication process.
I will focus on changes that were made to endpoint
definitions over the course of the trial.

The original CEC charter was created prior
to the initiation of the EMPA-REG study for use
with another anti-diabetic development program.
The empagliflozin program was added to the existing
CEC charter. There were at least nine versions of
the CEC charter, and here I will limit my
discussion to changes that are potentially relevant
to the interpretation of cardiovascular outcomes.

Changes to the endpoint definitions were
made in version 6 of the charter. Selected changes
are presented on the slide. The myocardial
infarction definition was liberalized from
requiring symptoms of ischemia and two supportive
criteria, ECG, biomarker, or imaging, to requiring
symptoms of ischemia, and only one additional
supportive criterion.

Also, the cardiac biomarker elevation
criterion was changed from one requiring the
99 percentile of the upper reference limit to one
requiring that cardiac biomarker be found to be
above the upper reference limit for assay.

The hospitalization for unstable angina and
hospitalization for heart failure were also
expanded to allow for overnight admission to chest
pain observation units in addition to emergency
room visits and hospital admissions. The stroke
definition only had minor modifications.

These changes were implemented prior to the interim analysis, were dated February 18, 2012, and were made at the recommendation of the CEC members.

The definition of hospitalization for heart failure was again changed in version 8 of the CEC charter on April 4, 2014. This change further liberalized the definition and allowed for inclusion of hospitalization for heart failure events if oral diuretics had been increased or initiated. Prior to that, only events resulting in initiation or increase in IV therapies qualified. This definition change occurred after interim data unblinding.

In version 9 of the charter, dated December 2014, subdural hematoma was removed from the definition of ischemic stroke. Although the endpoint definitions changed significantly during the trial, events were adjudicated according to the applicant based on the last definitions.

It should be noted that some definitions differ significantly from the current standardized
definitions for cardiovascular disease and stroke endpoints, and in particular the definition for hospitalization for heart failure. The impact of these changes on study results could not be quantified retrospectively.

I will now discuss key changes made to the statistical analysis plan for this trial. The statistical analysis plan for the EMPA-REG study was finalized prior to the interim analysis. In the initial statistical analysis plan, dated August 24, 2012, the primary endpoint was listed as 3-point MACE, which is a composite of cardiovascular death, non-fatal myocardial infarction excluding silent MI, and non-fatal stroke.

The secondary endpoint was 4-point MACE, which is a composite of cardiovascular death, non-fatal MI excluding silent MI, non-fatal stroke, and hospitalization for unstable angina. Silent MI and heart failure requiring hospitalization were other secondary endpoints.

A revised statistical analysis plan, dated...
May 2015, was received in November 2015. The plan did not make any changes to the primary and secondary endpoints, but added a significant number of other exploratory endpoints, some of which I will discuss later in this presentation.

This version, for the first time, specified the silent MI definition that was to be used for analysis. The statistical analysis plan also added language to clarify that non-assessable deaths are presumed to be CV death and therefore included in the analysis of the primary and secondary endpoints.

I will now describe baseline demographic and disease characteristics of the study population and summarize the disposition of the participants. Baseline demographic characteristics and diabetes history were generally balanced in the analysis population.

In both arms, the mean age of the study participants was 63 years, and the mean hemoglobin A1c was 8.1 percent. The majority of subjects had been diagnosed with type 2 diabetes for more than 5
years, with over half of subjects diagnosed for more than 10. The proportion of patients reporting diabetic complications was also balanced between the treatment groups.

The treatment groups were also similar in terms of baseline cardiovascular disease history. The history of coronary artery disease was present in around 76 percent of individuals at trial entry, and almost half of the subjects reported a history of past myocardial infarction. Around 23 percent of patients had a history of stroke, and around 21 percent a history of peripheral artery disease.

I will now turn to participant disposition. My slide shows participants in the analysis population, which excludes participants from sites with significant good clinical practice violations and 8 individuals who were randomized but never treated.

In the analysis population, 7,020 subjects were randomized and treated with at least one dose of the intervention. Information on MACE was available for 6,809 subjects at the end of the
study, or approximately 97 percent of the original cohort. 211 subjects prematurely discontinued the trial and did not have complete data on 3-point MACE. These individuals account for the missing information on MACE.

Information on vital status was available for 6,967 subjects, approximately 99.2 percent of the original cohort. Only 53 subjects did not have information on vital status. These individuals account for the missing information on death in this study.

I will now turn the presentation over to our biostatistician, Dr. Jennifer Clark, to review the statistical assessment of the EMPA-REG safety trial.

**FDA Presentation - Jennifer Clark**

DR. CLARK: Thank you. Good morning. My name is Jennifer Clark. I will be giving a presentation with the statistical assessment of the EMPA-REG trial. I'll start with a brief overview of aspects from the trial that were important for the statistical assessment. This includes the
analysis population, trial duration, and
statistical methods.

I will go over the trial results and the
testing hierarchy that was used to preserve the
type 1 error. In order to understand the results,
we will take a closer look at the primary 3-point
MACE results and the components that make up this
composite endpoint. I will also consider all-cause
mortality within the trial, as well as follow up
for mortality and cardiovascular MACE endpoints.
Sensitivity analysis results will be shown for both
MACE and death.

This trial was part of a diabetes drug
requirement to rule out a 30-percent increase in
cardiovascular risk when compared to placebo. This
is the first of the diabetes safety trials where
superiority is being considered for what were
initially safety endpoints.

The duration of this trial was less than
five years, with first randomization occurring on
September 15, 2010, and the last on April 19, 2013.
Those with last study visits on or after
December 15, 2014 were considered completers for the study.

The primary objective of this trial was to demonstrate safety with non-inferiority of empagliflozin against placebo for major adverse cardiovascular events, or MACE. Type 1 error was controlled for multiplicity through a prespecified testing hierarchy. This hierarchy only controlled error for the 3 point and 4-point MACE endpoints.

The first step in the hierarchy was testing an increase in CV events using a non-inferiority margin of 1.3 for 3-point MACE. If a 30 percent increase in CV events with 3-point MACE is ruled out in the first step, then 4-point MACE would be tested using the same non-inferiority margin. If a 30 percent increase in CV events was ruled out for both MACE endpoints, then similar methodology could be used to test for superiority, first for 3-point MACE, and then 4-point MACE.

Patients were randomized to either placebo, 10 milligrams of empagliflozin, or 25 milligrams of empagliflozin. The treated set was prespecified.
for the primary analysis. There were 8 subjects, which were randomized to a treatment arm but failed to start treatment, so they were not included in the primary analysis. There was 37 subjects who started treatment but were excluded due to site non-compliance or other issues. The two empagliflozin treatment arms were prespecified to be pooled together for the primary analysis.

A Cox proportional hazard model was used for the primary analysis. The model included adjusting baseline covariates for age, sex, BMI, HbA1c, eGFR, and geographic region, along with study treatment. Hazard ratios were based on Cox model results.

There was an additional adjustment to the type 1 error due to the prespecified interim analysis, which had a data cutoff of June 22, 2012. Because of this, the final analysis results here will be based on 95.02 percent confidence intervals instead of the usual 95 percent confidence intervals.

Here, I will be presenting the MACE results within the prespecified testing hierarchy. In
order to be considered non-inferior to placebo, the upper bound of the 95.02 percent confidence interval for the hazard ratio had to be below 1.3. When testing for superiority against placebo, a similar methodology is used with the upper bounds, but here the bound must be less than 1.

Three-point MACE has a hazard ratio of 0.86 with an upper bound of 0.99. 4-point MACE has a hazard ratio of 0.89 with an upper bound of 1.01. When looking at this in the testing hierarchy, we see the first hypothesis for ruling out a 30 percent increase in CV events is met since the upper bound for 3-point MACE is below 1.3. Non-inferiority against placebo for 4-point MACE is also achieved since this upper bound is also below 1.3.

The next step in the hierarchy is to move to testing for superiority. Since the upper bound for 3-point MACE is below 1, the criteria for showing superiority is met for this endpoint. This means this result supports cardiovascular benefit when comparing empagliflozin to placebo. This is
important because it's what the efficacy claim is
based upon.

When we progress to the last step in the
hierarchy, we see that superiority is not met for
the 4-point MACE endpoint since the upper bound is
not below 1, but non-inferiority still holds. At
this point, all alpha for type 1 error is
considered lost.

Formal statistical testing would stop here
had there been any further hypotheses prespecified
in the hierarchy. Since the 4-point MACE endpoint
did not attain superiority, it will not be further
discussed in this presentation.

Kaplan-Meier curves of the 3-point MACE
endpoint for both the pooled empagliflozin arms and
the placebo arm are shown here with the number
remaining at risk at the bottom of the table. We
see a separation of the curves around 3 months,
which continues throughout the study.

The total number of patients experiencing a
MACE event during the study period was 490 for
empagliflozin and 282 for placebo. Since the
number of patients in the pooled empagliflozin arms was approximately double the amount on placebo, this translates to 12.1 percent of subjects having a MACE on placebo and 10.5 percent of subjects with MACE on empagliflozin.

Three-point MACE is a composite made up of CV death, none-fatal MI, and non-fatal stroke. In order to understand what this and the primary hazard ratio results actually mean we need to look at these and the related counterpart outcomes.

The first component of CV death had the largest difference between treatment arms and the number of patients experiencing an event; 172 or 3.7 percent of patients had a CV death in the empagliflozin group versus 137 or 5.9 percent of those on placebo.

When looking at the related endpoint of all-cause death, of which CV death is a subset, we see that the difference in the number of patients experiencing an event is proportionally similar between the two treatment arms; 5.7 percent of subjects on empagliflozin died versus 8.3 percent
on placebo.

The second component of 3-point MACE was non-fatal MI. While there were more of these events in total than there were of CV deaths, they were not as disproportionately dispersed between the two treatment arms; 4.5 percent of patients on empagliflozin experienced a non-fatal MI versus 5.2 percent on placebo.

This close disbursement of events was similarly mirrored when looking at patients who experienced any sort of MI event, fatal or non-fatal, with 4.8 percent experiencing an MI on empagliflozin and 5.4 percent on placebo.

The third component of 3-point MACE was non-fatal stroke. This had the smallest number of events with a slight imbalance favoring placebo; 3.2 percent of subjects on empagliflozin experienced a non-fatal stroke versus 2.6 percent on placebo.

This imbalance is again reflected in the related outcomes of all strokes. Here, we have 3.5 percent of subjects on empagliflozin
experiencing a stroke versus 3 percent on placebo.

It is clear from these figures that the biggest
difference between the treatment arms lies in the
CV death component, which favors empagliflozin.

The estimated incidence rate is based on the
number of events, shown on the previous slides, and
the total number of patient years observed in the
study. For the primary 3-point MACE endpoint,
there were 3.7 events per 100 patient years on
empagliflozin and 4.4 per 100 patient years on
placebo.

The difference between the treatment arms
for the estimated incidence in each of the
components reflects the differences that were seen
in the number of events in the treatment arms. CV
death and its corresponding all-cause death outcome
show the greatest difference between treatment
arms.

We used the same Cox proportional hazards
model from the primary analysis to estimate hazard
ratios and 95 percent confidence intervals for each
of the components of 3-point MACE. The results are
shown here on this plot with a line drawn at 1 to show where hazard rates would be considered the same between the two treatment arms. The hazard ratios are in line with what we would expect based on the number of events occurring in each arm.

The primary 3-point MACE endpoint is shown in blue with each of its components in maroon. Both non-fatal stroke and non-fatal MI have upper bounds that are greater than 1. The superiority result for 3-point MACE is clearly due to the imbalance of events seen in CV death.

The results from the MACE components also translate to the related endpoints of all strokes, all MIs, and all-cause death. This plot shows the Cox model results for these related outcomes. Given that the ratio of patients experiencing an event was similar to the corresponding MACE component, results here are comparable to these related endpoints.

The results for CV death, which had a hazard ratio of 0.62 with an upper bound of 0.78, is also reflected in all-cause death with a hazard ratio of
0.68 and an upper bound of 0.82.

Since the death results were found to be the main factor for the MACE findings, we will be looking more closely at this. An inherent flaw in the prespecified primary analysis is that those who have a non-CV death are assumed to have a time until MACE equal to those who are alive and censored in the same treatment arm. For the EMPA-REG trial, this bias favored the placebo arm.

There were 154 non-CV deaths, of which 19 subjects had a non-fatal MACE, so there were 135 additional events when including all-cause death in the MACE endpoint, 51 on placebo and 84 on empagliflozin. This led to a hazard ratio of 0.85 with an upper bound of 0.97.

The Kaplan-Meier plots for CV death and all-cause death are seen here. We again see a separation of the curve starting out within the first few months of the study, like what was seen in 3-point MACE. The overall estimated incidence for CV death was 2.02 for the placebo arm and 1.24 for the pooled empagliflozin arm. When looking at
all-cause death, this went up to 2.86 for placebo and 1.94 for empagliflozin.

We examined follow-up in the trial for both MACE and death, which will be presented along with sensitivity analyses. There were approximately 2.9 percent of patients, which prematurely discontinued for MACE in the placebo arm, and 3.1 percent in the pooled empagliflozin arm.

Follow-up for death was more complete with 0.73 percent on the placebo arm prematurely discontinued, and 0.77 percent in the empagliflozin arm. Premature discontinuation did not seem to affect the overall picture of empagliflozin.

Sensitivity analyses for 3-point MACE were run using multiple imputation for the missing follow-up in those prematurely discontinued. The assumed event rate for the missing follow-up until December 15, 2014 was equal to the estimated incidence in the off-treatment patients of the empagliflozin and placebo groups.

Results from this analysis using the imputed and observed data were relatively unchanged from
the primary analysis results.

Sensitivity analyses using an extreme scenario favoring the placebo arm for CV death and death were also run. Results from these analyses showed an upper bound that remained below 1. These results are based on a single cardiovascular outcome trial, which was required to allow a 30 percent increase in cardiovascular risk. This trial was a success in ruling out the 30 percent increase for both the primary 3-point and secondary 4-point MACE endpoints.

Additionally, the trial showed superiority against placebo for 3-point MACE only. Currently, there is no precedent for when these types of safety studies show superiority. The superiority of the 3-point MACE is due to the differences seen between the treatment arms and CV death. When examining the amounts of premature discontinuation for this, we did not find that it affected the results.

Thank you. And I will turn this back to Dr. Lungu.
DR. LUNGU: Thank you, Dr. Clark. I will now discuss some additional issues related to the primary endpoint finding. While the results of the primary analysis demonstrated a statistically significant effect of empagliflozin on 3-point MACE, there are important additional considerations, which were needed to be weighed when deciding whether this trial provides the level of evidence necessary to form the basis for a new claim.

I covered the conduct of the trial in my previous presentation, and I am now going to cover some issues that are specifically related to the interpretation of the results. These issues attempt to address the following two questions. Did the trial win on the primary endpoint, which is arguably the gatekeeper for looking at cardiovascular death, a component of the primary endpoint? And does the trial provide substantial evidence of effectiveness in reducing cardiovascular mortality?

Perhaps the most obvious consideration is
that the EMPA-REG study was a single study, and as you have heard this morning, two studies are generally needed to form the basis for a new claim. The EMPA-REG study was also primarily designed to rule out excess cardiovascular risk and not specifically to establish a benefit. In that sense, it is a new paradigm and the persuasiveness of the evidence generated by such a trial could be affected.

It may also be relevant to consider that the trial was primarily focused on evaluating a specific type of cardiovascular risk. The p-value for superiority regarding the 3-point MACE endpoint was 0.04. As you have heard this morning from Dr. Guettier, the level of evidence for a single trial needs to be very statistically persuasive. You will need to consider whether the statistical results represent a very statistically persuasive finding and whether their persuasiveness is sufficient to form the basis for a new claim.

Dr. Clark has reviewed the impact that missing data had on the trial's statistical
significance. There may be other clinical issues, such as handling of the silent MIs, which could also impact your level of confidence in the primary results. This will be discussed in a later part of my presentation.

As you have also heard, the MACE result appears entirely driven by the effect of empagliflozin on cardiovascular mortality, with little effect on non-fatal stroke or MI. This suggests the benefit of empagliflozin may not reflect atherosclerotic cardiovascular risk reduction, the main risk evaluated in this trial.

In the next few slides, I will focus on two specific clinical issues, which may impact your level of confidence in the results. One is the potential lack of complete ascertainment for some myocardial infarction events, namely silent MI events. The second is an absence of information to confirm that the large proportion of cardiovascular deaths were indeed cardiovascular deaths.

I will now discuss the myocardial infarction endpoint focusing on the issue of the silent
myocardial infarction. Like clinical MI, a true clinical event of silent MI is associated with poor prognosis, and is a clinically important morbid event. Diabetic patients may be particularly prone to these events as they may not present with typical clinical symptoms associated with MI due to their underlying disease.

Because this event is an asymptomatic event, it is a challenging event to capture, fully ascertain, and analyze in a cardiovascular outcomes trial. In terms of precedent, some applicants have included these events in their primary endpoint while others have not. To some extent, this has depended on the types of procedures in place to prospectively capture these events.

The EMPA-REG study design suggests that the applicant, at least initially, intended to collect silent MI events. With regard to trial planning, silent MI was an event initially defined in the clinical trial protocol and ECGs were collected routinely at prespecified intervals throughout the trial. In addition, a central ECG vendor was
retained to analyze ECG changes and flag concerning changes to investigators.

Moreover, the case report forms included silent MI as an outcome event to be sent for adjudication. The adjudication committee was to evaluate events, which could be consistent with the primary endpoint, and silent MI was a trigger event that should have been evaluated.

In the nine versions of the CEC charter reviewed, we did not identify a specific definition for silent MI. Potential outcome events coded to the preferred term of silent MI, and referred to adjudication, had to fit one of the MI definitions in the charter to be counted as a primary endpoint event.

Silent MI was designated as a secondary endpoint in the protocol, but it is unclear whether these events were to be included in the primary endpoint, as these were not explicitly excluded from the initial versions of the protocol.

In version 4 of the protocol, which was implemented in 2011, the applicant explicitly
specified that silent MI was not to be counted as part of the non-fatal MI component in MACE. This occurred prior to the interim analysis.

The final statistical analysis plan, which was implemented in 2015, completely redefined silent MI and changed it from a clinical event to an event purely based on ECG criteria. As I have mentioned, the CEC was to adjudicate investigator-reported silent MIs, but the provided definitions in the CEC charter did not include a specific definition for silent MI. The CEC adjudication committee asked the applicant for clarification with regard to silent MI, and it is not clear that a response was ever provided.

As I have said, the final event definition for silent MI used in analysis was implemented in a 2015 change to the statistical analysis plan. This was to be used in secondary analysis only, and is based on ECG criteria. I would like to stress that this definition of events did not include any input from investigators or adjudicators, and is subject to these very important limitations.
Only about half of the analysis population had data that could be evaluated for the occurrence of silent MI using the applicant's definition. This was because of ECG abnormalities at baseline, absence of post-baseline ECG evaluations, or occurrence of intervening ECG changes unrelated to silent MI event in some participants. This further limits the conclusion that can be drawn from this endpoint.

We did explore the impact of inclusion of these events as defined by the applicant in the primary endpoint, and I will show you this on the next slide. Of the 7,020 subjects in the analysis population, only 3,589 had data that could be assessed for a silent MI based on the applicant's final definition. Using this definition, 15 events were identified in placebo and 38 in the empagliflozin group. The incidence of silent MI was slightly higher in the empagliflozin group at 1.6 percent compared to 1.2 percent in placebo.

Acknowledging that there are major limitations to these data, inclusion of the silent
MI in the primary endpoint data leads to rejection of superiority for 3-point MACE. The division concludes that there is missing information on clinically meaningful silent myocardial infarction events in the EMPA-REG OUTCOME study.

We would like you to opine on whether the absence of reliable information on silent MI is important in your overall assessment of the persuasiveness of the primary results.

The second topic that may impact the persuasiveness of the results is related to the cardiovascular death findings, and specifically related to the large proportion of cardiovascular deaths, which were deemed non-assessable by adjudicators.

In the assessment of death, adjudicators were tasked with assigning the cause of death to either a cardiovascular or a non-cardiovascular cause. Deaths assessed as cardiovascular death were further subcategorized as shown in the table. As Dr. Clark discussed, fewer deaths were observed in the empagliflozin group. The difference was
primarily due to deaths categorized as cardiovascular death.

The largest subcategory for cardiovascular death was non-assessable death. This subcategory made up 40 percent of events, and was followed only by sudden cardiovascular death and death due to heart failure or cardiogenic shock as the most frequent causes of CV death.

Given the proportion of events that were reported as non-assessable, it is unclear whether all events were truly cardiovascular death events, or what potential mechanism is responsible for the observed reduction in cardiovascular death.

In cardiovascular outcome trials, it is generally acceptable to categorize non-assessable death as presumed to be cardiovascular death, however the expectation is that this category would represent only a small proportion of these events. This is not the case in the EMPA-REG OUTCOME study.

Given the relatively large proportion of non-assessable deaths in the EMPA-REG study, we examined the type of information available to
adjudicators on these events. Less than half of the patients had a death certificate or proof of death available, and none had autopsy. Although the preferred terms reported by investigators for these patients were suggestive of cardiovascular death in the majority of cases, there is really no information to confirm or refute this from the trial documents.

Given the uncertainty around these deaths, we analyzed the impact of excluding non-assessable death on 3-point MACE and cardiovascular death. Excluding non-assessable death from the primary endpoint changes the results from superior to no longer superior.

The hazard ratio for 3-point MACE excluding these deaths is 0.9 with a 95 confidence interval of 0.77 to 1.06. Excluding non-assessable death from analysis of cardiovascular death would not alter conclusions on cardiovascular death where the hazard ratio not including these events is 0.59 with a 95 percent confidence interval of 0.44 to 0.79.
We would like you to opine on how these events impact your interpretation of the persuasiveness of the results.

I will next discuss the exploratory endpoints of heart failure and renal endpoints. I will start with heart failure. As a reminder, the hospitalization for heart failure definition was liberalized over the course of the trial to allow for the addition of oral diuretic, as well as short-term stays in chest pain observation units.

This made the definition less specific and potentially allows for inclusion of episodes that were either milder forms of heart failure or not heart failure at all. The definition used in analysis is less specific and represents less severe events than the current standardized definition.

Analysis for hospitalization for heart failure events were purely exploratory in this trial, and type 1 error was not controlled. The direction of change in the results of analysis suggest that treatment with empagliflozin could be
associated with a decrease in risk for hospitalization for heart failure or hospitalization for heart failure and death due to heart failure.

In addition to the exploratory nature of these analyses, the agency has other reservations with regard to the adequacy of this study to definitively establish a benefit of empagliflozin on heart failure and heart failure outcomes.

One reservation is that the study was not designed with the specific intent of demonstrating an effect on heart failure. Only 10 percent of subjects had a reported history of heart failure at baseline.

For subjects with a history of heart failure, information on the type of heart failure and the severity of heart failure was not collected. For example, there are no data on ejection fraction or New York Heart Association functional classification for these patients. Furthermore, it is not clear how well heart failure was managed at baseline and throughout the study.
While a majority of patients were receiving renin-angiotensin system antagonist and beta blockers, there are no data to evaluate whether the dose of these drugs and other drugs used in the chronic management of heart failure were optimized. While the heart failure findings are interesting, the division believes that data on heart failure are exploratory and should be confirmed in studies designed to specifically assess this outcome.

I will now discuss the results based on renal-related endpoints and the issues that affect the interpretability of the results. The applicant proposed changes to the prescribing information to reflect the results of these analyses.

Analyses based on renal endpoints were exploratory in the EMPA-REG study. There was no plan to control for type 1 error rate across these analyses in any version of the protocol or statistical analysis plan.

The definitions for various renal endpoints were changed significantly throughout the trial.

Specifics of renal endpoints to be used in the
final analysis were defined late in the trial in
the final statistical analysis plan submitted after
the interim analysis, and after the trial had
ended.

In addition, the clinical trial protocol or
statistical analysis plan did not specify processes
for identifying or confirming potential renal
events. These endpoints were identified based on
investigator-reported adverse events or in
laboratory findings, and were not adjudicated.

I will next discuss the results and
interpretation of the renal endpoints analysis.
Exploratory renal endpoints used for the final
analysis included new onset albuminuria, new onset
macroalbuminuria, and the composite endpoint new or
worsening nephropathy, which included new onset
macroalbuminuria, doubling of creatinine with eGFR
less than 45, initiation of continuous renal
replacement therapy, and death due to renal
disease.

I will first discuss the albuminuria related
endpoints, followed by the composite of new or
worsening nephropathy. Albuminuria was assessed using a single spot urine albumin to creatinine ratio measured by a central laboratory. The analysis for new onset albuminuria included subjects without albuminuria at baseline.

Similarly, the analysis for new onset macroalbuminuria included subjects without macroalbuminuria at baseline. No notable difference was observed between treatment groups for new onset albuminuria. New onset macroalbuminuria was more frequent in the placebo group.

Because empagliflozin causes intravascular volume contraction, we explored the potential impact of hemodynamic effects on urine albumin excretion. I will show this in the next slides. This slide shows changes in spot urine albumin to creatinine ratio over the course of the trial. It suggests lower albuminuria in the empagliflozin group.

The changes in glomerular filtration ratio and systolic blood pressure suggests an acute
hemodynamic effect with empagliflozin, as shown by
the acute drop in eGFR shown on the left, and the
acute decrease in systolic blood pressure shown on
the right. The changes in eGFR and systolic blood
pressure suggests that changes in albuminuria could
be driven by these factors.

The effect on albuminuria reflects a
pharmacodynamic effect. This slide shows urine
albumin to creatinine ratio at baseline, last value
on-treatment, and after a 30-day wash-out. The
effect of empagliflozin on albumin to creatinine
ratio appears to go away with wash-out, as
illustrated by the increase in urine albumin to
creatinine ratio from the last value on-treatment
to 30-day follow-up. This suggests that the
underlying renal disease is not changed.

The applicant used a composite endpoint
consisting of new onset macroalbuminuria, doubling
of serum creatinine with eGFR less than 45,
initiation of continuous renal replacement therapy,
and death due to renal disease, to define an event
of new or worsening nephropathy. Using this
definition, fewer patients in the empagliflozin
group seemed to experience this endpoint. I would
note that the component that contributes the most
to this difference is new onset macroalbuminuria.

As I mentioned previously, this may be in
great part attributable to the hemodynamic effect
with empagliflozin. The other components
contributed a small number of events, and 2 out of
4 components had too few events to draw any
meaningful conclusions.

I will review my concerns regarding the
components of doubling of serum creatinine and
renal replacement therapy over the next few slides.
I will not be discussing death due to renal disease
as there were only three events reported in the
trial.

The doubling of serum creatinine was used as
a marker of diabetic kidney disease progression.
This criterion required only a single post-baseline
serum creatinine value more than twice baseline,
with an eGFR less than 45 on the same date without
confirmation that the decline in renal function
persisted after a specified time period.

As a result, the endpoint might capture both acute reversible changes in renal function, such as acute kidney injury, and chronic irreversible changes in renal function, such as developmental progression of chronic kidney disease.

To explore this issue, we look for a confirmatory creatinine value more than twice baseline and eGFR less than 45 at any time in the 30 days following an initial event. The decline was confirmed in fewer than half of subjects with an event, suggesting that most of the initial events may have been cases of acute kidney injury. This endpoint does not appear to be specific in capturing events of nephropathy as defined by the applicant.

I will now discuss the evaluation of events classified by the applicant as continuous renal replacement therapy. The definition of this endpoint was not clearly specified, and we reviewed the narratives and CRFs for a random sample of events to gain additional insight. Most of the
cases reviewed represented events of acute kidney injury requiring temporary dialysis, and none of them represented chronic dialysis that could be attributed to progression of underlying kidney disease.

This suggests that in EMPA-REG, this endpoint was not able to differentiate between acute reversible need for renal replacement therapy and end-stage renal disease. While acute kidney injury events requiring dialysis are clinically significant, they do not necessarily represent events attributable to diabetic nephropathy disease progression.

In conclusion, renal endpoints in EMPA-REG were exploratory. There was no control of type 1 error. The endpoints selected differ from those typically used to establish efficacy of drugs to treat diabetic nephropathy or to assess effects on irreversible loss of renal function.

The endpoints were redefined during the trial and processes to identify and confirm renal events were not defined. The endpoints captured
effects on albuminuria, which appear to be a reversible hemodynamic effect, and may not predict treatment effects on renal outcomes.

The agency has not accepted on-treatment effects on albuminuria as a surrogate for clinical outcomes in diabetic nephropathy, in part because therapies can have acute reversible pharmacologic effects on albuminuria that may differ from the long-term effects on the irreversible loss of renal function and underlying disease progression.

Having completed my discussion of exploratory endpoints, I will turn next to highlight some differences that we noticed between the treatment arms. The study was designed to compare empagliflozin to placebo in addition to local standard of care. The expectation is that this would minimize confounding in the interpretability of the results.

Glycemic control was different between the treatment arms. While an early difference in glycemic control was expected, given that one group was randomized on active drug while the other group
was not, it was expected that adjustment of co-
administered anti-diabetic therapy would minimize
differences as the trial progressed. However, as
seen on this slide, the difference persisted over
the entire course of the trial.

The reason for this difference is unclear.
Anti-diabetic therapies appear to have been
intensified more in the placebo group, yet these
changes did not result in minimizing glycemic
control differences. Background therapies were
increased more, and new anti-diabetics were started
more frequently in the placebo-treated subjects.
Insulin DPP-4 inhibitors and sulfonylureas were the
most frequently used concomitant anti-diabetics.

As observed in the original submission,
small dose-dependent increases in mean cholesterol
were seen with empagliflozin. In addition, small
increases in hemoglobin and hematocrit were
observed and more patients shifted from normal
hemoglobin or hematocrit to values above the upper
limit of normal.

As expected, based on blood pressure
A numerical imbalance was also seen with the other two members of the class at the time of the approval and appears inconsistent with the observed findings of blood pressure lowering with this agent.

The adjudication of stroke in the EMPA-REG study was based on all available data. Standardized assessments for stroke events, such as
clinical assessment or specific imaging, were not required or specified in the protocol or the CEC charter. The CEC charter outlined four criteria for identification of stroke.

There needed to be a rapid onset of a focal or global neurological deficit. The deficit needed to last for more than 24 hours, unless it met the criteria shown in parenthesis. The deficit could not be attributed to a non-stroke cause, such as brain tumor or head trauma. Finally, the diagnosis needed to be confirmed either by a specialist by imaging or by lumbar puncture.

Two hundred and thirty-three strokes were observed with 210 of these reported as non-fatal strokes. More strokes occurred in the empagliflozin group, and the hazard ratio for stroke, including fatal and non-fatal stroke, and for non-fatal stroke alone, was about 1.

These findings raise question with regards to the role of empagliflozin on reduction of atherosclerotic cardiovascular disease progression and cannot definitively exclude the possibility
that the drug could cause a small increase in the
risk of stroke events in certain individuals.

This is somewhat of an unexpected finding in
light of the blood pressure differences observed in
the trial. Both systolic and diastolic blood
pressure were decreased in the empagliflozin group
compared to placebo, as shown on this slide. This
might be expected to reduce one's risk of a stroke
event.

The final topic that I will cover are the
non-cardiovascular safety findings from the
EMPA-REG OUTCOME trial. Specifically, I will
discuss the findings for events designated as
adverse events of special interest, and that are
not already in the product label.

Analysis of non-cardiovascular safety was
based on investigator-reported adverse events,
review of laboratory data, and safety endpoints
predefined by the applicant. The adverse reactions
listed in the right-hand column are serious drug-
related adverse reactions already featured in the
product label. My talk will not focus on these.
I will spend the next few minutes reviewing the findings related to fractures, hepatic injury, malignancies, and venous embolic and thromboembolic events in the EMPA-REG study.

Overall, fracture events were balanced between the treatment groups. However, an increased risk in upper extremity fractures not favoring empagliflozin was seen. A similar observation was made in another SGLT2 inhibitor program. While bone density was not assessed as part of the study, the term osteoporosis was reported more commonly in the empagliflozin group.

A signal for hepatic injury was identified in the development program, and was to be followed in the ongoing EMPA-REG study. A hepatic adjudication committee was tasked with determining whether events of hepatic injury were causally related to study drug and could represent events of drug-induced liver injury.

Events referred for adjudication included serious hepatic events and certain events based on laboratory test profiles. More events were
referred for adjudication with empagliflozin, though nearly all events were adjudicated as unlikely related to study drug.

A profile consistent with biochemical Hy's law was reported more commonly with empagliflozin than with placebo. However, alternative etiologies more likely than drug-induced liver injury were identified for these events. Overall, the findings do not suggest that empagliflozin carries a heightened risk for drug-induced liver injury.

The overall incidence of malignancy was similar between treatment groups. The applicant was asked to look at some selected types of cancers as events of interest, which were in balance at the interim assessment. The number of events for each of these is presented here.

Note the 2 to 1 randomization. The incidence for these malignancies of interest were generally balanced between treatment groups. The one cancer with a suggestion of increased risk was pancreatic cancer. However, the total number of events is small and confounding risk factors, such
as obesity, diabetes, or smoking were identified in review of these cases.

While no imbalance for venous embolic and thromboembolic events has been seen with empagliflozin in the original NDA review, the observed increase in hemoglobin and hematocrit in the development program raised some concerns. The assessment for these events was based on a review of investigator-reported adverse events and excluded stroke events. No increased risk of venous embolic and thromboembolic events was seen.

In summary, the EMPA-REG OUTCOME trial suggests that empagliflozin is both non-inferior and superior to placebo on the 3-point MACE endpoint. The results appear to be entirely driven by the cardiovascular death component. Empagliflozin was found to be non-inferior but not superior for the 4-point MACE. The review of the non-cardiovascular safety was generally reassuring.

There are certain important factors that affect our interpretability of the study results and the persuasiveness of the evidence generated by
the study. First, it is unclear whether this single study is sufficient to support a new efficacy claim. The EMPA-REG study was designed primarily as a safety trial, and the p-value for superiority for 3-point MACE was 0.04.

The handling of silent MI in this trial was not optimal, and we would like to hear your opinion on this aspect of the conduct as it relates to your overall assessment of the evidence.

Analysis of 4-point MACE and the individual components of the 3-point MACE do not provide persuasive evidence that the effect of empagliflozin is attributable to atherosclerotic cardiovascular disease risk reduction.

The results of the study hinge on the finding of cardiovascular death. While we recognize that this is perhaps the most important of all the endpoints in the composite, the high proportion of non-assessable deaths in EMPA-REG study may impact your interpretation of the evidence.

With the limited information on these
deaths, it is difficult to know for certain whether they were truly cardiovascular deaths, and what mechanism might be driving the CV death result.

Thank you, and we look forward to your discussion around these issues during this afternoon's session.

Clarifying Questions to FDA

DR. SMITH: Thank you. What we're going to do now is have again clarifying questions. We're coming to discussion, but this is not the discussion time. This is really for clarification. We'll focus first directing questions to the FDA, and then as time allows, and then later this afternoon, we'll come back to our remaining clarifying questions for the sponsor.

Dr. Hiatt, I think you had a question?

DR. HIATT: Yes. We didn't get a chance to ask this in the sponsor's presentation, but you did raise an issue about silent MI, so let me pursue that a little bit further. In preparing for the meeting, I sort of challenged myself to see how that's handled in other cardiovascular outcome
trials.

In looking at a number of recent trials who were testing more anti-thrombotic therapies, it's typically included in the definition of myocardial infarction. And the CDISC definition that's most commonly used would not include it as an acute MI, but a significant new Q-wave that occurs after randomization would be considered to be a new MI of clinical significance that would be adjudicated by the CEC as a myocardial infarction.

So the question then comes up, how do these events stack up? And if you look at again trials where these are triggered as just part of the clinical review and not sought after as a specific case report form in a core lab, they're typically few in number. They carry clinical significance. And they typically track the primary endpoint's response to therapy. But if not ascertained properly, they can just add noise.

So the question here about ascertainment bias is clearly relevant, and I think in this trial went a little bit further in terms of the number of
the events that were captured. It was a little higher than one would expect. So you take that kind of information, and look at the MACE primary, and then ask, well, had that stayed in the definition, it changes the significance of the MACE primary and makes it no longer significant, which would kind of stop your thinking about superiority.

How do you view that? I mean, you put it out as a question, but it seemed that the charter was changed, and I didn't see if the FDA commented on that. It wasn't clear why the CEC did that. And so I'm just wondering what your thinking is about how that changed the approvability around the primary endpoint.

DR. GUETTIER: So as Dr. Lungu stated, the silent MI were a topic of confusion in the actual review of this application. And the final definition that was used for analysis of silent MI was purely ECG based, and that was something that we had to clarify with the applicant.

What seemed to have happened in the trial is that there was no real specification in the
protocol, other than for a secondary endpoint, that silent MI would be looked at. Again, we don't really have, as Dr. Lungu stated -- it was never a topic of discussion between the division and us with regards to inclusion of silent MI in the primary endpoint, but it wasn't specifically or explicitly excluded from the endpoint. The protocol itself, as Dr. Lungu stated, suggested that actually the applicant collected information that could be used to analyze silent MI.

What was not clear to us is whether the CEC and the CEC charter ever had knew what to do with these events. They received some silent MI events by investigators because the case report form actually had a checkbox for silent MI. If the checkbox was checked, it would make it to an adjudication package, and an adjudication package would be sent to the CEC members.

There was some confusion as to what to do with these events based on our review of the record, the CEC record. And again, it's not really clear how that was resolved, but at some point,
before the interim, the applicant specifically excluded silent MI from the primary endpoint.

Again, the CEC members and the applicant can best speak to this, had very specific event that they had to check off after they were done with their adjudication.

Silent MI was not one of those events. If they had received a silent MI as the outcome term, they had to transform that into one of the MI definitions, and that happened for some events. And we asked that, for the applicant, the applicant have these data. So there were some silent MI events that were captured and sent to adjudication and were reviewed.

Again, it wasn't really ever a point of discussion between us and the applicant in the negotiation for this trial. And the trial was complex and we were dealing with other things, including how to --

DR. HIATT: Just to briefly clarify. If you're going to run a cardiovascular outcome trial and you present the charter to the FDA, the sponsor
can decide whether silent MI is in the definition or not. Because I've looked at other trials. It's kind of variable. It's not 100 percent. And in this case, it sounds like it wasn't a review issue. Is that correct?

DR. GUETTIER: I mean, that's a topic for discussion that we'd like your input on, whether or not these types of events should be included in all cardiovascular safety trial, and that's why we pose the question to you today.

Again, we know, at least as a division -- and we have our cardiology colleagues here as well who can answer -- that they're difficult to assess because some people have baseline ECG abnormalities and so you're going to get missing data in those patients. Unless the trial pre-specifies and has specific procedures in place to collect ECGs, then you might just get random events.

Then the CEC charter has to have a definition or else the CEC members are at a loss and will code events to something else if they can
find something or will just throw out these events. That's maybe something the applicant can clarify.

DR. SMITH: So, Dr. Hiatt, did you -- to expand on your question, it would be helpful for you to ask for any more clarification at this point from the sponsor or should we move on with the FDA clarifications?

DR. HIATT: I think that answers it. It's a discussion point for later, but I'd rather not prolong that now.

DR. SMITH: Okay. Dr. Heckbert?

DR. HECKBERT: Yes, Susan Heckbert. My question was really on the same topic, on silent MIs. And just as a point of clarification, on slide 7 and 8 of Dr. Lungu's presentation, where she's talking about silent MI, she says that in the data that were available on silent MI, which was not complete and was perhaps suffering from a variety of different limitations, you indicated that only about half of patients could be analyzed for silent MI due to baseline ECG abnormalities, absence of post-baseline evaluation, or intervening
ECG changes.

I just wondered, with the data that were available, two questions. Were they read finally by a central ECG reading center? There was a mention that one was retained early on, but I wasn't sure whether these -- what data we do have were read by a central center, whether that was subject to individual people at each center.

Then the second question was, what proportion of the -- so half the patients had ECG data and half of them didn't. What proportion of those was it missing because of an absence of post-baseline evaluation versus pre-existing ECG abnormalities or intervening ECG changes?

DR. LUNGU: To answer the first question, they were read by the central ECG vendor. For the second question, I think we have this information in a slide. Let me look and I'll get back to you.

DR. HICKS: Good morning. My name is Karen Hicks. I'm a medical officer in the Division of Cardiovascular and Renal Products, and I'm the cardiology consultant to the Division of Metabolism
and Endocrinology Products on the review of this application.

Dr. Hiatt, you're correct. There are arguments for and against including silent MI -- can you hear me now? There are arguments for and against including silent MI in overall adjudication of non-fatal MIs in cardiovascular outcome trials. There is thought to be a prognostic significance, especially in diabetic patients. And for the most part, silent MIs may comprise 9 to 37 percent of non-fatal MIs in a particular clinical trial, and it's thought that there is a worse overall prognosis.

LaToya, if I could please go to background slide number 8 to further address the ascertainment issues. So, EMPA-REG used an algorithm for silent MI that likely did not identify all potential events. They excluded patients with right bundles, left bundles, look, there are ways to evaluate Q-wave myocardial infarctions even in those settings.

I reviewed some of the silent MI data,
didn't have to go very far to figure out that the
time-to-event data were completely unreliable. We
were told, and we queried this many, many times,
and we were told that silent MI was not adjudicated
in this trial. And that's not completely true
because just last week, we found out that some
silent MI was adjudicated.

So here's what happened in the trial, as I
understand it, and perhaps the applicant can
clarify things a little bit further. So silent MI
was included as a definition in every single
protocol. And in the CEC charters, silent MI was
not explicitly excluded. And in fact, it was
listed as one of the trigger terms to collect
events. So when they were scouring through the
safety database, they were picking up these events.

My understanding, based on the response we
received just last week, is that the silent MIs
that were being identified by the investigators as
a trigger term -- that those events went to the CEC
for adjudication.

In most cases, they were adjudicated as MIs
due to PCI, MIs due to CABG, some of them weren't adjudicated as events at all. Some were thought to be a heart failure event. And my understanding is -- but we should check with the applicant -- that those MIs were included in the overall primary endpoint results.

The other way that silent MIs were identified was through the electrocardiographic core laboratory. And what should have happened in this trial but didn't, is that if the core lab had identified a Q-wave MI, that an adjudication package should have been put together, everything should have gone to the CEC for evaluation, and then things would be good.

But what happened was the EKG core lab, all they did was send the EKG back to the investigator for sign off, and I don't think that investigator had all the 12-lead electrocardiograms in front of him or her. And so that's why the time to event data were not reliable.

So in summary, the trial used an algorithm that didn't likely capture all potential events.
The time-to-event data were unreliable. There was reportedly no oversight by the CEC of these events. Again, I think prior to study initiation, there really needs to be a discussion is silent MI in or is it not in, so that if it is in, then it needs to be adjudicated.

CECs do things a little bit differently, but some CECs, if it's identified by a core lab, will have at least one CEC member adjudicate it. If it's something that's identified by the investigator, they'll have two members of the CEC evaluate it. And lastly, a number of patients lacked baseline 12-lead electrocardiograms.

So in summary, it's unlikely that any four of these factors led to differential ascertainment. Silent MI was really not well handled at all. And I know at least some of the silent MI findings are unreliable. It's not like we'd want any of that in the primary endpoint anyway. And I hope that answers everyone's questions.

DR. HIATT: So just to understand, all the points you raised were pretty clear in reading the
material, so the lean goes in the wrong direction. Numerically, it's unfavorable. But it sounds like your interpretation is that's most likely a random result because if you include those kinds of events, it's most likely random events, so it can go either way. It doesn't really mean anything.

DR. HICKS: Yes. I have to say that we were concerned about the exclusion of silent MI from the primary endpoint in amendment 3. The timing of all of that and the timing of other things that happened in the trial after the CEC became involved -- any time we see changes, there are red flags that go up. And unless things are really well documented about why there are changes, we have to be concerned about other things. Thank you.

DR. SMITH: So would the sponsor want to make any additional -- I'm really just looking for clarification, not just carrying on.

DR. WOERLE: Hans-Juergen Woerle. We have to say, we added to the confusion around the topic silent MI with the following thing we did. We
captured ECG pathological Q-waves, single ECG pathological Q-waves and flagged those as silent MI and reported those cases as silent MI.

In hindsight, that term should not have been used because these were single ECG measurements where we have no confirmation. We have very little information on the clinical relevance of these findings. But we added this to the statistical protocol as a term, the single ECG finding as a silent MI.

That was a secondary endpoint to clarify that what we had defined in the protocol is actually not part of the primary endpoint because it's not an adjudicated event. It's a single ECG finding. We clarified this in a protocol amendment prior to the interim analysis. But it was very clear, when you look at the adjudication charter, that there was no qualification for such events to be made to the CEC member.

Now, Jim Januzzi adjudicated these events, and he can add some clarification, which may help to shed some light on this.
DR. JANUZZI: Thank you very much for the opportunity to speak on this. This is obviously an important topic, as Dr. Lungu pointed out in a very nice presentation. Silent myocardial infarction is an important topic. Dr. Hiatt, you made this clear as well. And as Dr. Hicks said, there are substantial challenges to recognition and actually how to adjudicate these in clinical trials.

So while it is clinically relevant, it's challenging to recognize and adjudicate. And relevant to this, the current guidance from ACC/AHA, Dr. Hicks was an author, is actually well silent to the definition of silent MI in clinical trials' definitions.

When considered as an endpoint in clinical trials, silent MI, the standard approach has traditionally relied, as has been said, on a baseline electrocardiogram, then showing a change, which is then stable on subsequent measurements. So it requires serial ECG measurement in order to demonstrate stability of the finding.

This in part has to do with the fact that
the changes consistent with MI are evanescent. They may be non-specific and may not be stable over time. Therefore, single tracing should not, by any means, be used to make the conclusive diagnosis of a myocardial infarction in the absence of clinical signs, symptoms or other data. As a consequence, most CVOTs do not include silent MI.

May I have slide CV-243, please? That would be slide 1 up. Just an example of the current cardiovascular outcomes trials looking in diabetes since 2008 as well as some of the older, just illustrating the heterogeneity with respect to inclusion or not of silent myocardial infarction.

Now, relevant to EMPA-REG -- and this slide can be taken down, thank you. Relevant to EMPA-REG OUTCOME the term silent MI, as Dr. Woerle just said, is somewhat inaccurate, actually, and somewhat of an unfortunate term that was used.

From the CEC's point of view -- and I'd emphasize that if an investigator felt a patient had suffered a myocardial infarction, silent or otherwise, they could send the information to the
CEC for adjudication. So, in point of fact, if triggered, MI would have been evaluated for on our level and the CEC, so silent MI technically was not excluded from consideration.

What was triggered here were electrocardiograms that were flagged as showing a possibly new abnormality. And in these 53 tracings that Dr. Lungu analyzed in her sensitivity analysis, of those 53 flagged electrocardiograms -- I choose that terminology rather than calling them silent MI -- the site was actually queried by the central ECG lab, core lab, reviewed by the investigator and their cardiovascular consultant.

Only 3 were sent to the clinical endpoints adjudication committee for review. Of those patients, only 1 was adjudicated as having hospitalized heart failure, and none were adjudicated as having a myocardial infarction.

One last point I think that would be helpful for reassurance -- may I have slide 1 up, please -- is relevant to the outcomes of the
patients that had these flagged ECGs, these 53 patients. We see, on the top line, the patients that suffered subsequent mortality and then adjudicated cardiovascular death. The flagged ECGs are in the right column. We see a substantially lower mortality rate overall compared to patients with adjudicated non-fatal MI using standard criteria.

Of the CV deaths, there were two, one was sudden cardiac death and the other was death due to pulmonary thromboembolism. So at least this provides some information on the CEC level at least relative to these specific patients, more specific data from the sponsor. I'll defer.

DR. SMITH: All right. Dr. Konstam, do you have a follow-up question on the same topic?

DR. KONSTAM: Yes, I do. So, Jim, just to kind of nail down a couple of things. Okay. The ones that you did adjudicate that were sent in from the investigators as silent MIs, what happened to those patients? Are they in the final results or not?
DR. JANUZZI: So they were as you saw in that last slide I showed. So yes, in fact, if a patient was sent in with a trigger term, silent MI, we would look at the event that the investigator was triggering. And as I said, and as Dr. Hicks indicated, in some cases they were MIs related to PCI, they were CABG MIs. So they were considered, sure.

DR. KONSTAM: So if you considered them MIs, they were counted as MIs.

DR. JANUZZI: Yes, exactly.

DR. KONSTAM: Okay. So this relates to my next question. So let me just say, I can go either way with inclusion or exclusion. What's more concerning is changes that occur from prespecified terms, and in practice during the course of the trial. So at some point in time -- we hear it was before the interim analysis -- the sponsor determined that silent MIs would not be included.

Now, I want to understand a few things about that. Why did you do that at that point in time? Who participated in that decision? And how did you
deal with the fact that they had already included silent MIs, as long as they came from the investigator?

DR. WOERLE: To clarify, we used an unfortunate term. What we wanted to clarify in the protocol that single ECG abnormalities, not being confirmed by the adjudication committee as an MI, should not be included. When you look -- and I would need Jim for confirmation here -- when you look at the CEC charter, the initial CEC charter did not have the term silent MI, but all pathologic clinically relevant MIs were included in the analysis. But please, Jim, would you comment?

DR. JANUZZI: I can keep it simple on the answer. Any case that was sent to us that looked like an MI was adjudicated as an MI, and was therefore included.

DR. KONSTAM: Even if the investigator indicated silent MI?

DR. JANUZZI: That's correct. And that speaks, to some extent, to the investigator at the sites accuracy for the event itself.
DR. KONSTAM: What is it that made you make that decision at that point in time? Was it based on what?

DR. WOERLE: Can we have the slide on the initial protocol? Slide 1 up, please? That's the definition of the initial protocol. It states primary event comprises adjudicated composite of CV death, non-fatal MI, and non-fatal stroke. To the adjudication committee, any potential trigger which would qualify for these endpoints had been sent.

Now, from this protocol amendment, from this protocol, it's not entirely clear whether silent MI is included or not. And type 4 and type 5 MIs, which can be considered as MI, are actually included. Now, to clarify this, we have to have a look at the CEC charter.

Slide 3 up, please. That's an extract of the CEC charter. And here in the CEC charter, the adjudication committee has the opportunity to potentially trigger the following events, hard-outcome events. And as you can see, the term silent MI, what we defined as a single ECG
abnormality which was not confirmed by subsequent
ECG and additional information, which then had been
adjudicated as an MI, are not mentioned here.

To clarify this, we in the first protocol
amendment, made the clarification that silent MI
are excluded from the primary endpoint. It would
have been hindsight, much clearer to state --

DR. KONSTAM: But not silent MIs that were
identified by the investigator? Is that what you
said in the amendment? Did the amendment say
silent MIs would be excluded unless they are
identified by the investigator and sent into the
adjudicating committee?

DR. WOERLE: Dr. Broedl, would you like to
comment?

DR. BROEDL: Uli Broedl, Medicine. So to
address your question, what did we state in the
protocol, we unfortunately stated simply excluding
silent MI. But I want to make it very clear -- and
I would like to have the process slide -- how the
adjudication went. What Dr. Lungu nicely presented
was what happens if you add 53 flagged ECGs, which,
as Dr. Hicks pointed out, mixed with the primary endpoint.

What happened in terms of silent MI analysis is the following. Can I please have the silent MI assessment? If an investigator considered something like a silent MI as an adverse event, this adverse event had to be entered into remote data capturing system.

The vendor then assessed whether the term that was reported matches a trigger list of terms and be used on the SMQ, ischemic heart disease. We used cardiac failure, and we used Torsade de points. SMQ, the SMQ ischemic heart disease, included the trigger term, silent MI.

However, the CEC charter, which you can find in our briefing book, clearly specified for type 1 and type 2 MI, which were by far the most frequent MIs in this trial, that patients have to have symptoms, in addition to one of the following criteria, including biomarker, ECG changes, or imaging, and this is depicted in slide 1. Up, please.
So in essence, the vendor checked whether a trigger term was reported by the investigator. The trigger terms included, for instance angina pectoris, but also silent MI. The CEC panel, based on the criteria that you can find and that I alluded to, assessed whether this is enough evidence of a cardiac event in a clinical setting, but only had the opportunity to report those non-fatal cardiac events that are listed here, non-fatal MI, hospitalization for unstable angina, coronary revascularization, stent thrombosis, and heart failure hospitalization. The panel did not have the opportunity to specifically vote for silent MI.

DR. HICKS: May I clarify?

DR. SMITH: Yes.

DR. HICKS: So, in response to Dr. Januzzi's comments, I just wanted to clarify that we actually do have a definition for silent MI. It's called prior myocardial infarction. And the electrocardiographic criteria for prior MI are exactly the same as silent MI. And this is not
only in our data standards paper. It's also in our
draft definitions dated August 2014.

By the way, these definitions are in each
CEC charter. And there were a lot of charters and
a lot of changes in the charters that went on
during the course of the study. Thank you.

DR. SMITH: Thanks. I'm going to move us
along. We may return to this point later on.

Dr. Schambelan, you had a question?

DR. SCHAMBELAN: My question is long.

DR. SMITH: Dr. Neaton?

DR. NEATON: My question is simpler. So in
the FDA briefing document, on page 75, where
there's a description of some of the missing data,
there's reference to 74 subjects with positive
MACE, non-fatal events that could not be assessed
by the clinical events committee. Could you shed
some light on what you're referring to there and
the significance of that?

DR. GUETTIER: That's Dr. Hicks's --

DR. NEATON: I believe that was Dr. Hicks's
document, yes.
DR. HICKS: Yes, thank you. Actually, that was another information request we had made to the applicant, and they had prepared that data. So there were, in addition to the 200 or so patients who were missing MACE for a period of time, there were also approximately 70 patients who had potential MACE events. Not all of them, but some of those 70 had potential MACE events and Jennifer Clark has done a number of sensitivity analyses looking at that.

DR. NEATON: When you say a potential MACE event, there was some information, documentation provided that they were ruled out as a MACE event?

DR. HICKS: No. Actually, the CEC was unable to adjudicate these particular cases. So they were potential events, but because the data were incomplete, they couldn't finalize the adjudication diagnosis.

DR. SMITH: What I would like to as -- there are more questions. If there are any committee members that have questions that may lead to some preparation of some data that might be brought to
us later this afternoon, I'd like to prioritize that. Anybody have a question in that regard? Dr. Wilson?

DR. WILSON: So I believe in the product label, it says contraindication for an eGFR less than 45. And I'm not able, from the materials I've heard so far, to figure out what happened to those people and how many people were in that group in the trials.

There's a creatinine doubling and eGFR less than 45, but in fact the label says eGFR less than 45, you're not to prescribe the drug. So I think that would be worth knowing, that safety data, and whether it affects -- and what happened to those people? Did they stop using the drug?

DR. SMITH: Yes. Do you have the data now or would you like to pull that together for this afternoon?

DR. WOERLE: We can put this up for you. Let's remember, the trial was started, part of the phase 3 program, when it was not clear what limitation in terms of renal function we would see.
We specifically looked into patients who have eGFR less than 60 but also less than 45. And my colleague from drug safety, Gabriel Kim, looked into this and will walk you through what we saw.

DR. KIM: Gabriel Kim, Boehringer Ingelheim.

We have specifically analyzed patients with CKD stage 3b. Overall, in terms of the general safety findings, it was consistent what we observed for the overall population. In addition to this, we also looked at the MACE endpoint and cardiovascular endpoints, which also suggested a similar effect within this population.

Finally, in terms of a renal endpoint that we looked at, there was also shown a similar effect in terms of renal safety and endpoints.

DR. WILSON: Just as a follow-up, the people continued taking medication whether -- they weren't pulled out. They weren't censored? Just so I know.

DR. KIM: So, to answer your question, they were not instructed to discontinue medication once they have reached a CKD 3b eGFR category.
DR. SMITH: Thanks. We do have some more questions. I apologize to people who haven't been able to ask their questions. We will come back to those, but by the clock we really need to break for lunch at this point.

So, we're going to take a lunch break.

We're going to come back here, we're starting a little late, so we'll come back at 10 minutes past 1:00. And please take any personal belongings you may want with you at this time. Committee members, please remember there should be no discussion of the meeting during lunch, amongst yourselves, with the press, or with any member of the audience.

Thank you.

(Whereupon, at 12:20 p.m., a lunch recess was taken.)
A F T E R N O O N  S E S S I O N  

(1:22 p.m.)

Open Public Hearing

DR. SMITH: I encourage people to take their seats so we can get started here. We're still missing a few and we'll wait a minute or two. Okay. I think we'll start the afternoon part of today's work.

Both the Food and Drug Administration, the FDA, and the public believe in a transparent process for information gathering and decision making. To ensure such transparency at the open public hearing session, which is what we're going to have scheduled next, the FDA believes that it's important to understand the context of an individual's presentation.

For this reason, FDA encourages you, the open public hearing speaker, at the beginning of your written or oral statement, to advise the committee of any financial relationship that you may have with the sponsor, its product, and if known, its direct competitors. For example, this
financial information may include the sponsor's payment of your travel, lodging, or other expenses in connection with your attendance at the meeting.

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

The FDA and this committee place great importance in the open public hearing process. The insights and comments provided can help the agency and this committee in their consideration of the issues before them.

That said, in many instances and for many topics, there will be a variety of opinions. One of our goals today is for this open public hearing to be conducted in a fair and open way, where every participant is listened to carefully and treated with dignity and courtesy. Therefore, please speak only when recognized by the chairperson. Thanks
for your cooperation.

At this point, will speaker number 1 please step up to the podium and introduce yourself? Please state your name and any organization you are representing for the record.

DR. GRUNBERGER: Thank you, Mr. Chairman. It's always fun to be the first speaker after lunch. I'm George Grunberger, and I represent the American Association of Clinical Endocrinologists. The sponsor played no role in financing my trip here.

The American Association of Clinical Endocrinologists is the largest organization of clinical endocrinologists in the world, which comprises over 7,000 members from 97 countries. The mission statement of AACE is that it's a professional community of physicians specializing in endocrinology, diabetes, and metabolism, committed to enhancing the ability of its members to provide the highest quality of patient care.

I don't have to emphasize to this audience the burdens of type 2 diabetes, whether financial
or in terms of suffering, the complications. As we know, even today, it remains the leading cause of blindness in adults, leading cause of kidney failure, the leading cause of lower limb amputations. Relevant to this meeting, the patients with diabetes are much more likely to have hypertension, cardiovascular disease, stroke, and other complications.

We know that diabetes doubles the risk of vascular events, whether it's coronary heart disease, cerebrovascular disease, cardiovascular deaths. We know that diabetes confers the highest lifetime risk of coronary heart disease of any single risk factor. And we've seen that diabetes is associated with significant loss of life years, and the vascular deaths contribute the majority of this life lost.

We are here because the 2008 FDA guidance for diabetes drugs was designed to assure cardiovascular safety. It was not designed to demonstrate benefit. So all trials we have reported out to date show that safety until we have
EMPA-REG OUTCOME, which sort of is the reason we are here for today.

We've seen that in the primary outcome, the 3-point MACE, there was a significant 14 percent reduction of 3-point MACE, which was driven largely by the cardiovascular death with its 38 percent reduction. And also hospitalization for heart failure was significantly improved in patients on empagliflozin. To make it more interesting, only one of this 3-point MACE was significantly affected. It was a cardiovascular death with no significant improvement for patients on empagliflozin when it came to non-fatal MI or stroke.

So every day, after September 17, 2015, I've been asked three questions. Is this is a class effect? Should all patients with type 2 diabetes be put on SGLT2 inhibitor or empagliflozin specifically? And when will AACE, in view of a positive important outcome in our leader results, change its diabetes algorithm?

Well, let me start with the last one.
That's easy. AACE does not have to change anything. GLP-1 receptor agonists and SGLT2 inhibitors have already been the top two choices in our diabetes algorithm initially or right after metformin for the last three years.

This is a 2016 version of AACE comprehensive diabetes algorithm, which you can see whether for monotherapy, dual therapy, or triple therapy, GLP receptor agonist and SGLT2 inhibitors have been chosen as the two leading categories in trying to control glycemia in patient type 2 diabetes.

We do have obviously very strict recommendations when it comes to management of dyslipidemia and hypertension also for patients with type 2 diabetes because we believe every patient with this condition is at high and very high risk for cardiovascular complications.

If you look on the next slide and the 2016 diabetes algorithm, they already included when we talked about the profiles in diabetic medications a possible benefit for SGLT2 inhibitors for congestive heart failure, but we believe there is
no effect on progression of atherosclerotic cardiovascular disease. So we recognize possible outcomes from the EMPA-REG study.

So is this a class effect? Well, who knows? As you know, the canagliflozin, empagliflozin studies will not be out for the next couple of years. And they have recruited somewhat different patient populations, so it might be difficult to compare those results. Again, remember that the key inclusion criteria for EMPA-REG study included not only long-term uncontrolled type 2 diabetes, but established cardiovascular disease.

So should every patient with type 2 diabetes today be placed on empagliflozin or at least a SGLT2 inhibitor? We believe that would be premature because we have only one study and only one component of the primary endpoint, which was significantly positive.

So at this point we can recommend that only patients who fit the inclusion criteria of EMPA-REG OUTCOME should be considered, and for others let's wait for results of studies and real-world
postmarketing data in patients with type 2 diabetes who did not meet the specific inclusion criteria for EMPA-REG OUTCOME study.

So in closing, the American Association of Clinical Endocrinologists does not advocate for approval or indication for any specific drug. However, we recognize there is a great need for new drugs to help manage the ever-increasing burden of type 2 diabetes. And we certainly need more effective and safer medications to improve metabolic control of our patients without the risk of hypoglycemic weight gain, and now hopefully with cardiovascular benefit. Thank you.

DR. SMITH: Thank you. Will speaker number 2 now please step up to the podium and introduce yourself? Please state your name and any organization you may be representing for the record.

MS. GAO: Good afternoon. My name is Helen Gao, and I'm here today representing the diaTribe Foundation, a diabetes patient advocacy non-profit based out of San Francisco. Donors to
the diaTribe Foundation include the Helmsley Charitable Trust and many others, including today's sponsor. By way of disclosures, I also work for Close Concerns, a healthcare information company focused on diabetes and obesity. My colleague, Emily Reiger, will review Close Concerns' disclosures later this afternoon.

Today, I'd like to speak about an update indication reflecting a cardiovascular benefit for JARDIANCE in high-risk patients and what that could mean for patients and physicians. I was at the presentation of the full results from EMPA-REG OUTCOME at EASD last September, and vividly remember the air of excitement and wonder as a risk reduction for each endpoint was revealed to thunderous applause. It was truly a historic moment.

Coming out of the session, I was surprised and disappointed to see that the results weren't front page news in the mainstream press. I wondered how the average patient or the average primary care physician, who has so many demands
beyond diabetes, will learn of these unprecedented results.

An updated indication clearly stating that JARDIANCE can reduce mortality in high-risk patients through the reduction of cardiovascular death and heart failure would go far in publicizing with historic and clinically meaningful benefit. A 38 percent relative risk reduction in cardiovascular death is huge, and patients and healthcare professionals need to know about it.

Furthermore, many patients with type 2 diabetes might not even fully understand the link between diabetes and increased cardiovascular risk. If you put clear language about cardiovascular outcomes on the label of a diabetes drug, you increase the chances of a conversation about that risk between doctors and patients.

For the physicians and patients who are already aware of that cardiovascular risk but have not been able to mitigate it as trial after trial showed little impact on cardiovascular outcomes, JARDIANCE offers a new hope. It could mean a
longer, healthier life with less time spent in hospitals.

That's a big win for patients, and a big win for the morale of healthcare professionals. For too long, those in the diabetes field have felt, in the words of the musical Hamilton, outgunned, outmanned, outnumbered, and out-planned. I sincerely hope that the committee will vote today in favor of letting healthcare professionals and patients know that they have one more tool in the fight against diabetes. Thank you.

DR. SMITH: Thank you. Will speaker number 3 now please step up to the podium and introduce yourself? Please state your name and any organization you are representing for the record.

MS. REIGER: Good afternoon, and thank you for the opportunity to speak today. My name is Emily Reiger and I am here representing Close Concerns, a healthcare information company that aims to improve patient outcomes by making people smarter about diabetes and obesity. As far as disclosures go, today's sponsor is one of almost
300 organizations that subscribe to our fee-based newsletter, Closer Look.

Today, I'd like to take a step back and offer a few big picture thoughts on what JARDIANCE and especially the benefits demonstrated in EMPA-REG OUTCOME mean for patients with diabetes and their healthcare providers.

Thanks to the FDA's recommendation that these major outcomes trials be done, we now have valuable information on JARDIANCE that can help healthcare providers as they choose among the many different options for patients with type 2 diabetes. We know this trial required a substantial investment of time and money, and we want to ensure that the diabetes community can gain as much value as possible from that investment.

People with diabetes want to live long and they want to live well. Few people expected that findings of lower cardiovascular risk would emerge so clearly in this trial, even if for only a small group of patients. Given the staggering prevalence and cost of diabetes and its complications, a
product that reduces the risk of a negative outcome for even a small percentage of patients could have a significant impact on the healthcare system.

We understand that any label update will need to include statements about the limitations of the data and how much it can be generalized. And we'll leave it to the experts to debate the more technical questions related to trial design.

However, we think it's important for today's busy healthcare providers to have the opportunity speak broadly with patients at high risk of cardiovascular disease about the potential benefits of this medication.

We would love to see increased awareness among people with diabetes about the risks of cardiovascular disease, and we think that including this data on the label can be an important component of that effort. So thank you again for the opportunity to speak and for your consideration of these matters.

Clarifying Questions (continued)

DR. SMITH: Thank you. The open public
hearing portion of this meeting has now concluded and we will no longer take comments from the audience. And what we're going to do first is to resume the opportunity to ask any more clarifying questions from the committee. And before that, the sponsor has asked if they could briefly provide some additional clarifying points on questions that were raised this morning. So we'll go to that step now.

Dr. Woerle?

DR. WOERLE: More patients, and we weren't sure if we delivered you the right answer. These 74 non-fatal adjudicated events, which were not assessable, these patients were included in the analysis, but the events were not counted as outcome events. I hope this helps to clarify this point.

The second point, we wanted to make clear on the topic of silent MI. We conducted routine ECG measurements and this should be seen as casting a net most sensitive to get as many information as possible. However, only events who have been
raised by the investigator have been reported to
the adjudication committee. And of course all
events who have been reported to the adjudication
committee have been adjudicated as outcome events
or not.

The third point I wanted to make on presumed
cardiovascular death. While the number appears
somewhat high, I wanted to remind the committee
that in a patient population with established
cardiovascular disease where you can assume that
this was not a non-cardiovascular event, it is
reasonable to conclude that these patients should
be included as presumed cardiovascular death and
actually that this is standard in most of the
cardiovascular outcome trials.

DR. SMITH: Okay. Thank you. So, back to
questions. And I'm going to start with people who
got left out earlier for schedule reasons. So,
Dr. Everett, we'll lead off with you.

DR. EVERETT: Thank you. Actually, my
question pertains to the last comment that the
sponsor just made, and also maybe applies to
something that the FDA mentioned earlier. In particular, I think this idea that the patients who had a presumed and non-assessable death, that was then presumed to be cardiovascular in etiology, I tend to agree with you that it might be a reasonable presumption that those are cardiovascular deaths.

So my question to the sponsor is, why were those deaths not included in the primary endpoint as initially specified, either in the protocol or, as I understand it, until the statistical analysis plan was modified on May 21, 2015, after the trial was closed, last patient, last visit was completed?

DR. WOERLE: Please let me clarify. In the adjudication charter issued in 2010, it was already clearly stated that presumed cardiovascular death events would be counted as part of the primary endpoint.

DR. EVERETT: So where's the confusion that this was not part of the primary endpoint until 2015 then? Or sorry, yes, May 21, 2015. Why does there seem to be a discrepancy about that?
DR. WOERLE: I'm not entirely clear where this confusion comes from. I can show you the adjudication charter where it clearly states that presumed cardiovascular death would be part of the primary endpoint.

DR. EVERETT: Is it not in the statistical analysis plan? I mean, that's what the FDA's documentation would appear to say. And the reason why I mention this, of course, is because when you remove them from the primary endpoint, there are changes in the level of significance as seen on one of the FDA slides, which I can find for you here in a moment. Slide number 12.

DR. WOERLE: So allow me to clarify. From the very beginning the CEC charter -- and we have documented this -- in the initial version of the CEC charter stated that CV death includes presumed CV death. That was in 2010. This was added then in the TSOP for clarification to align with the CEC charter. But it was always the intention of the sponsor, at any point in time, to include presumed CV death.
DR. EVERETT: Does that answer the FDA's concerns about this issue, about how the presumed cardiovascular deaths were not included and then were included?

DR. GUETTIER: So we are in agreement with the sponsor that, in the CEC charter, the CEC charter from the beginning, at least the nine versions that we've reviewed, stated that non-assessable deaths would be counted as CV deaths. I believe the CEC form that was used for adjudication included CV death as a criteria for deaths.

The statistical analysis plan change was a specification that it would not be included, and that was done late. But we have no basis for thinking that whatever happened in the CEC charter wasn't going to be followed. I mean, at least that's the division's perspective. There might be some difference of opinion within the agency.

DR. HICKS: Thank you. This is Karen Hicks. I'm a medical officer in the Division of Cardiovascular and Renal Products, and I'm the
cardiology consultant to the Division of Metabolism and Endocrinology Products for this application.

I can't believe I'm actually helping the applicant here, but the CEC charter actually did have a line in one of the first paragraphs. And so I can attest to that. But I'd like to carry this a little bit further because it was never prespecified in any of the protocols. And this change right before the end of the trial unblinding, it was just kind of strange to see in the statistical analysis plan.

So if I could, LaToya, may I please have slide number 10 up? Here's what our FDA draft definitions paper says about undetermined deaths. And as you know, undetermined deaths were presumed to be cardiovascular deaths in this trial.

There are a couple of caveats. So in general, we believe that most deaths can be classifiable as either cardiovascular or non-cardiovascular. But the use of this category of death, undetermined cause of death, should be discouraged, and should apply to few patients in
well-run clinical trials.

124 patients with undetermined deaths is not a small number of patients. That's actually a large number of patients. And we go on to state -- slide number 11 please -- that it's a common analytic approach for cause-of-death analyses. And the caveat is that the approach should be prespecified, and we actually encourage that. It is not only in the CEC charter, but also in the protocols and also in any of the other trial documentation and statistical analysis plan.

So moving forward to slide number 12, please. So in summary, it's something that should be prespecified, but the caveat is, is that the number of undetermined deaths should be few. And when we see this many undetermined deaths, we know that there's missing data. I read the CEC minutes. The CEC members were marveling at how much data were missing for these undetermined deaths, and they kept requesting further information.

If you look at these cases, these deaths are completely uninterpretable, and they could be
non-cardiovascular deaths. They could be cardiovascular deaths, and there could be many possible etiologies, cardiovascular death due to myocardial infarction, due to stroke, due to heart failure, due to sudden death, but it was really a mixed bag. And I think, had there been more rigorous follow-up of this death information, that we may have a better idea about the mechanism behind cardiovascular death in this trial.

I think the take-home message here is, all of that said, if you do an extreme analysis and exclude all these undetermined deaths from the analysis, the effect on CV death is still robust.

Thank you.

DR. SMITH: Thank you. Dr. Rosenberg, you'd had a question hours ago this morning, and I've not forgotten you if you still have a question.

DR. ROSENBERG: Well, they have accumulated since this morning, of course. If you will allow me, I will take a little more time. I try with one not-simple question, but one question and then go back to the rationale and the background behind
that simple question.

I would like to ask the sponsor, given all the problems in collecting deaths and documenting events, as Karen just outlined, why not come back to a little bit of sanity here and a little bit of simplicity about what clinical trials used to be about, improving patient survival? So why not propose an indication for overall survival in this trial as you obviously had the power for it?

So now let's go back to what's behind this. The undetermined cause of deaths, and I think Karen just brought back the issue and it's still strong here. But it brings back to fact that it's very hard to classify deaths, but still, as has been said, it's really too high a number. And then there's the problem of the silent MI. And even the best expert in the world, like Marc Pfeffer, just really don't know how to classify silent MI, so that should that even be included there?

I'm not sure we need to spend hours discussing silent MI that shouldn't even be considered probably in this type of trial if you
don't have a prespecified process, very carefully
designed case report form, to collect the
information systematically. And then maybe the CEC
can do a decent job and even there it's very hard.

Then finally, the adjudication process -- I
think, at some point, we need to revise what we are
doing here in terms of adjudication. The Cochrane
just published a meta-analysis of 47 studies of
277 -- 375,000 patients comparing adjudication
versus no adjudication versus investigator-reported
deaths.

There's absolutely no difference, except, as
a caveat, maybe when the investigators are
unblinded. And here, we can assume that probably
most of them were unblinded even if it was
supposedly a double-blinded study, but they knew
what the glycemia was, if it was going down or not,
so just a little caveat.

But I think we have to question the value of
adjudication here. Karen asked you and probably
won't answer how many millions you spent on
adjudication that may be used to form better
research here. And I think the FDA leadership has
a very strong idea there, so we'll move on and
maybe we'll get -- the FDA will get away with
adjudication in most trials. So maybe you can
answer the question.

DR. WOERLE: Okay. CO-72, please. Slide 2
please up. This is the finding on all-cause
mortality. Our interpretation was, since CV death
had a hazard ratio of 0.62, with a highly
significant p-value, whereas non-CV death had a
hazard ratio of 0.84, but not significant, that the
benefits we see in overall mortality are largely
driven by CV death.

It's also our understanding that, very
unfrequently, overall mortality indications are
granted and we focused on the most relevant finding
of the EMPA-REG OUTCOME trial, which was the
finding on CV death.

I could not agree with you more, and
probably Jim Januzzi will say the same, on the
difficulties of adjudicating outcome events, in
particular causes of death. I wanted to point out
that we had an extremely rigorous process. The adjudication committee, whenever they felt they would not have enough information to determine the cause of death, asked the sponsor to reach out to the investigator. We had three attempts to get as many information back to the adjudication committee as possible.

Then I would like to add one point on presumed CV death. If the committee is concerned about the finding on presumed CV death and exclude that from the CV death findings, since these people have died, there is no doubt, it would be logical then to count them as non-cardiovascular death, and I would like to show what happens then. That's slide 2 up please.

Obviously, the finding on all-cause mortality remains highly significant. CV death remains highly significant with a hazard ratio of 0.59. But now non-CV death becomes significant, which gives you an idea on the robustness of the finding we have observed in the EMPA-REG OUTCOME trial.
The last point you made on the CEC charter, I'm not sure. You asked something on the CEC charter. Have I clarified or commented on what you wanted to know there?

DR. ROSENBERG: Well, I don't think you can answer the question. Whatever the charter says about classification, if you haven't collected the information, there's no way you get something valuable.

DR. SMITH: Okay. Dr. Wilson, you had had a question much earlier.

(No response.)

DR. SMITH: Dr. Hiatt, you had had one much earlier resolved.

(No response.)

DR. SMITH: Dr. Proschan, also much earlier? I'm aware of that.

DR. PROSCHAN: Right. No. I was interested in the stroke analyses. There was an increase in stroke, and I absolutely agree with the sponsor that looking at overall stroke, fatal and non-fatal, is what we should be doing. It does not
make sense to look at non-fatal stroke if you have to censor fatal stroke. That just makes no sense.

But there was that stroke disadvantage despite an advantage in terms of the blood pressure. And I'm just wondering has anyone done any kind of analyses where you look at differences in stroke. And I know this is a dangerous thing to do, but trying to adjust for post-randomization blood pressure, or doing some sort of landmark analysis where you look at how much blood pressure they've lost in a certain period and then move forward from there to look at strokes and to try and adjust for those differences.

DR. WOERLE: Yes, certainly we have in depth looked into the relationship of drop in blood pressure and potential occurrence of stroke. And our colleagues in drug safety have done numerous analyses and could not find an association. We looked into outlier analysis in extremes. Those patients with the highest drop in blood pressure, whether those were those patients who had a higher risk of stroke, we didn't find any of these
associations.

I fully agree with you that, in general, when you see a larger blood pressure drop, that you should anticipate a benefit on stroke. However, what is important to keep in mind is the magnitude of the blood-pressure lowering and also the baseline blood pressure.

When you look into the recent ACCORD trial, which is probably the best evidence we have, ACCORD blood pressure, where a similar blood pressure lowering was seen as in the EMPA-REG trial in the diabetic population, I highlight on the diabetic population. Actually, no benefit on stroke could have been observed. And if any, one probably had to wait much longer than the three years what we have observed in this trial.

DR. SMITH: Dr. Budnitz, you had a question.

DR. BUDNITZ: Yes. This is a question regarding CV death in subgroup analysis, as presented in the sponsor slide CO-54, but I use that just as a starting point. That shows the overall hazard ratio of 0.62 in various subgroups.
that are around 0.62. But I think one subgroup that isn't in the slide but in the very nice materials, is by region. And it seems like there's much greater heterogeneity by region in not just the cardiovascular death outcome, but I think all outcomes, ranging from a hazard ratio 0.35 in Asia, up to 0.81 that includes one in the confidence interval in North America.

So I'm just wondering if I'm interpreting this correctly. I'm less concerned with the confidence interval and more about the point estimate being much closer to unity for North America, and Europe as well. It seems to be a consistent finding across the outcomes. And I'm wondering if the sponsor has insights and make sure I'm understanding this correctly.

DR. WOERLE: Overall, our interpretation on the cardiovascular death finding is that there's actually very little heterogeneity with no significant p-values for interactions. We also looked into the regional distribution and the p-value for interaction was 0.14, not indicating
major interactions.

We looked into what from our point of view is an at least as important analysis. When you look into white Caucasians, living most frequently in North America and Europe, versus other populations and the hazard ratio for the white population is 0.64 with the upper bound being below 1, with a confidence interval of 0.83.

DR. BUDNITZ: So just to clarify, you don't think it's a significant difference between a hazard ratio of 0.35 in Asia and 0.81 in North America?

DR. WOERLE: We have extremes. We always have certain extremes. The test we apply is significant interactions. And when we look into what drives the EMPA-REG OUTCOME results, the majority of events come from North America and Europe. So the number of events being contributed actually from Asia is less than 50. And then when you get into relatively lower numbers, you see certain heterogeneity in the point estimate with widening of confidence intervals.
DR. BUDNITZ: Although there are equal numbers from Asia and North America in the study?

DR. WOERLE: What I'm saying is, when you combine North America and Europe, it gives you the most reliable estimate. And when you combine North America and Europe, the hazard ratio is virtually identical to the 0.62. And these are the two most robust cohorts we have.

DR. SMITH: All right. Diana Hallare, you had some questions much earlier today. Again, apologies for taking so long.

MS. HALLARE: Hello. This is also a follow-up to Mr. Budnitz. I would like to ask about CO-54 and also CO-41, wherein the African-American or black population, for instance, had a higher hazard ratio. And also if you would consider the ages 30 to 59, especially the ages 45 to 59 population, they are more leaning towards the placebo than towards the empagliflozin.

Also, I was wondering how could you explain that effect? I know there may be extremes in certain populations, but do you think, for
instance, that the lipid effect could be a factor in it or what could be possibly the background processes behind the effects that are seemingly extreme at this point?

DR. WOERLE: Let's review briefly 3-point MACE subgroup analysis, slide 2 up from the core presentation. It's correct that we saw some heterogeneity when it comes to subgroups. Subgroup findings become particularly difficult to interpret when you have about 50 or less events.

Now, we recruited in this trial 5 percent of the overall population who identified themselves as black or African-American. These 350 patients allowed us to make a reasonably robust assessment on the overall safety of the compound, which did not appear to be different from the overall population.

We have also information from our phase 3 program where we also recruited African-American and black population. We did not identify any differences in terms of blood pressure or response, if any slightly more, but no major differences in
LDL or HDL cholesterol.

The most important and the striking finding of the EMPA-REG OUTCOME trial is the CV death finding. And when you look into the CV death finding -- slide 1 up please -- all hazard ratios are on the left side of unity.

Now, a problem that I just mentioned is the number of events and the widening of the confidence intervals and the reliability of the point estimate. And one can see, for the African-American population, the population seems to have a similar effect as the overall population, but given the lower number of events, there is greater uncertainty.

But our overall conclusion is that in terms of general safety, there is no difference between the African-American population and the overall population. And in terms of CV death benefit, the same statement holds true, with limitation based on low numbers.

MS. HALLARE: I would also like to ask about the effects or interactions with other medications,
for instance with the beta blockers and the ACE inhibitors, for instance. So I was wondering if that could have an effect on different side effects or whether that could have had effect on the lipids, for instance.

DR. WOERLE: We looked into subgroups prespecified as for ACE and beta blockers. There was no significant interaction with very similar hazard ratios, 0.65 and 0.61 for ACE and the same for beta blockers. So there was no interaction.

MS. HALLARE: Thank you.

DR. SMITH: Dr. Neaton?

(No response.)

DR. SMITH: Dr. De Lemos?

DR. DE LEMOS: Yes, a question for FDA. Can you expand a little bit on the reluctance to consider all-cause mortality in general as an indication and specifically in this case where you have a cardiovascular death finding that's out of context already?

So how is that qualitatively different than all-cause mortality, particularly given all
the challenges we've just talked about with regard to classifying death?

DR. GUETTIER: I'm going to call Dr. Temple to the microphone.

(Laughter.)

DR. TEMPLE: I'm not sure historically how true that is, but you can think of a reason. It's hard to think of a drug that effects all causes of mortality, so it doesn't make a whole lot of sense. But we frequently look at all-cause mortality to see if the beneficial effect on one subset is countered by another. So we are interested in it.

If you had a profound effect on cardiovascular mortality and overall mortality went the other way, you'd agonize a lot about what the heck was going on. But I think that's the reason. It doesn't make sense that a drug would affect specifically every single kind of mortality, and I think that's why. But I'm not sure we've -- we've put it in labels. We've shown the results. So I don't think there's a systematic conclusion.

DR. SMITH: Dr. Cho?
DR. CHO: Just a quick question about regional variation in non-assessable cardiovascular death. Was there one particular region with higher rates of inability to assess death?

DR. WOERLE: I'm looking at my colleague from drug development, Uli Broedl. Can you answer the question, please?

DR. BROEDL: Uli Broedl, Medicine. We looked at presumed cardiovascular death, not only in the overall population, but also per region, and there was no difference in the occurrence of presumed CV death per region, no difference.

DR. SMITH: Dr. Schambelan?

DR. SCHAMBELAN: So this refers to FDA's slide 12 in the clinical assessment and to the point about censoring the non-assessable deaths. So, as a cardiovascular consultant pointed out, the findings were still robust for CV death. But if I understand that slide correctly, there's no longer superiority demonstrated for the 3-point MACE. So is that simply a matter of numbers, or what are we to take from that difference?
DR. CHONG: I just want to make sure I have the question right. So you want to know why on slide 12 the statistical significance is lost for 3-point MACE when you exclude non-assessable death. Correct?

DR. SCHAMBELAN: The point was made earlier that it's still a robust finding if those people are excluded whom we don't know the cause of death. The CV death is still a robust finding, but we lose the superiority upper bound for 3-point MACE. Is that simply a matter of numbers? I want to try to understand that. And that's my guess.

DR. SCHAMBELAN: Right. Fewer deaths that feed into the 3-point MACE which is driving --

DR. WOERLE: Lack of power.

DR. SMITH: Okay. That's resolved by the data, by looking at these data again? Or you're --

DR. GUETTIER: You basically have less endpoint events, and that's what you get when you do the analysis whether or not -- of course it's a retrospective analysis. It also makes certain assumptions, which may be valid or not valid, and
that's for everybody to decide.

DR. SMITH: Dr. Rosenberg?

DR. ROSENBERG: Thank you. I'd like first
to push back a little against what Dr. Temple just
said, with all the respect I have for him. The
patient who takes a new drug, he wants to know is
he going to live longer or not. That's the bottom
line.

I understand from the drug development
clinical trialist point of view, it only makes
sense to look at overall mortality if you know the
drug has specific targeted effect if there's no
competing risk. It looks like for this drug, this
class of drug, cardiovascular mortality is
overwhelming. So I don't see any problem for using
overall mortality and getting rid of all this
adjudication and all this nonsense.

Now I have two questions, back to the data,
one about stroke. I really don't understand how
and why the sponsor didn't look more carefully at
stroke in terms of separating hemorrhagic stroke
and ischemic stroke. I haven't seen the data
clearly.

From the CEC point of view, sponsor, I don't know why you included TIA, which has no clinical meaning whatsoever when you compare it to the other strokes. It doesn't make sense to pool TIA with stroke, so I'd like to see clearer on that.

The second question was regarding if we started splitting and slicing data, let's continue. And I'd be interested to see data by country, country-specific data rather than region. For example, I don't remember having seen Western Europe versus Eastern Europe, but where it was all coded into different countries.

DR. WOERLE: We've done these analyses. We have presented in the briefing book, but we're happy to provide you the data. And Dr. Sven Koehler, who did extensive assessment will walk you through.

DR. POCOCK: So first of all, to clarify, TIA was not included in the primary endpoint. Only strokes were included in the primary endpoint.

Second, if we looked into types of stroke,
hemorrhagic strokes were reported in less than 10 percent of the cases. Slide 2 up, please. So the vast majority of strokes indeed were ischemic strokes.

We went back retrospectively and further evaluated the types of stroke and asked the adjudication committee to assess the respective types of strokes. And I would like to ask Dr. Bernstein to comment on that.

DR. BERNSTEIN: Can we have slide 2 up, please? Richard Bernstein, stroke neurology from Northwestern. If you look at the vast majority of strokes in this study, they were ischemic. And as the panel knows, ischemic stroke is really not just one disease. It's several diseases.

So, for example, large artery atherosclerosis is a quite different disease than small vessel occlusion, which are both quite different than cardioembolic stroke, which is due usually to atrial fibrillation. And undetermined etiology, which accounts for a sizeable number of strokes here, is a grab bag of different biological
processes.

What you can see here is the relative number of events in the placebo and empagliflozin arms, and there's a non-specific increase in all types of events, which to me is not a biologically plausible result since all of them are different mechanisms.

DR. SMITH: Is that a follow-up related to this, Dr. Good, related to this same point?

DR. GOOD: Yes, just a follow-up of that. I read in one of the reports that the drug group had a lower use of anti-platelet agents and anticoagulants and had a higher risk of atrial fibrillation. And I wondered if that might possibly explain some of the apparent difference, maybe not the whole thing.

DR. WOERLE: It's correct, we saw some subtle differences in the use of anticoagulative medication. We saw some differences in the regional distribution. You saw the data this morning in Europe that the stroke rate was much lower in the placebo group, being suggestive that they might have received more comprehensive
preventive medication. But when we looked into the
data, we could not establish a clear association.

DR. SMITH: Dr. Palevsky, you had a
question?

DR. PALEVSKY: So we haven't talked much
about the kidney disease aspect. And looking at
your presentation versus the agency's presentation
and your briefing materials, in your combined
outcome of new or worsening nephropathy, I'm
confused by the renal replacement therapy, which is
described in some places as continuous renal
replacement therapy and in others as renal
replacement therapy.

As the agency points out, continuous renal
replacement therapy is a treatment that is almost
exclusively used in patients with acute kidney
injury in the setting of critical illness. So can
you clarify that aspect of the endpoint and what
you were really assessing there?

Then the other is the doubling of serum
creatinine and eGFR of less than 45, which again
can represent either acute disease or chronic
disease. So can you provide additional granularity on how you know or don't know that this is progression of chronic disease for these patients?

DR. WOERLE: Certainly. Can I have slide 2 up? And may I ask Dr. Kim to join me? And one of the limitations of the data we have on the EMPA-REG OUTCOME trial that indeed events were not adjudicated, which leaves us with a certain degree of uncertainty when it comes to the initiation of renal replacement therapy. What is the striking finding that all components of the composite are going in the same direction, but we are aware of the limitation given small numbers.

But you specifically asked on acute versus chronic kidney renal replacement therapy, and Dr. Kim will provide you the answer.

DR. KIM: So to clarify the first question, which was pertaining to is this actually chronic renal replacement therapy or acute renal replacement therapy, it was a mixture of both. Slide 3 up, please. So, when we did the analysis, both acute initiation of renal replacement therapy
and chronic renal replacement therapy were included. And the endpoint that would show us in terms of initial renal replacement was a combination of both.

When we look at specifically chronic renal replacement therapy, indicating a start and continuation of replacement therapy until the end of the trial, we saw fewer numbers, but the hazard ratio was still similar to what we observed for the total. The second question was pertaining to GFR doubling of creatinine.

DR. PALEVSKY: So verification that the doubling of serum creatinine and eGFR was a persistent decline of that rather than representing either biological variation or true acute kidney injury.

DR. KIM: The doubling of creatinine was also a singular measurement of doubling of creatinine, and that would have made that endpoint.

Slide 2 up, please. So the first row indicates the number we have presented, the doubling of serum creatinine, and lower than 45 mLs per min per 1.73
meters square.

When we look at sustained doubling of creatinine, which would be defined as a measurement after 30 days having a continuous doubling of creatinine, the numbers do decrease.

When we look at other modern endpoints, such as sustained 50 percent decrease in eGFR or sustained 40 percent of decrease in the eGFR, the numbers again increase. But looking at the hazard ratios, they continued to be consistent with the initial finding based on single creatinine measurements.

 DR. SMITH: Dr. McBryde?

 DR. MCBRYDE: I'm going to follow up on Paul's inquiry about the kidney. I had a question for both the sponsor and for FDA. For the sponsor, a couple of quick questions. One, I was curious which MDRD formula that you used. I presume it was probably demographics and serum variables because I'm curious if there were urinary assessments done as part of the evaluation.

 DR. WOERLE: Dr. Broedl will provide you the
answer.

    DR. MCBRYDE: Thanks.

    DR. BROEDL: Uli Broedl, Medicine. You are correct, there were no urinary samples or data that were included.

    DR. MCBRYDE: Then I couldn't find in the materials the urine albumin to creatinine ratios. How were those standardized, and do you know what assay was used for the measurement of the albumin? Was it a turbidimetric, a nephelometric, radioimmunoassay?

    DR. BROEDL: So I can't answer this on top of my head. All I can say right now is this was based on spot urine measurement and at regularly defined visits where we took safety levels.

    DR. MCBRYDE: In that spot, do you know if they were first morning urines, or if they were random samples?

    DR. BROEDL: Our trial protocol, the visit usually should have been in the morning, so the presumption is first morning visits, but this was not fully standardized.
DR. MCBRYDE: Thank you. And then for FDA, I wanted to follow up on slide 23. You had analyzed the data after the wash-out looking at the albuminuria. I was curious if you had done a similar analysis looking at the eGFR measurements and the blood pressure measurements.

DR. LUNGU: The eGFR does go up. I don't recall for blood pressure, but I think it does as well. But the eGFR we did not present, but it does go up.

DR. MCBRYDE: Thank you.

DR. GUETTIER: I think the sponsor had the eGFR figure in their talk.

DR. WOERLE: Yes, slide 1 up, please. So does this speak for itself, or do you have questions?

DR. MCBRYDE: Thank you.

DR. WOERLE: We did similar measurements for blood pressure and blood pressure values came back to baseline. And the difference at the end of the trial is approximately 4.5 milliliter.

DR. MCBRYDE: Thank you.
DR. SMITH: So where we are -- yes, Dr. Palevsky?

DR. PALEVSKY: On that last slide, just since you have switched from using MDRD at another point and using CKD-EPI here, which CKD-EPI equation? Is this creatinine? Is this creatinine cystatin-C? So which equation are you looking at, and why are you switching eGFR calculation methods?

DR. WOERLE: Okay. We have Dr. Wanner with us. He's the first author on the recently published New England Journal who is best equipped to provide you this answer.

DR. WANNER: Christoph Wanner, University of Würzburg. I appreciate to respond to your question. I think the first one you were asking was CRRT, continuous renal replacement therapy. It was a language issue. When I came in, I corrected this. This is dialysis. It's not CRRT on the intensive care unit.

The albumin measurement was nephelometric. The four laboratories forming the core laboratory were using the same assay. The creatinine was
enzymatic measurement and it was isotope dilution, 
mass spectrometry corrected, so standardized. And 
the MDRD formula was introduced in 2010, and we 
decided to do so because at this time, the CKD-EPI 
was not validated across the globe and the MDRD for 
an entry criteria.

But then the reviewer from the New England 
Journal paper asked us subsequently, you can use 
the CKD-EPI to create the slopes since the 
creatinine is standardized. And then we switched 
to creatinine. And cystatin-C, we only measured in 
a thousand patients and this is not in this 
analysis.

DR. SMITH: So given where we are on the 
clock, I would like to move to the discussion 
questions. There were a few people who had their 
hands up for clarifying questions. So I want to 
move forward, but if there are questions remaining 
that you feel are important before we move related 
to the discussion questions, we'll entertain those. 
So is that -- yes, Dr. Konstam?

DR. KONSTAM: I don't know if this really
even needs an answer, but with regard to what occurred at the time of the interim analysis, the sponsor said something to the effect of a Chinese wall or complete protection between the people who were unblinded and not. The FDA presentation said there were 230 individuals unblinded.

It's kind of inconceivable to me that 230 people can keep their mouths shut for two years or whatever it took. I mean, that could be a comment. If the sponsor wants to answer it, I guess that would be okay.

DR. WOERLE: We have implemented, after the agreement with FDA on the conduct of the trial and the interim analysis, a comprehensive plan to maintain trial integrity. All individuals had to sign confidentiality agreements.

DR. KONSTAM: No, I'm sure you did the best you can.

DR. WOERLE: We undertook every effort, but at the end of the day, you're right. Effort is nice, but what is the outcome of your effort? And what gives us the greatest reassurance, when you
look into the primary endpoint and the components of the primary endpoint and when you look at patients being included prior to the interim analysis and those patients being included after the interim analysis, that you basically get almost virtually identical hazard ratios.

Questions to the Committee and Discussion

DR. SMITH: Okay. We will now proceed with the questions to the committee and the panel discussions. And I just want to remind public observers that while the meeting is open for public observation, public attendees may not participate except at the specific request of the panel.

We'll go to the first discussion question, and I will read this. Discuss your interpretation of the EMPA-REG OUTCOME study conduct. Please comment on whether interim unblinding, or changes made to the protocol, endpoint definitions, and analysis plan, for example specific exclusion of silent MI from the primary endpoint, during the course of the EMPA-REG OUTCOME study alter or do not alter your level of confidence in a conclusion
that excess CV risk was excluded, and CV benefit was established.

So I'll open that for discussion or for comment from members of the panel. Dr. Hiatt?

DR. HIATT: Thanks. Well, these points have already been made. And this trial was primarily designed as a trial to exclude a certain level of cardiovascular risk, and it was repowered to be able to establish superiority. When you sort of play with those numbers, the event rate on placebo was 4.4 percent, which is kind of a standard event rate. It's not super high.

If you're trying to power for efficacy in a cardiovascular outcome trial with say a hazard ratio of 0.85, you need about 1200 or 1300 primary events. So 600 events seems kind of low to me for a primary MACE endpoint, which left me a little bit confused as to really what the intent of the trial was. I think it's easiest to interpret the trial on establishing non-inferiority, and it's a little more challenging with the MACE primary to establish superiority.
I think the other components of these questions, including silent MI, may be easier to answer in number 2.

DR. SMITH: Yes, Dr. Konstam?

DR. KONSTAM: Well, the question basically is, are we influenced by these things we've been concerned about in the trial. And I guess I am really principally because so much rests on the p-value for the superiority for the MACE 1 endpoint that sits at about 0.04.

There are issues about maintaining the blind during the interim analysis. There are issues of missing data for the MACE 1 endpoint, MACE 3 endpoint. There are issues about silent MI, which I'm not so concerned about as an individual thing.

But I guess, in aggregate, I guess I am concerned about them really because, you know for that endpoint of MACE 3, which I believe is really important here, it's a marginal figure to begin with.

DR. SMITH: So again, to follow up that, I think in the discussion question itself, at least
as the FDA wrote it, they've really focused -- they
don't mention MACE 3 and they're really focusing on
cardiovascular risk.

I back off. I've reread the question.

Cancel my last two sentences. Other comments on
this? Dr. De Lemos?

DR. DE LEMOS: I just think with the
indication, it's not relevant. Right? I mean,
we're talking about a cardiovascular death
indication and these are all important issues. The
silent MI is an important issue, but we're talking
about cardiovascular death, not the primary MACE
component. So frankly, the blinding issues are
less important there, too. And whether there was a
firewall leak, when we're talking about a death
endpoint, I think these issues kind of wash out.

DR. SMITH: Other comments? Dr. Neaton?

DR. NEATON: I mean, I've been involved in a
number of trials where we've measured silent MI. I
would agree with what the sponsor and I think the
FDA is concurring, that they just dropped that. I
think that's a bit of a mess, and so I think they
just went about doing it incorrectly, and possibly labeled it incorrectly.

When I measure silent MI, you want to make certain the surveillance of people in the 2 groups is done comparably over the whole course of follow-up, which typically means ECGs at standard intervals, and comparing serial ECG change. And so they didn't do that from the start.

I share Dr. Konstam's concern about the MAC 3. I have to say though, I was reassured with two things in the FDA analyses. One, they did address, at least to a limited extent, the impact of missing data. And they're right on the border. The upper bound of that confidence interval was 0.996.

The other thing they did, which relates to some of the earlier discussion, which I think was actually reasonable potentially in a trial where you expect a large fraction of the deaths to be cardiovascular, is simply change your composite to look at fatal/non-fatal MI, and stroke, and all-cause mortality. And in that situation, the result
is actually stronger.

    I think while trials can always be done
better in terms of follow-up, the sufficient
analyses led me to kind of have some reassurance
and confidence in at least the outcome that we're
going to be talking about later.

    DR. SMITH:  Dr. Cooke?

    DR. COOKE:  I think the aspect of the study
design or conduct that bothers me the most was that
large percentage of deaths that were unassigned or
uninterpreted. I forget what the exact term is.
But the fact that 40 percent of the deaths were not
defined and had to be put into the presumed
cardiovascular death, the fact that's such a large
percentage, does bother me.

    I think the FDA made the point that what you
do with those -- it's important what you do with
them, but those numbers should be relatively small.
And especially if we're trying to discuss an
efficacy trial where 40 percent of the deaths we
don't know what the cause is, that bothers me.

    Now, that's mitigated a lot by a number of
the other sort of analyses. The fact that all-cause mortality was significantly lower with the treatment, and the fact that the sensitivity analyses in a number of ways, whether you assign them to all cardiovascular or not cardiovascular, or eliminate them completely, the data are supportive of efficacy. But again, I think that number does stick out a lot to me.

DR. SMITH: Dr. Rosenberg?

DR. ROSENBERG: Yes, I agree with Dr. Neaton and Dr. Cooke's comments. I mean, that bothers me not because it seems to change the outcome, but it bothers me about what else has happened during the study conduct that we don't know.

The fact that such a large trial with so much monitoring, they see such a level of uncertainty about so many things related to the outcomes, and I assume other things that we're not sure how the data was collected so that we're not able to validate those critical elements, that does bother me.

DR. SMITH: Dr. Proschan?
DR. PROSCHAN: I was initially troubled by the fact that it seemed like perhaps people were seeing results and that's what caused them to resize the trial to show benefit. But after hearing the FDA's presentation, I'm convinced that that's not what happened because the FDA actually encouraged them to increase the size of the trial rather than have a separate trial. So that doesn't bother me anymore.

Also, the undetermined cause of death doesn't bother me that much because the all-cause mortality result was significant. And I disagree a little bit with Dr. Temple. I usually don't disagree with Dr. Temple. I know better.

(Laughter.)

DR. PROSCHAN: But I think when you say that there's a benefit on all-cause mortality, you're not saying, this benefits every single cause of mortality, you're just saying there's a benefit on mortality. Most of the mortality is cardiovascular, and so I think it is meaningful if you prefer to not say all-cause, just say there's a
benefit on mortality. I think that's substantiated from the results here.

DR. SMITH: Thank you. And Dr. Everett?

DR. EVERETT: Just a quick comment that I think the fact that the primary, the endpoint that we're discussing here is a mortality endpoint. As Dr. De Lemos said earlier, it makes this much easier in some respects because, as the sponsor has pointed out, it's not like we're misclassifying somebody as alive versus dead.

We're convinced they're dead. It's just a question of classifying them as cardiovascular versus non-cardiovascular. And depending upon how you draw that line, the benefits are present in both groups or more substantially in cardiovascular death.

I think the wandering road that the trial took organizationally is notable. And I bet if the sponsor could have a do over, they probably wouldn't have as many amendments and readjustments and realignments. That said, I'm not convinced that those had a material effect on the outcome or
in my confidence of the outcome. If we were
talking about a different endpoint than mortality,
my answer there might be substantially different.

DR. SMITH: Dr. Cho?

DR. CHO: Thank you. I think for me, the
most interesting part was reading the recently
published LEADER trial, which also had a placebo
cardiovascular death rate of 6.0 percent, which is
very similar to the death rate in the EMPA-REG
OUTCOME study, which actually reassured me a great
deal. Even though there was a 40 percent presumed
non-assessable cardiovascular death, it's
remarkably similar to another diabetic trial.

DR. SMITH: Dr. Fradkin, did you have a
comment?

DR. FRADKIN: So I just looked up the
methods, the design paper for the study, which was
submitted in May of 2014. And even then, they were
describing as a significant secondary outcome the
time to individual occurrence of silent myocardial
infarction. So I am confused by this sort of
wandering course.
I'm just wondering why, if shortly before the study ended, they said that this was in fact a significant secondary outcome and they were doing these EKGs, I don't understand how in this kind of timeframe it sort of -- we lost silent MI. And I am concerned because it looked, from the data, like silent MI might have been a hazard signal.

I feel like even though what they're asking for is cardiovascular mortality, from the statistical perspective, you wouldn't be able to consider that sort of part of a composite outcome if you didn't think that the composite outcome was solid. So I think I get back to what Dr. Konstam said. The 0.04 is not super strong.

Now, that was not including -- they never included silent MI, as far as I can tell, in this design paper in the composite, but I think sort of the whole variability of it does lead you to question whether the composite outcome was strong enough to get back to what was said in the initial FDA presentation about not needing two trials.

DR. SMITH: Yes, Dr. Konstam?
DR. KONSTAM: I just maybe want to engage my colleagues a little bit because I'm hearing people viewing what's going on here in different ways. And one of the comments around the table has been well, you know, the real meat is going to be in the cardiovascular death endpoint, so the MACE 3 endpoint may not be as important. And I'm kind of not thinking exactly along those lines.

Let me just ask folks this. If the p-value for the MACE 3 endpoint were 0.07 instead of 0.04, and the CV death and all-cause death were exactly the same as they are, would we be saying the same thing?

I think the answer is no because I think you won't have hit the endpoint that gets you to look at the cardiovascular death, and you've lost the hierarchy even one level up. So I think, at least the way I'm thinking about it, whether or not the MACE 3 endpoint is positive or not to me is real important in terms of whether we should even be considering the cardiovascular death endpoint. But I don't know if -- I'm willing to be dissuaded from
that if people think --

DR. SMITH: Dr. Thomas?

DR. THOMAS: I actually agree with you completely. If the MACE 3 endpoint did not meet significance, we probably wouldn't be having this meeting. You would have an intriguing finding that there's less cardiovascular deaths and less deaths overall, and you'd probably require another trial to prove the point.

So there's a little bit of luck that these cases that were deemed to be cardiac were put into that category. If they were excluded as non-cardiac deaths, as the FDA data shows, your point estimates become larger, you're not significant, and that's probably just due to having less events to decide this.

So in a way, the sponsor is lucky that it fell in a significant way. And I think that's the ultimate decision, do you feel that that's enough to sway with all the other evidence. It did meet the primary endpoint, but there are some issues with the data and the data collection, and how much
of that is going to play the role.

    The only thing is no one's addressed the other part of this question, which is, has CV risk been excluded. And I think, actually, that seems to be fairly confident that that has been excluded. But I agree, I think the 3 endpoint has to be the crux of whether you make this decision or not. If you're comfortable with that being significant, then you can go along the rest of the pathway.

    DR. DE LEMOS: Why? I mean, why would we interpret it that way with the p-value? I mean, it's a probabilistic statement. The finding is completely out of context, the mortality finding, and we're evaluating an out-of-context standalone mortality finding with an incredible p-value and trying to decide. They're not asking for an indication for reduction in the MACE endpoint. They're asking for an indication of cardiovascular death.

    I guess it fails the common sense test to me to say that we will consider this. Would we rather be considering an indication where the primary
endpoint hit at \( P \) equals 0.48 and the cardiovascular death and total mortality were borderline non-significant? We give them that indication, but we won't give them a mortality finding?

This is a much stronger result for patients and for the drug. Whether it's enough to stand alone as a single trial is a separate point. I mean, that's a different question. But it makes zero sense to me to take this -- this finding is clearly out of context with the composite. There's no explanation within the composite for why one sees this degree of mortality benefit. It's unexplained and it stands alone, not as a component of the composite.

DR. SMITH: Yes, Dr. Konstam?

DR. KONSTAM: Well, it's the last question where we'll start drawing conclusions, right. So I'm listening to my colleagues as you are. But there are a number, I'm sure, as you know, of prior cardiovascular trials where there was a statistically significant benefit on either
mortality or cardiovascular mortality that was not the primary endpoint and was not in the hierarchy of secondary endpoints.

It might have been a down-the-road secondary or an exploratory endpoint. It might have been something that was looked at post hoc. It might have been something that came out of a subset analysis. And now, in all of those circumstances, well, at least the ones I'm thinking of, when it was replicated in a trial specifically designed to look at mortality, it didn't show up at all.

Now, there are things maybe that can get us there. Okay? But I mean, I would stick to the point that we approach questions in a clinical trial in a hierarchal format in order to retain alpha. And for this particular endpoint of cardiovascular mortality, we're out of alpha. So why would you -- let's say if it were something else, I mean, I think we're affected by the fact that it's mortality, but we still want to get it right.

Let me just make one more point. I think
there are things that could influence us despite that to say, yes, but I'm going to go with it anyway, for example, a very, very, very low p-value. The problem I have with that -- and I respect Dr. Pocock, but I don't think that the real p-value is calculable in a setting like this where we just have an endpoint that's floating, where we really don't have it in the hierarchy, but maybe that could influence this.

The second thing that might influence us is the number of events. Okay? So there are 300 some-odd events, cardiovascular mortality events. And we might say it really is about the number of events, that's a critical factor. Those could help us get there I guess, but I really would stick -- if you want to be rigorous about it, you believe something if you got to it in a rigorous statistical way.

DR. SMITH: Dr. Thomas, respond.

DR. THOMAS: I think to just add to what you mentioned is really the crux of this in terms of the final decisions that are made, is this
sufficient based on what we have to be a single trial to give this indication. I was on the committee for the JUPITER trial, which was a much larger trial and I think it was much clearer. And even then, there was some hesitation about a single trial, no matter how well done, to get an indication like this.

So really, that's going to be the end question that people ask, is this one trial, based on these factors, even though the cardiovascular reduction and mortality is, from what the data is, is dramatic, going to be enough to require another trial or is this sufficient.

DR. SMITH: Dr. Wilson?

DR. WILSON: So to build on what was just said, I think what we're missing is a plausible really well-defined and substantiated mechanism. So we have glucose, and it doesn't appear to be glucose. We have blood pressure. The blood pressure effects are not real big. We have heart failure.

Throughout this meeting, the heart failure
information has really been downplayed because it's not like what we've seen for some of the other studies for heart failure. And some of the definitions of heart failure admissions are rather short-term admissions. And also some of it's change in medication programs.

I think that's why we're circling back to what Marv Konstam's been saying, how did we get there? And can we really say this is it for a death outcome, but we're not sure how we got there? So I still have some concern about how we got there and raise the hierarchal issues that were originally raised almost as the very first question of the day by Will Hiatt. Important issue.

DR. SMITH: Who had his hand up again, Dr. Hiatt?

DR. HIATT: I feel we're wandering between question 1 and question 2. And so in terms of strength of the evidence and trying to dissect what this means, I think I'll hold off until question 2.

DR. SMITH: Dr. Palevsky?

(No response.)
DR. SMITH: Dr. Rosenberg?

DR. ROSENBERG: Yes, maybe I should hold off, but I'll go anyway. Well, in fact, it's two questions. One, related to what's just been said, I really agree that we didn't get a clear explanation of the mechanism. And especially when you see at the divergence of the curve within the first three months, which is almost unheard of, I heard talking about hemodynamic but really without clear explanation about what that means that explains such a drop in mortality. Patients in this study didn't seem to have severe heart failure, or any other problem that would explain that. So I'm really puzzled by that.

My second question, again, may happen later with answering the question is to the FDA.

DR. SMITH: I'm hesitant to ask more questions because, really, at this point, if it's something that will critically affect this discussion, okay.

DR. ROSENBERG: Yes.

DR. SMITH: But we've sort of moved beyond
questions to us discussing what we got.

    DR. ROSENBERG: I think it will, at least in my mind.

    DR. SMITH: All right.

    DR. ROSENBERG: Is there any precedent for this kind of study or other of the FDA having given regulatory approval based on the component, the positive outcome on one component of a composite endpoint when the composite endpoint was the one that was approved by the FDA as the primary outcome?

    DR. SMITH: Does FDA have a quick answer to that?

    DR. STOCKBRIDGE: I would say we routinely dissect out a component and try to figure out whether or not all of the components actually contribute to it. And I don't even think, when we do it, we insist that any one of the components actually ends up being statistically significant in order to name it. It's mostly about trying to weed out components that clearly don't contribute to the overall observed effect. It's done for descriptive
purposes.

DR. TEMPLE: I can give you an example.
What Norm said is right. When you have a composite
that's made up of multiple different things, why
would you even expect they all go the same way?
But the LIFE study comparing atenolol and losartan
had a composite endpoint the same as this, and it
won nicely showing that losartan was better.

When we look closely, all of the benefit was
on stroke. There was no hint of a benefit on MI or
anything like that, but fatal and non-fatal stroke
was responsible for the whole thing, and that's
what the label says. That's the claim they got.
We also noted that it didn't work in the black
population.

So you do look at these things, and how to
do it, and whether that makes sense, and whether
you've corrected for multiplicity is a very good
question. But as Norm says, you sort of have to
look.

DR. KONSTAM: Can I just follow through with
a question to you, Bob? But what trial were we
just talking about?

DR. TEMPLE: LIFE.

DR. KONSTAM: LIFE. But the primary endpoint of the LIFE trial was positive. Right? I don't remember the p-value. I think it was a pretty low p-value.

DR. TEMPLE: You're exactly right. You had to win on the primary endpoint --

DR. KONSTAM: Right.

DR. TEMPLE: -- which is what you said.

DR. KONSTAM: So you could have chosen how to word the indication, but what would you have done with that stroke finding if the primary endpoint did not hit statistical significance?

DR. TEMPLE: No, no. What you said before was right. It's presumed that you have to win on the primary endpoint.

DR. KONSTAM: That's right.

DR. TEMPLE: That's right. And that's one of the -- that's why you're asking that here.

DR. KONSTAM: Just to dissect that one step further, you also like to see two trials. So you'd
like to see -- if you're going to go with one trial, you'd like to see a pretty low p-value in that primary endpoint.

   DR. TEMPLE: Well, in LIFE, we didn't have a very low p-value.

   DR. KONSTAM: It wasn't that low?

   DR. TEMPLE: No.

   DR. KONSTAM: What was it?

   DR. TEMPLE: I mean, it wasn't 0.001 or anything.

   DR. KONSTAM: I'm going to look it up. I'm going to look it up.

   (Laughter.)

   DR. TEMPLE: It was better than 0.04.

   DR. HICKS: I really need to speak up about this because I believe that the heart failure findings in this trial are not reliable. I looked at a bunch of the CEC packets on heart failure and a lot of the events that were adjudicated as heart failure hospitalizations shouldn't have been adjudicated as such.

   Also, what I found in my review is that
events were being adjudicated as heart failure, and
they did not even meet the criteria of the CEC
definition. And based on my review, I found that
probably the definition that was used by the CEC
got further diluted to mean any ER visit and any
oral dose of Lasix.

Lastly -- and perhaps Dr. Januzzi can
provide some additional information on this -- it
appeared that these events were being
double-counted. So if patients came in with a
myocardial infarction and they had heart failure
secondary to the myocardial infarction, these
events were being double-counted as a myocardial
infarction and also as a heart failure
hospitalization.

So I think if the applicant can speak more
about this -- but this is what was found in the
adverse events datasets and also in the time-to-
event CEC datasets.

DR. SMITH: So does the applicant have a
response to that, just briefly?

DR. JANUZZI: Thanks for the opportunity to
respond. It is true that there were cases of patients that would present with heart failure and at the same hospitalization, they would have myocardial infarction adjudicated as well. One important thing to emphasize is that we adjudicated the myocardial infarctions within the context of the universal definition of myocardial infarction wherein type 2 or so called supply/demand infarct is not an uncommon scenario in the context of heart failure hospitalization. So that's just one situation I'd emphasize.

Relative to other specific individual concerns, I would say when we adjudicated cases, we adhered to the CEC charter for the definitions. I can't speak to individual examples, but that's obviously something you've had the opportunity to look at.

DR. HICKS: So yes, I did see those cases with type 2, but there were cases of type 1 MIs and heart failure was associated with the type 1 MIs, and there was double-counting of events.

DR. WOERLE: We conducted various
sensitivity on heart failure outcome. Slide 3 up.

We used the adjudicated events as being described in the main presentation. We used loop diuretic introduction as a proxy for hospitalization, for new hospitalization. And we used the introduction of loop or heart failure.

You see, we did the same analysis on e-reporting on heart failure and we get always very consistent findings. Slide 2 up. This is based on symptoms, so edema, serious adverse events and heart failure reporting. Slide down, please.

DR. SMITH: Dr. Proschan, did you have a comment?

DR. PROSCHAN: Yes. I think we are encroaching on question 2, and that one actually is encroaching on question 4. But I just wanted to comment on, y if the p-value for MACE 3 had been 0.07, would we even be here. I think we definitely would be here.

This secondary outcome is not feelings of dread and foreboding. It's cardiovascular mortality and it's overall mortality. I think we
would definitely be here. And I think we do need to consider it.

I know that there are people who say, okay, if you don't get a significant result on the primary, you can't look at anything else. But if you're looking at mortality and you're seeing a highly significant p-value, even after you make a pretty strong adjustment for mortality -- I mean for multiplicity, sorry -- I think that is pretty compelling.

On the other hand, the one thing I will say is there's always more multiplicity than you think because you could consider well all the different ways to analyze the data, the different components of each outcome, when you add all that up, I think it's more than 42. I'd be willing to bet a lot of money that it's more than 42.

DR. SMITH: Dr. Cho, did you have a comment?

DR. CHO: I wanted to say that the MACE that we're looking at in EMPA-REG is an athero MACE, atherosclerotic event MACE. And the mechanism of this drug is probably not its benefit on
atherosclerosis, and I think that we should all be sort of mindful of that.

Traditionally, we've always accepted MACE as sort of the athero MACE, the MI, stroke, and whatnot, but I think we come to an area now where we have some -- who knows what mechanism to decrease cardiovascular mortality through a different mechanism. So that's my one comment.

My second comment is, recently there's been a paradigm heart failure study looking at ENTRESTO and their primary outcome was death and CHF hospitalization, which is an accepted endpoint for heart failure trials. They redid the analysis looking at a softer heart failure endpoint, as was done in EMPA-REG, visits to the ER, increase in IV diuretics. And not surprisingly, the death rate with these soft endpoints was also reduced with ENTRESTO.

So I think that in the light of 30-day readmission and people trying to get patients not admitted to hospitals, looking at these sort of 24-hour emergency room visits is not out of
ordinary. And we will be seeing this more and more throughout other trials.

DR. SMITH: So I would like to summarize very briefly, perhaps not inclusively, on the discussion we've just had. And so what I heard was expressions of concern in regard to the various issues such as changes in protocol, endpoint definitions, analysis plan as posed in this discussion by the -- for this question by the FDA, that that presents concerns and issues in terms of interpretation of the data.

A couple members of the committee expressed a high level, or a pretty high level, of confidence in the conclusion that excess cardiovascular risk was excluded, and that wasn't opposed by comments from other members of the committee.

In regard to cardiovascular benefit, there was a lot of discussion. There's substantial concern expressed by various members of the committee. One of the points that was raised was that the fact that there's some uncertainty in regard to the MACE 3 primary outcome raises some
fundamental questions about the amount of emphasis that should be placed in interpreting the cardiovascular mortality data. That was not a universally expressed opinion. There was controversy on that from within the committee.

Certainly, there were comments made on the importance clinically and the clarity of definition of mortality, all-cause or cardiovascular mortality. There were issues raised about the non-assigned deaths that then ended up as presumed cardiovascular mortality. That's an issue, but it was noted that, in that regard, that all-cause mortality still upholds significance, and that was somewhat reassuring.

Another point that was made, which is one of the overriding issues that adds to some of the difficulty of interpreting these data, is that we haven't been presented with a plausible mechanism for the effects of this drug on cardiovascular mortality.

So that's my summary. Did I leave out anything critical that people would like to state?
(No response.)

DR. SMITH: I think we'll move to discussion question 2.

Dr. Rosenberg, you had a comment?

DR. ROSENBERG: Yes, just a slight modification on your last point. The mechanism of action may probably not on atherosclerosis because it's very early and we don't have a clear explanation of this early benefit.

DR. GUETTIER: Can I get just a little bit more clarification, Dr. Smith? It's Jean-Marc from the FDA. So this trial, to some extent, was a child of the 2008 guidance and a lot of the issues, at least in the conduct that we sort of came about, came out because actually this is really a safety trial.

I think I heard sort of at least the specific discussion point was really focused on whether or not what we presented this morning in terms of the changes. What is your level of worry about the quality of the data that comes out of this trial? A lot of the discussion, actually,
from this discussion point, were more towards discussion point 2 and 3, but in terms of the level we heard from Dr. Konstam, I was wondering if anyone else had any other comments on that.

DR. SMITH: Would anyone like to respond to that? Yes, Dr. Rosenberg?

DR. ROSENBERG: I think I have no concern about safety.

DR. EVERETT: I'll answer.

DR. SMITH: Dr. Everett, I got the wrong person there.

DR. EVERETT: My confidence in the mortality data is reasonably high. My confidence in some of the other endpoints, myocardial infarction and heart failure in particular, is much lower. And I'd be very wary about -- because of the design issues we've discussed, we haven't really even addressed the heart failure issues at length yet in this discussion, but that would be my short answer to your question.

DR. SMITH: I might give a longer answer, which is that I -- so the 2008 guidance includes
within it language that focuses on MACE or a MACE-plus. And when we look at those specific endpoints that were within the 2008 guidance, and then if we look at an extension to an efficacy trial, we obviously encountered some uncertainty with the data within this study. Depending on how we choose to analyze, include, or exclude in terms of the analyses of some of these factors that were raised as concern, not only MACE 4 but MACE 3 comes under question. So if we simply look at the simple extension of the 2008 guidance into an efficacy superiority trial rather than a non-inferiority trial, we have encountered some issues in the context of this study.

We happen to have had another event happen here, which is that we have this rather dramatic cardiovascular mortality data. And in a way it's sort of on and off topic to me in terms of your question because that's a component of a MACE, but instructions aren't clear from that 2008 guidance. So I think there are complexities in trying
to move from a guidance that was generated to help in the design of studies with a primary goal of establishing safety to something that is looking at superiority, and we've encountered some of those difficulties. I don't know if that's helpful.

Yes?

DR. KEWALRAMANI: From the perspective of the design of trials, whether the original intent was the exclusion of the hazard ratio of 1.8 or 1.3, or the construct of such a trial for superiority over standard of care, I think that the study here was powered. While there were changes, these changes were made apparent.

There was a CEC involved, and I think whether or not silent ischemia was in or should have been in, whether it was designed as a safety study or efficacy study, that would have been the same. And the non-assessable -- I think, while the numbers may be higher than we would like, I don't think that that is a function, per se, of safety study for the exclusion of a 1.3 or 1.8 or for the demonstration of superiority.
So I think if we really look under the hood of this study, it would look like a study that we would design for the purposes of demonstration of cardiovascular morbidity and mortality.

DR. SMITH: So, I would like to move to discussion question 2. I'll read this. Please discuss the persuasiveness of the statistical results for the primary analysis. Please also comment on how results for the individual components in the primary composite endpoint impact your level of confidence in the study findings.

Finally, comment on concerns you may have related to potentially incomplete ascertainment of some myocardial infarction events, i.e. silent MI, in this trial, and whether these concerns, if any, alter your level of confidence in the results for the primary analysis.

So we've obviously in part discussed this, but let's briefly at least revisit those issues. Yes, Dr. Hiatt?

DR. HIATT: I thought a little bit more clearly I think in terms of trying to frame the
issue. So you come into this with a guidance and
the idea that we're trying to exclude the risk of
MACE and now show that we have superiority on MACE.
And if you think about the history around that,
around drug trials using anti-thrombotic agents or
statins, really what drives that historically is a
reduction in fatal and non-fatal myocardial
infarction.

That's how you sort of come into this. And
when you then go back to the primary MACE endpoint,
which is on a single pivotal, not very convincing,
and then look at a couple of other analyses, MACE
on-treatment, which should be a little stronger,
has an upper bound of 1.02, MACE-plus silent MI,
upper bound of 1.06.

Fatal and non-fatal MI, you know the thing
that you think would drive it, has an upper bound
of 1.09. And there's no benefit on stroke,
combined with the idea that I think we get to CV
death as a post hoc exploratory analysis. So
that's the problem I'm having with this.

So in some ways, I think, Michael, you've
described this, you sort of have to sort of set all
that aside and say it's just CV mortality, that's it. And I don't know if there's a mediation
analysis or something looking at what's driving that. I just don't think we have a clue.

It doesn't look like it's the kind of mechanisms that most cardiovascular outcome trials
are designed to test, which is plaque rupture and thrombosis. I don't think it's that. It might have something to do with that, but it's driving mortality down for some other reason.

I think that's where my discomfort with trying to wrestle with what to do here, is really strong because starting out this discussion, I couldn't get to CV mortality very easily because you went on MACE and then there's nothing beyond that that really logically takes you to a hierarchal analysis to move into the components. But you sort of say you're just being awfully technical, and sort of acting like what I learned from you guys, Norm and Bob, on cardio renal, so I'm a little hung up on this point.
But that's the problem here. Because it's mortality, you said what the heck? Just throw all that other stuff out and it's just death, and nothing else really matters. But then, if that were confirmed by another agent in the class, say, wow, there's something going on here, and maybe we'll be smarter in five years and figure out what the mechanism is and say we should have thought about that, but it certainly makes sense.

The history of statins were sort of getting on the market because they changed a surrogate. And then suddenly, they show cardiovascular benefit later and things came together. So there's precedent for that. But I think the statistical part of this is weak. I just don't think it gets me to this single-component analysis.

In that regard, I'm sort of left with just saying, I don't know, I'm not convinced, but there's a CV mortality. And it's just hard to kind of get your arms around that in some rational way.

DR. SMITH: Dr. Neaton?

DR. NEATON: Let me just ask a question
about the history and the focus on MI. So my
history and trying to discern whether a person died
from fatal MI or sudden death, that's very hard.
So they've actually classified the deaths by fatal
MI, so when they look at fatal versus non-fatal MI,
I'm assuming that's what they're using, that
classification.

But presumably there are sudden deaths in
these not accessible deaths that are being judged
cardiovascular that would probably fall into that
category. I mean, the fatal/non-fatal MI results
are trending in the right direction at least, the
same as the overall outcome. So I think the
overall -- my summary of this is that you're on a
very fragile area for the primary MACE outcome.

I don't know, Marv, how I would react if it
was 0.06 or 0.07 versus 0.03. I think the range of
p-values we're seeing is somewhere in the range of
maybe 0.02 to 0.10, based on the various
sensitivity analyses. But it's being driven by an
outcome that really makes a difference. And every
way you look at that outcome, it adds up as having
an impact on mortality.

So I don't think you can ignore it is the way I think about it. And you should view these p-values as giving you some strength of the evidence and not getting locked in on the magic 0.05.

DR. SMITH: Dr. Konstam?

DR. KONSTAM: Well, first of all, everybody I listen to changes my mind.

(Laughter.)

DR. KONSTAM: So be careful what you say. But just as a point, I want to sort of go after this point about, well, this is mortality. This is really important.

I think the question in front of us is do we believe the finding. I mean, that's the question. It is not this is a really important finding, therefore we ought to believe it. I mean, that's backwards. Right? We ought to first decide do we believe it, and then we decide that it's important.

I think that the issue around, when we get to it, about the cardiovascular mortality -- so
first, let me just say, this is a drug that's already approved. Okay? There's a paper in the New England Journal about this that says that the drug reduces mortality. So we're not talking about approving the drug or not approving the drug. It's already in the academic literature that that's what it does. It reduces mortality.

The question here is, is the FDA going to approve it for that indication, for that goal. And the way I think the FDA thinks about this in its core is, do I believe the finding or not. Is this something that we agree with so strongly that we're going to say to the prescribing public this drug reduces cardiovascular mortality? So I think that discussion is independent of this is a really important endpoint, to me.

DR. SMITH: Dr. Budnitz?

DR. BUDNITZ: I was going to add that the statistical analyses that most intrigued me was actually the on-treatment analyses, which seemed to make no difference for the endpoints in MACE of MI, and if I'm recalling, I think stroke. But where
they did make a difference, and I think the about 25 percent discontinuation rate, so I would expect there would be more efficacy for the on-treatment subgroup.

   It did show up for cardiovascular mortality, where the overall rate of hazard ratio I think was, for example, on the order of 0.66. But on treatment, it was 0.52. So for me, that was reassuring that this drug might do something, and it was notable that the on-treatment results were completely the same for the other MACE endpoints.

   DR. SMITH: Dr. Schambelan?

   DR. SCHAMBELAN: I just wanted to add, I did look up the data for LIFE and the p-value was 0.021, but the point estimate was almost identical to what -- this is for the MACE outcome -- was I think 0.85 or 0.86, very similar to what we saw here. So I don't know, Marv, how you feel about going from 0.02 to 0.04 to 0.07.

   DR. KONSTAM: I don't know.

   (Laughter.)

   DR. KONSTAM: I think -- like I'm trying to
be rigorous with what I've learned from the FDA over the years they want. Right? That's all. And what they want is two studies. And if they don't have two studies, they want to feel good that they still have the right answer. And one of the important things they look at is how low is the p-value.

So, I just ask the question, if the p-value of the primary isn't that low, how do you know it's right, okay, with a single trial, and does that affect your willingness to approve it for a component of the primary endpoint. That's what they did. Maybe they need to explain themselves, but that's what the issue is.

DR. SMITH: Further discussion on this point? Yes?

DR. YANOFSKI: So I'm really struck by the fact that the evidence I'm hearing doesn't seem like this is driven by a reduction in atherosclerotic cardiovascular disease in terms of the reduction in those MACE components, in terms of the timeline. We don't really know the mechanism,
but where you're really having your bang for the buck is in this really robust reduction in cardiovascular mortality.

I was reassured by the fact that even when you did the sensitivity analysis, not looking at those where you couldn't really -- where you didn't know if it was cardiovascular or not, it didn't really change your findings. So everything I'm hearing tells me that this is a real finding. We don't really know why. We don't know the mechanism. Do we need to know the mechanism to approve the indication? I guess that's the question.

DR. SMITH: Dr. De Lemos?

DR. DE LEMOS: Yes. I think just to summarize, there's some excellent points. I think that I'm more persuaded by a weak p-value for the primary endpoint and a really robust p-value for the mortality endpoints than I would be for two P 0.02s for both. I think that's what -- some people would be happier with that result in terms of allowing the drug to get a cardiovascular
indication.

I'm also very favorably influenced by the multitude of influences on blood pressure and favorable influences on renal function and heart failure that, while by no means conclusive and terribly limited, all line up in the right direction. The non-fatal effects on myocardial infarction -- with the exception of stroke, everything lines up in the right direction to support a cardiovascular mortality benefit that's favorable.

It's not conclusive, but then what's really buttressed is the all-cause mortality. It's not going to be the indication, but as Brendan said, I think that's what gives you -- that markedly increases my confidence in the CV mortality finding.

DR. SMITH: Yes, Dr. Good?

DR. GOOD: Yes, since stroke was just mentioned here by my colleague, perhaps we should talk about that just for a minute. I think the most important thing is that there wasn't any
statistically significant difference in HR between the treatment groups and the placebo, and that's a fact.

I think there's a very soft signal that perhaps there's a slightly higher risk of stroke, but I think the point is that it was across the board, across all stroke pathophysiologies was pretty interesting. It's lacunar stroke, it's cardioembolic stroke, it's atherosclerotic stroke. Maybe that kind of feeds into this mechanism issue as well.

I have a little bit of concern about whether there is a regional difference in diagnosis of stroke. I have no idea why the stroke incidence was higher in Europe than it was in other areas, and the sponsors had no idea either. I'm not quite sure what to do with that.

Regarding the TIA issue, a TI is usually considered in the same category as ischemic stroke. It's a little bit harder to diagnose because there are many causes of TIAs. They're not always ischemic stroke. I'm not sure about rolling that
in with ischemic stroke and looking at then how that changes the HR, whether that's really legit or not.

I think in the end, I'm not terribly concerned about the stroke, the non-fatal stroke. I do think that there's a little signal we have to think about, but I don't think we should focus extensively on that point.

DR. SMITH: Yes, Dr. Fradkin?

DR. FRADKIN: Thanks. I agree with Dr. Yanovski that the fact that we don't know the mechanism shouldn't be a reason not to approve it. Clearly, it's something different than atherosclerosis just based on the time course, based on the fact that we started to see the lines diverge from the very beginning. I think we would have to be looking at some other mechanism, something about survival of an MI rather than atherosclerosis.

To me the entire issue though boils down to whether when it's one of multiple secondary endpoints, are we sure, based on the fact that the
death, robust as it looks, is a secondary endpoint. Can you approve it with one trial? And that to me is the crux of the question.

DR. SMITH: Anyone have a specific comment in regard to a component of this discussion point in terms of the silent MI? We've already talked about this extensively, but any further comment that anyone -- Dr. Hiatt?

DR. HIATT: Yes, I sort of added that in my earlier comments. But I think there are a number of things that make the p-value dance around 0.05 for the MACE primary, and that's all it does. I think we talked a lot about it earlier that it's probably not reliable. It's probably mostly noise.

But, you know, the on-treatment analysis takes it to non-significant, and the fatal/non-fatal MIs are not significant. So I think it's all kind of telling us that this is not having a profound effect on ischemia per se, it's probably something else. And that's all I've learned from that discussion. I don't think it's a fatal flaw.
DR. SMITH: Yes, Dr. Proschan?

DR. PROSCHAN: Yes. I absolutely agree that the evidence for MACE 3 is not that strong and you don't have to do much to tip it the other way. And I am aware of the fact that there have been other trials where you see a very small p-value on a secondary endpoint, you try and repeat it, you can't repeat it.

Again, my explanation for that is that there is a lot more multiplicity than you think. If you had seen that the 10-milligram dose had shown a really strong benefit, you would have emphasized that. So there's a whole lot more multiplicity than you think.

However, I still believe that, based on the evidence that we've seen, that the mortality finding, cardiovascular mortality is real. Now is this estimate of the hazard ratio biased?

Absolutely, yes it is. So it's not going to be as strong as what's indicated here. I feel pretty confident about that. But I do also feel pretty confident that it's not a hazard ratio of 1, it's
less.

DR. SMITH: Dr. Rosenberg?

DR. ROSENBERG: I would slightly disagree with Dr. Fradkin. When we're talking about a secondary endpoint that is strongly positive like it is observed, you really would like to have a mechanism to understand why it's such. That's the only thing I would add.

DR. SMITH: So I might summarize that in responding to this question, we a little bit tracked over the same set of opinions and views that have been presented with the first discussion question, which is that there is concern or uncertainty about the primary combined MACE endpoint. And that's affected by a number of variables, including looking at the individual components of that, which further highlight the issue that we don't really understand the mechanism of the effect on mortality.

Again, we heard a divergence of opinions from the committee in terms of the mortality data. And that is not so much questioning what those data
are, but the actual strength of that as a single study result. And I think with a greater acceptance of that as a strong indicator of benefit by some members of the committee and much stronger expressions of uncertainty about other members of the committee as to whether that in fact would be reproduced in a second trial.

Anyone want to add anything, or FDA settle for a quick summary on that point?

(No response.)

DR. SMITH: So we'll go to the next discussion question. Again, we've covered some of these points, but this I think will bring back some emphasis on others. Discuss the persuasiveness of the mortality findings in the EMPA-REG OUTCOME study.

In your discussion, please address any potential limitations of these data, including but not limited to issues raised in discussion point number 2, the proportion of deaths that were determined non-assessable by adjudicators, the lack of granular data on potentially important
information such as baseline heart failure history
and dose of relevant baseline and concomitant
medications, and the lack of prespecified alpha
adjustment for this endpoint.

So I'll open that for comment. Yes,
Dr. Konstam?

DR. KONSTAM: Actually, I have new
information to bring to the group, okay, because
I've been working here. So with regard to this
question, most of my focus has been on lack of
prespecified alpha adjustment for this endpoint,
and I think there's really concern about that.
Now, I want to say something, a few things on the
positive side.

There are things about this that drive you
toward -- help me get a little bit more comfortable
that this is a real finding. One is the very low
p-value. The second is the number of events, okay,
and I'll get back to that in a second, 309 events.
I'm focusing on the cardiovascular death endpoint,
so 309 events.

I'm also influenced by the fact that you
could argue there's sort of two trials in here because there are two different doses, and both of those doses show the same thing. So that actually is the one piece of information in the dataset that moves me along in thinking about this.

I looked up three trials where there was something like this, that is, there was a mortality endpoint that jumped up and wasn't expected either because it wasn't the primary, it was a subset analysis, some secondary analysis or something.

The first one was PRAISE with amlodipine. And the finding was that in dilated cardiomyopathy patients as opposed to ischemic heart disease patients, it reduced mortality. That finding was not confirmed subsequently with a larger trial.

In that trial, the first one, the PRAISE trial, there were a total of 103 events compared to what we have here, which 309 events. It was associated with a pretty low p-value. They just quoted as less than 0.001, so the p-value didn't help very much, but maybe the fact that it was a third as many events may influence us.
The second one is with losartan against captopril. There was an ELITE trial for which the primary endpoint focused on renal function comparing captopril and losartan, but a mortality benefit for losartan popped out of nowhere.

So there, first of all -- by the way, in both of these trials, in the first PRAISE trial and in the ELITE trial that I'm mentioning, the primary endpoint was non-significant. Okay? That's another point. But this finding with losartan versus captopril was not replicated in ELITE II. In ELITE, the p-value was only 0.035, and there were only 49 events.

The third example is the carvedilol trials in the mid-90s where there was a huge amount of controversy about whether the mortality data should go into labeling for the drug because the finding was a composite of four different trials, none of which was designed to look at mortality. That one was approved after a second panel with 53 events. Why was that approved?

(Laughter.)
DR. KONSTAM: Fifty-three events and a
p-value of less than 0.001. And then, actually,
Bob mentioned the LIFE study. Well, that's
different because it didn't show mortality benefit.
I won't talk about that.

My take home out of this is that if you look
at all of the information in those three trials,
none of them were as robust as what we've got in
terms of both the p-value and the number of events.
So I think those compared to these other findings
that turned out not to be true looks a little
stronger.

DR. SMITH: Any other discussion of this
question? Dr. Thomas?

DR. THOMAS: First, I just wanted to echo
what Dr. Cho mentioned, which is, at least in the
United States, I don't know about the rest of the
world, with observation status being required by
Medicare, and many insurers also having that, you
have to include those patients, even if they're not
hospitalized because you're very unlikely to get a
straightforward heart failure patient being
admitted nowadays in a hospital that has
observation care because they can only get
readmissions. And with the new insurance rules,
you won't be able to get them hospitalized.

So it is something probably future trials
will have to look at because otherwise, if you just
go by hospitalizations, you're going to miss all
these cases.

That's an important point and it addresses
this heart failure issue. This wasn't set up to be
a heart failure trial, so in a way it doesn't
completely matter if all the data is not as
complete or rigorous as it would be in a heart
failure trial. All the other cardiovascular
information is more persuasive in terms of
cardiovascular death.

Then the last thing is we don't have a
mechanism. But if you think about mechanisms, not
glucose, because with glucose, we've had large
trials that have not shown glucose to work,
atherosclerotic is unlikely. It's so quick.
Right? Blood pressure, it had a blood pressure
effect, but it's still pretty quick. We should see
similar effects of blood pressure agents.

One thing we've been skirting around -- and
I think the reason I bring this up is hopefully, if
there is a trial that's done or other analysis
probably needs to be looked at, is we've been
mentioning the word dead in bed a lot. And that
usually relates to people who have autonomic
neuropathy, poor RR variability, and that's not
something that was probably looked at in this trial
but is something we probably should address.

Is there some effect of these agents on
autonomic neuropathy or some other surrogate that
is having an impact on sudden death or arrhythmias?
We won't need the mechanism right now; shouldn't
preclude us from approving something or not
approving something because, you know, aspirin was
around for 80 years without a mechanism.

DR. SMITH: Dr. Budnitz?

DR. BUDNITZ: I think the concern I have
about the mortality finding relates to the third
bullet there, the lack of granular data on
potentially important information based on heart failure and concomitant medications, particularly because so much of the data come from outside the U.S., like 80 percent outside North America. And we just don't know what the -- or at least I don't know what the treatment and baseline characteristics of these folks are, whether it's diet, whether it's the dose and duration of other medications, particularly when we don't have a mechanism.

So it's not an issue of statistical interaction. It's an issue of generalizability where I'm not quite sure what these folks look like, if they look like the patients that are being treated in the U.S.. And so this lack of information about if the mechanism is some sort of hemodynamic and heart failure mechanism, we don't have information about what medicines folks are taking for heart failure and their baseline echoes or other characteristics.

DR. EVERETT: Dan, could I just ask you a follow-up? Are you worried that the randomization,
that those issues are not balanced on either side of the randomization? Or what are you worried about? Because to me, that's a concern in terms of defining mechanism, but you have a randomized trial, so theoretically at least, they should be balanced.

DR. BUDNITZ: Right. And so it's not that there's no effect, but the magnitude of the difference. So for example in Asia, the hazard ratio was 0.3. In the U.S., it was 0.8 for cardiovascular mortality. So you can do a test for interaction, and that can be negative, but clinically I think that might be different. And so I do think that there's the potential for there to be differences by region. Does that make sense?

DR. SMITH: Dr. Proschan?

DR. PROSCHAN: Not to throw another statistical issue in, but people have talked about trying to repeat a result and you don't get the result the second time. That's not necessarily saying that it wasn't true, right, because you have two trials. One says there's something going on,
the other one says there's nothing going on. You don't know what the right answer is still.

Moreover, when you do the second trial, what people usually do is they say let's assume the treatment effect that we saw in that last trial, and let's power it for that. That treatment effect is going to be biased because you did notice that because it was a secondary outcome, and, wow, look at this. So that treatment effect is almost surely going to be too strong. That's not really what the true effect is. So really, you should take that into consideration.

Then, when we stupidly a lot of times power trials for 80 percent, which means even if you're correct that the treatment effect was what you saw last time, you have a 1 in 5 chance of not getting a significant result. So the fact that another trial didn't confirm it doesn't really tell you that it's not true.

DR. SMITH: Dr. Palevsky?

DR. PALEVSKY: A couple thoughts. As a trialist, we use composite endpoints because we
don't think that any one of the individual endpoints in the composite is going to have a high enough event rate for us to have a significant outcome with a trial that is going to be feasible to conduct.

If we were to ask a patient what's important and ask them the hierarchy of the endpoints in the composite in either MACE or MACE 4, I think I'm just going to speculate that they'd put mortality at the top of that list and take non-fatal events as of lower importance.

So although I have concern over the fact that the mortality is moving in a different direction from the overall composite, which may or may not have quite reached the p-value of 0.05, I am quite impressed by that. I don't think that the sponsor would have dared propose a trial just looking at cardiovascular mortality.

That said, I am a little bothered by the lack of the mechanism. And one concern, not so much the difference in the co-management at baseline, but could there be some effect of
interventions after the patient's on trial, either alternate management of glycemic control that may have diverged because this is a hypoglycemic agent and what are the co-management? Or diuretic management since it does have a diuretic effect?

One of the questions that I have is, what were the impacts of the co-management on the trial on this mortality outcome? And could we be missing something that it may not be entirely the benefit of this agent, but a mixture of benefit of this agent and some detriment to the co-management that was occurring?

DR. SMITH: Yes, I think that's an interesting comment. And just to sort of expand on that a little bit what some of those factors potentially might include would be the difference in hemoglobin A1c, which was modest, somewhere in the range of 0.5 percent, plus or minus difference, depending on how you choose to look at the data.

There was a difference in blood pressure that again was only a few millimeters of mercury, but it went in a direction downward, which one
might think would be favorable. And there was a
difference in additional medications used to manage
glycemia, even though the glycemia wasn't as
effectively managed so there was more insulin and
there were more sulfonylureas.

I find it difficult to focus on any one of
those, and not even particularly comfortable to
pool them all together and say, "Ah ha, there's the
explanation." But the fact is they do again add a
little discomfort. It's another version of the
mechanism argument. It adds a little discomfort to
the strength to place on the mortality data
themselves. Dr. Thomas, you had a comment?

DR. THOMAS: So I think those are all
important points. Actually, when the guidance
first came out and the way these trials are
designed, physicians have the judgment in these
trials to adjust medications as they feel fit.

There's no predefined protocol as in other
trials. So, for example, in most trials like
ACCORD, you had a target. So if you blood pressure
didn't hit that target, you had to do something,
either to increase the medication, you just
couldn't sit as is.

Here, there is no guidance. You just do
what's the norm for your community. So the fact
that actually anything shows up amongst all this
noise of treatment is to me actually quite
astounding.

I actually was very surprised that anything
would show up in these trials beyond
non-inferiority because there's so much noise in
the management. Someone could up the statins,
someone could not up the statins. They're not
treating towards our guidelines or treatment. They
should, but there's nothing requiring anyone in any
of these management of the patients, to do anything
other than give the placebo or the drug.

DR. SMITH: Dr. Li-Ng?

DR. LI-NG: I just wanted to point out my
general concern about the indication that's been
sought by the sponsor. The SGLT2 inhibitors are a
relatively new class of medications that have been
added to our diabetes treatment regimen.
In regards to some of the studies that Dr. Konstam had brought up, there were other studies that showed some benefit in terms of ACE inhibitors or ARBs or beta blockers to add the indication for cardiovascular event prevention.

My concern is that given this one study again showing very persuasive evidence that there is a decrease in cardiovascular mortality and all-cause mortality, I can't -- again, my concern is that this is one study in a new class of medications and we haven't seen any other studies showing the same thing.

DR. SMITH: So what I would like to do is summarize what's been said so far, and any critical additional comments can be made. I'm anxious to move us along because we have two more discussion questions and two voting questions and I don't want to be rushed at the end here.

So in regard to this discussion question, again, we've a little reiterated some of the same issues. We've heard some arguments about the persuasiveness of the mortality data in terms of
the very low p-value itself, the substantial number
of events in this trial, particularly in comparison
with several historical trials, and the fact that
the two doses of the drug showed really very
similar effects on the mortality endpoint.

We, again, members of the committee
expressed concern about a non-understanding of
mechanism. And the suggestion was made in regard
to potential effect on autonomic nervous system and
the relationship between disturbed autonomic
signaling and sudden death.

But also the point was made that it may not
be essential to know the mechanism in order to
place strength in the mortality data. And in that
same context, the point was made that among all the
elements of the composite endpoint, the one that
clinically certainly would be viewed as being the
most important is the mortality data itself.

Again, in regard to how the FDA may consider
responding to these mortality data, the point was
made that this does represent the first member of a
drug class. So again, it's another version of
being a single study, one might say. I don't want
to open a bag of worms with that comment.

Any other comments before we move on? So
let's go to the next discussion question. And I
don't want to short-change any of these questions,
but I'd like us to try to be a little bit brief so
we have adequate time to deal with the voting
questions.

So this question, number 4, discuss the
heart failure findings in the EMPA-REG OUTCOME
study. Please comment on the potential limitations
of these data, if any, and on whether the results
of this study establish a benefit of empagliflozin
on heart failure and heart-failure-related
outcomes. Yes, Dr. Konstam?

DR. KONSTAM: This one, I have a bigger
problem with. Okay? So it's nowhere to be found
in the hierarchal testing. It's not a component of
the primary endpoint. The p-value is not as
impressive. So as has been said, you really have
no idea what it really is. It's nominally 0.0017,
but given the multiplicity involved here, it's
really undeterminable. And for what it's worth,
it's not death.

I'd like it to be true. I kind of think it probably is, but it doesn't reach my standard of provability.

DR. SMITH: I'll try Dr. Thomas.

DR. THOMAS: I would just add to that and then FDA can correct me if I'm wrong. If this was being done as a heart failure study, this agent would have to be an add-on to currently approved standard of care. We have no idea what was really being done throughout this trial. So it's an intriguing finding, and if it really needs to be explored, it probably needs to be explored in a different study that's formally done for heart failure.

DR. SMITH: Dr. Everett?

DR. EVERETT: I just want to preface my comments with the fact that heart failure is a very prevalent problem. It's an important problem. It's increasingly important. And to echo Dr. Cho's comments earlier about a shift in the way that
these patients are managed from hospitalization, to outpatient treatment units, to potentially even -- I have an email in my inbox about a randomized trial of treating them at home. So there's a lot of impetus to try and treat these patients outside of the typical healthcare setting, so it's an important endpoint.

But, when you have a definition that shifts over the course of the trial and is not anywhere near close to the primary endpoint, to echo Marv's points, or a component of the primary endpoint, my willingness to accept these results a priori, I think, without the benefit of a second heart failure-focused trial where done by investigators practiced in the art of doing those trials, which are very challenging, I'm reluctant to accept the results at face value.

DR. SMITH: Other comments? Yes, Dr. Palevsky?

DR. PALEVSKY: So I will echo the comments about the rigor. And just another point to remember, this is an agent that is a diuretic as
one of its mechanisms of action. And we just need
to bear that in mind that it's going to have an
effect on volume management that may not have
been -- you know, basically the patient was on a
higher dose diuretic than otherwise considered.
I'm not very impressed by the heart failure data.

DR. SMITH: Dr. Wilson?

DR. WILSON: Just a follow-up. I think we
lack modern metrics, and we need some new metrics
for this. And perhaps, for instance the person who
goes into the ED observation unit and goes home,
that's one hospital day, if that, versus those on
the wards, we have hospital days, D-A-Y-S, plural.

So that's another thing to think forward,
because this is going to keep coming back to these
sorts of committees, what is the heart failure
metric in 2017, 2018, 2019. So one of them is
days. The other is major changes in outpatient
doses. But I still go back to some sort of symptom
complex and weight loss with the diuresis over a
short term. We need better metrics for evaluation
of this outcome moving forward.
DR. SMITH: Yes? Dr. Proschan?

DR. PROSCHAN: So I agree that I think the results are not nearly as compelling for heart failure. But I also wonder why is this even a question if they are no longer asking for that, anything to be said about heart failure.

(Laughter.)

DR. SMITH: Well, I think the FDA poses these discussion questions not limited to a specific indication that may be under consideration, but something that may help to inform them in a background way for that decision, but also in terms of future issues. This is a drug where there may be a future heart failure study, for example. So I don't think they're restricted to working around their indication request.

Dr. Rosenberg?

DR. ROSENBERG: So maybe I'm old fashioned, but a change in class of heart failure is something that we used to look at and we don't even have baseline here, so it's hard to say anything.

DR. SMITH: Dr. Cho?
DR. CHO: I just want to reiterate that this is clearly not a heart failure trial and they didn't even have New York Heart class association as part of the trial. But I just want to -- I mean, it's a very interesting finding, and I want to encourage the Boehringer Ingelheim group to study it in the heart failure patient population because it would be so interesting to look at it in that particular population.

DR. SMITH: To summarize what's been said so far here, committee members have expressed the view that this is an important clinical endpoint. I'll add to that, we haven't talked in detail here in this discussion about the prevalence of heart failure within type 2 diabetes, but it's high and it's clinically significant. It's limiting clinically to people who have heart failure functionally. It's limiting. So it's an important endpoint.

It's been noted that these data are quite intriguing, and so I suppose we can interpret that if in fact a beneficial effect on heart failure
were adequately substantiated, that would be an important effect of the drug.

The point was noted that this drug has diuretic properties. So in that context, one would want to know whether they were using a diuretic and whether there were alternative diuretics or whether this is an effect coupled with other actions of the drug or independent, which would be important to consider in clinical care.

That notwithstanding, there were a lot of concerns raised about interpretation and the strength of the heart failure data within this study in that it was not designed primarily as a heart failure study. There is data that one would like to have that aren't there in terms of baseline heart failure data and other modes of assessment of heart failure issues about the definition of heart failure. And the point was even raised that we may need better metrics for heart failure. So certainly, in considering future studies, that's something that one has to grapple with as best as one can.
I would say that's a summary. Anything to add? Dr. Rosenberg?

DR. ROSENBERG: It's just a little side comment. If you accept that the drug is effective in reducing cardiovascular mortality, you have to consider in which population you'll be able to ethically randomize patients to the drug if you want to study heart failure.

DR. SMITH: FDA, is that enough input from us or do you want more discussion of this?

(No response.)

DR. SMITH: So we'll go to the next discussion question, which is to discuss the renal findings in the EMPA-REG OUTCOME study. Please comment on the potential limitations of these data, if any, and on whether the results of the study establish a benefit of empagliflozin on kidney disease related to diabetes. Dr. Palevsky?

DR. PALEVSKY: So I don't think that the renal endpoints were rigorously enough collected and adjudicated to really come to strong conclusions. It's certainly intriguing, the slope
of the eGFR change. But slope of eGFR change, I believe the FDA would not accept as an endpoint.

Certainly, the albuminuria seems to be a transient effect. Stopping the drug, the albuminuria comes back together. We know that diuretics do that, so I'm not sure that it's telling us anything about structural disease.

If this is going to be looked at for a renal endpoint, it's going to have to be looked at with a very rigorous adjudication of whether the change in kidney function is persistent, whether it meets the accepted standards of either a 50 percent decline in -- or reducing the 50 percent increase in serum creatinine or 40 or 50 percent change in eGFR, which was not rigorously done.

There will need to be very careful adjudication of AKI versus progression of CKD, which clearly was not done, and understanding of acute versus chronic dialysis, which seemed a bit hazy.

I think that there's very intriguing data here, but this is far from the level of data that
will establish a specific renal benefit. And we'd love a drug with a specific renal benefit.

DR. SMITH: Yes, I agree. And again, just to somewhat reiterate, it's important to separate transient and reversible effects from sustained and longer-term effects. I know you made that point, but there is a suggestion in the data with the limitations in the data, that there may be some sustained effects, unlike the effect on urine albumin excretion.

This is a context again where the other co-variables within the study are critical, as you and I have discussed before. They are ever more critical because we're looking at microvascular disease where we have a stronger argument for a role of glycemia. And we know that there was a difference in A1c within the study that one would not want in a study that's evaluating effects on renal complications in diabetes.

Similarly, there was an effect on blood pressure, which one would not want in this context. So those issues become, I think, even somewhat
greater in terms of limiting the interpretation.

Nonetheless, I agree as well that it's an intriguing finding that if it in fact were validated with another study -- and I don't well know that it would be -- that would be a potentially very useful effect of the drug.

DR. PALEVSKY: I would also add, in addition to co-management, we don't know about dosing of RAS blockade, which is the standard of therapy for nephroprotection in these patients, and were there differences.

My own personal interest in acute kidney injury -- there is the concern always with an agent that has diuretic effects that there may be an increased risk of subtle AKI, and we have to worry, as we do with RAS blockade, about potential benefit being offset by those risks. So those are things that really need to be very carefully looked at in order to define the drug as beneficial for kidney disease.

DR. SMITH: Right. And in fairness to the sponsor, this was not intended to be a study on
renal complications and the effects. It's interesting to see the data, and so these criticisms are not one that are false in this study, they're just things that should be addressed if this were pursued in another study.

Dr. De Lemos, were you going to make a comment? No? Other comments on this? Yes, Dr. McBryde?

DR. MCBRYDE: I just wanted to reiterate what Paul said, a couple of things that struck me. One was estimated GFR. The graphs and the statistics look really impressive, but these are estimating equations.

They have inherent inaccuracy and imprecision that you can't resolve to 7 mLs per minute, to 1.73 meters squared. MDRD, even CKD-EPI, you've looking at 20 percent variability with 90 percent confidence. And at an estimated GFR over 70, your error is going to be far greater than what you're resolving in this, and so I think the data is really confusing the way it was expressed.

One thing that I was interested in was the albuminuria issue. What strikes me is it doesn't
look like it's really hemodynamic. There's only
about a 4-millimeter systolic blood pressure
change. This is not the SPRINT trial where you got
20 millimeters of reduction and you would expect to
see a dramatic reduction in albuminuria and it
resolved relatively quickly.

One of the things that intrigued me about
this class of drugs is they are all very highly
protein bound. They are protein bound to albumin,
and with nephelometry, it could actually be giving
you an erroneous reading, and that would be
something I would throw out there.

Certainly calcium, some other drugs have
been implicated in throwing off turbidimetric and
nephelometric measurements. So thinking long term,
you might want to think about something more
sensitive, like HPLC or other radioimmunoassays
rather than this particular approach.

So I am not sure I have any confidence in
the data. It wasn't prespecified, so I don't want
to beat anybody up for it. And as Paul said, I
would love to see something that protects the
kidneys in this population.

DR. SMITH: So I might summarize that the data themselves, even though presented in a simple way, they may show evidence for benefit in terms of renal outcomes, that there are really major deficiencies in the study design and other aspects of the way the data were collected, and the data themselves that provide no confidence in terms of interpreting these data as a solid demonstration of renal benefit.

That clearly would be a very important accomplishment if that were shown in a subsequent study, but it's clear, to the points made by committee members, that really that would need much further study to approach that point. Okay?

So we need to take a break, and when we come back, we'll address the voting questions. And so we'll take I think about a 10-minute break now, if you would come back here within about 10 minutes.

I just again remind committee members that there should be no discussion of the topic or the elements of these considerations among yourselves,
other than in the open forum of the committee.

(Whereupon, at 3:45 p.m., a recess was taken.)

DR. SMITH: All right. So we're going to get started again. I notice in the morning when I say everybody should take their seats, in a couple minutes it gets instantly silent, even though we're not going to start right away. And about this time of day, nobody listens anymore, keep talking, and I don't blame you for that. So anyway, here we go again.

We're now going to proceed with the voting questions. We will be using an electronic voting system for this meeting. Once we begin the vote, the buttons will start flashing on the microphone units, and they'll continue to flash even after you've entered your vote. Please press the button firmly that corresponds to your vote. If you're unsure of your vote, or if you wish to change your vote, you may press the button you want until the vote is closed.

After everyone has completed their vote, the
vote will be locked in. The vote will then be displayed on the screen. The DFO will read the vote from the screen into the record. And next, we'll go around the room and each individual who voted will state their name and their vote into the record. And at that point, you can also state the reason you voted as you did if you want to. And we'll continue in the same manner until we've made it around the room for those comments and all the voting questions have been addressed.

So if there are no questions or comments concerning the wording or the question, we'll now open this first question for voting. I guess I'll read it. Question 6, vote.

Based on data in the briefing materials and presentations of today's meeting, do you believe the EMPA-REG OUTCOME study results have fulfilled the recommendations laid out in the 2008 FDA guidance for industry by demonstrating the use of empagliflozin to improve glycemic control would not result in an unacceptable increase in cardiovascular risk? And if yes, consider your
rationale for your vote. If no, provide your
rationale and comment on what additional data would
be needed.

So any clarification needed before we start
the voting? Okay, so the voting, if we could
activate the microphones.

(Vote taken.)

DR. BONNER: For the record, 23 yes, zero
no, zero abstain.

DR. SMITH: Okay. So we're going to go
around the room. We'll start with Dr. Schambelan,
if you would state your name into the microphone
and your vote for the record. And you may make
comment on why you voted yes.

Since it's 23/0 yeses, it probably doesn't
require a long comment from people. So if you
don't make real long comments, we'll get to the
next voting question and have more time to make
comments about it. Okay? Dr. Schambelan?

DR. SCHAMBELAN: My name is Morrie
Schambelan. I voted yes, and I have no comment.

(Laughter.)
DR. BUDNITZ: Dan Budnitz, voted yes. I'll save my comments for the second question.

DR. COOKE: David Cooke. I voted yes. I think it's clear that they met the criteria.

DR. NEATON: Jim Neaton. I voted yes. It is very clear they met the criteria.


DR. FRADKIN: Judy Fradkin. Also yes. Also ditto.

DR. ROSENBERG: I voted yes. Interesting they made it for the MACE and the MACE-plus despite the limitation for the last component of the MACE-plus.

DR. MCBRYDE: Kevin McBryde. I voted yes for the same reasons previously cited.

DR. EVERETT: Brendan Everett. I voted yes. I think they've demonstrated cardiovascular safety.

DR. SMITH: Robert Smith. I voted yes. And after I told everybody to not say very much, I was impressed not only at meeting the criteria, but in the various ways of reanalyzing the data addressing questionable groups, various sensitivity analyses.
I mean, this effect held up. So we're not moving through this, in my opinion, quickly because we feel like going quickly, I think it was really very clear viewed from multiple potentially critical directions.

DR. THOMAS: Abraham Thomas. I voted yes. And I really agree with Dr. Smith that, of all the things we discussed this afternoon and this morning, this is the clearest.

DR. KONSTAM: Marv Konstam. Yes.

DR. LI-NG: Melissa Li-Ng. I voted yes. And I concur with all the previous comments.

DR. GOOD: David Good, yes.

DR. DE LEMOS: James De Lemos, yes.

MS. HALLARE: Diana Hallare, yes.

DR. PALEVSKY: Paul Palevsky. Yes. No further comments beyond what's been made.

DR. WILSON: Peter Wilson, yes.

DR. HECKBERT: Susan Heckbert, yes. The data were very convincing regarding safety.

DR. YANOVSKI: Susan Yanovski, yes.

MR. LUMLEY: Dan Lumley, patient rep. Yes.
DR. HIATT: William Hiatt, yes. And it was convincing.

DR. PROSCHAN: Michael Proschan, yes. Because, duh.

(Laughter.)

DR. SMITH: You realize that's in the federal record?

(Laughter.)

DR. PROSCHAN: I've done a lot worse than that in the federal record. No problem.

(Laughter.)

DR. SMITH: So I think we'll go to the second voting question. So for this voting question, I will read.

Based on data in the briefing materials and presentation at today's meeting, do you believe the EMPA-REG OUTCOME study results provide substantial evidence to establish that empagliflozin reduces cardiovascular mortality in the population studied? If yes, please provide the rationale for your vote. If no, please provide the rationale for your vote and comment on what additional data would be
needed.

     Any questions, clarification questions

regarding the -- yes, Dr. Hiatt?

     DR. HIATT: I would just like to clarify

with the FDA on this question. Vis-a-vis a meeting
we had a year ago, if we vote yes, and this goes in
the label, and no other drug of this class
substantiates that, you're stuck with that. Right?
You can't withdraw that unless there's a new safety
concern that might change how you feel.

     But I just want to understand, if we put

this in, and based on this single trial, then the
sponsor gets that claim regardless of whatever
happens after that. Is that correct?

     DR. GUETTIER: It's very hard to take a

claim out once it's in the label because the other
trials that are negative may be negative for some
very good reasons. So I think what we're asking
you to opine on today is with the data that you've
heard today, whether or not you think this is
believable enough to be included as a claim in the
label.
DR. SMITH: So I'd like to ask for a little clarification on that because I can read this question one way, which doesn't mention a claim. Or I could read this question in a way that would say, do I think the FDA ought to put this claim in effect. And so I would like to clarify what you're asking us.

DR. GUETTIER: So this question was really meant to actually get at the issue of a claim. It talks about substantial evidence to establish that empagliflozin reduces cardiovascular mortality. And as you heard this morning, substantial evidence is the evidence necessary to form the basis of a new claim.

The new claim the sponsor is seeking, as of today, is a cardiovascular mortality benefit. It wasn't the initial claim that they sought, but this is the new claim. So really what this question is asking you is whether or not the evidence that is provided by this trial is substantial evidence and warrants inclusion in the product label.

DR. KONSTAM: Can I follow that for a
second?  Oops, I voted.  I might not have meant that vote.  So, if you decide not to give it the claim, is it likely, nevertheless, that the data would appear in the label?

DR. GUETTIER:  Basically without a claim, it's unlikely that the data would appear in the label.  The data in the label describes in general -- so section 14 of the label generally describes the trial that supports the indication.  And so if there is no indication, there is no claim.

That being said, this is now a safety trial and we would likely put at least the safety component of the trial.  How we would do that, we haven't really gotten to.

DR. SCHAMBELAN:  Could I ask a question as well?  So since the claim here is for cardiovascular mortality and not for a MACE outcome, are we in a position that we can, if we have concerns about the MACE outcome, that we can ignore that in answering this question?

DR. GUETTIER:  I think that your vote should
be informed by the discussion at this afternoon's session, and I think you've heard some thoughtful comments by other members of the committee regarding how to think about the cardiovascular mortality claim. And so we think that the discussion that occurred this afternoon should at least should have informed how you vote about this.

Again, this is asking about the cardiovascular mortality claim. You've heard both from the FDA presentation, from comments that were made by some of the members that in order for you to look at this particular component of the composite, you have to be convinced that you've won on the primary endpoint, et cetera, et cetera. So those are some of the things that you should be thinking about as you vote.

DR. SMITH: Further questions, clarifying questions.

DR. KONSTAM: You've got two over there.

DR. SMITH: Sorry. Dr. Good?

(No response.)

DR. SMITH: Dr. Li-Ng?
DR. LI-NG: Sorry, a question for the FDA.

So how would you interpret -- so the question that you're posing is, is there substantial evidence to establish that empagliflozin reduces cardiovascular mortality in order to for it to have this indication. So, because there are other ramifications of adding this to the label for this particular class of medications, how would you interpret then a yes versus a no. Or maybe that's not an appropriate question at this point.

DR. GUETTIER: I think you know what we've said early today is that, as important as the vote is in terms of yes or no, is the rationale for why you voted. And so you can basically explain why you voted the way you did. And based on your rationale, we do look at the record and the yes and no is not always black and white.

DR. SMITH: Yes, Dr. Good?

DR. GOOD: So the implication is that we feel that the recommendation would stand on this single trial and would not require a second trial really. That's kind of the implication of this.
DR. SMITH: I would interpret it that way because it's based on the data in the briefing materials as written in this question, yes. Other clarifying questions? Dr. Rosenberg?

DR. ROSENBERG: Yes. It's not a clarification but just confirmation that from the FDA they can always disagree with the recommendation. Right?

(Laughter.)

DR. SMITH: So I think we're ready to go ahead with the vote.

(Vote taken.)

DR. BONNER: For the record, vote question number 7, 12 yes, 11 no, and zero abstain.

DR. SMITH: So again, we'll go around the room. We're going to start this time on the other side of the room. Dr. Proschan, you're first, if you would state your name into the microphone for the record, your vote, and any comments you'd like to make on the basis for your vote.

DR. PROSCHAN: I'm Michael Proschan. I voted yes. I think the evidence was pretty robust.
I feel pretty confident that there's some benefit on cardiovascular mortality. I think it's likely that that 38 percent is an overstatement that's probably biased. But I think, even if it's 20 percent, you know that's still very good.

I would not have voted yes on the original claim they made, but I think, because of the many events, because of the sensitivity analyses, I feel pretty confident that there is some benefit on cardiovascular mortality.

DR. HIATT: William Hiatt. I voted yes. In reviewing the data before this meeting, I was not convinced. And in particular, I was concerned about the idea that the MACE endpoint barely made it in. But that in and of itself would probably not have been a convincing result.

The thing that swayed my thinking about this particular question was that cardiovascular mortality withstood all the sensitivity analyses, including the worst-case missing data analysis of assuming a complete reversal of events between drug and placebo. So with that kind of evidence, it's
hard to explain it away.

I don't understand the mechanism, but it seems that that is a highly clinically important result, and I would hate to sort of deny a claim based on the unique way this trial was run and how this finding evolved, because I believe at the end of the day, I think the result is true.

MR. LUMLEY: I'm Dan Lumley. I voted yes. As I indicated before, I'm the patient rep. I've had type 2 diabetes for 20 years. My disease is under control largely because of drugs, diet, exercise, a Kansas City caring doctor, and recently the Kansas City Royals winning last year's World Series. It made me feel a hell of a lot better.

But that's not the reason I voted yes. The reason I voted yes is, from what I heard today, I think it works, and it's safe.

DR. YANOFSKI: Susan Yanovski, and I also voted yes. I was impressed with the magnitude of the reduction of CVD and all-cause mortality, which I think really represents the hardest of hard outcomes. We were shown statistically strong
evidence of important clinical benefit in a large multi-center study, which was one of the FDA criteria for accepting data from one study.

I do understand the uncertainty that's imposed by having a fairly large proportion of non-assessable CV deaths, but I was reassured by the sensitivity analyses, particularly those showing a significant CV death risk reduction, even after you excluded those deaths.

DR. HECKBERT: This is Susan Heckbert, and I voted no. I agree particularly with the points that Dr. Hiatt brought up that this trial does provide evidence of reduction in cardiovascular mortality. I think it's intriguing and it would really be great if this is, as it appears to be, correct.

But I think the question we're asked to vote on today was whether the available data is adequate to support an additional indication in patients with established CVD to reduce the risk of cardiovascular death.

I believe that given the issues with the
non-assessable causes of mortality, the fact that cardiovascular mortality was not a primary endpoint, and the issues of multiple testing that we've discussed today, and also the fact that the FDA usually requires two well-designed and conducted trials to add this type of an indication, taking all those things together, my opinion is that although these data are intriguing and promising, a second study is needed before this indication would be added.

DR. WILSON: Peter Wilson. I voted no. This is the first compound in this class going for this indication. And I think we should have a higher bar for the quality of the evidence and the importance of the findings, they really be substantiated.

I think it's very hard to go from safety to superiority in one study. I have difficulty with a trial that has nine modifications as it went through the path to eventual conclusion and readjustment of some of the key issues that may have been intervening issues, and especially
concerning because the blood pressure effect, and potentially the heart failure effect, may be the most important pathways. And the heart failure one, we'd all like to see more information and better information.

I am concerned that if approved, the medication might be overused in younger patients at fairly low risk of heart failure and extremely low risk of cardiovascular death. So I think a second study that would really confirm these and firm up who are the candidates would be important.

Then finally, what trial would be the next trial? So I think as people with heart failure or at high risk for heart failure, with a variety of modern approaches and assessments.

DR. PALEVSKY: Paul Palevsky, and I voted no. I can see that the results on mortality are very intriguing. I think that there are a number of issues that have already been discussed about concerns regarding the unable to be classified cardiovascular deaths.

I'm concerned about the fact that we really
don't know what the mechanism is, and we don't know whether there may be other aspects of interventions in these patients that may have been contributing to the apparent benefit from the drug rather than it necessarily being a true drug.

This is a drug that's on the market. Prescribers will have access to the published literature about the drug. But I think that before it's labeled, we have to have confirmatory evidence.

MS. HALLARE: Diana Hallare. I voted no. This is because of the missing data, the unblinding of certain individuals, the uninterpreted deaths, silent MI controversy. And I would like to see more as additional data, a more diverse population in this study, including ethnicity, those with or without heart failure, those at risk for kidney disease, for instance. And I know currently that there is a clinical trial going on for a particular subgroup, and I think that is commendable. And I believe that a second trial would provide more clarity.
DR. DE LEMOS: I'm James De Lemos. I voted yes. I struggled as well with the classification of the cardiovascular death events and the likely misclassification of the events and acknowledge as well that the effect size is likely dramatically overestimated and probably wouldn't replicate to this degree.

But the robustness of the finding, the large number of endpoint events, and then most importantly, the fact that it's buttressed by a reduction in all-cause mortality, the ultimate endpoint to me, drove a narrow decision in favor of approval, for recommendation for approval.

DR. GOOD: David Good. I also came down narrowly on the side of yes, for pretty much the same reasons that already been elucidated by my colleagues in the room.

I felt that the cardiovascular mortality was always reasonable robust. If I had to vote only on the primary outcome measure of 3-point MACE, I would have voted no. I'm also concerned about the mechanism, but that didn't dissuade me from voting
yes after considering this quite carefully.

DR. LI-NG: Melissa Li-Ng. I voted no for similar reasons that have already been stated. For me, as an endocrinologist, the lack of a mechanism does bother me. I was thinking that clinically. I practice in a region where SGLT2 inhibitors were approved before the U.S. FDA approved them, so they've been in use for about three or four years already. And for me to tell a patient that now this has been shown to decrease your risk of dying from all causes based on one study, despite the robust evidence, again, it's very hard for me to be able to say that to a patient without a second study to confirm these findings.

DR. KONSTAM: Marv Konstam. I voted yes. First of all, I want to say to my colleagues on the FDA, good luck in figuring out what to do.

(Laughter.)

DR. KONSTAM: I think that the vote actually is, as it turned out, is very representative of what's been going on in my mind all day long, of going back and forth about what the right thing to
do here is. And I will say that every member to my
right that's voted no, I agree with what they said.
Okay?

(Laughter.)

DR. KONSTAM: So I agree with everybody.

(Laughter.)

DR. KONSTAM: I think that -- so it's tough, because we have a component of an endpoint, which
is a secondary endpoint, that hits its p-value
marginally. And so for that reason, despite, I
think, Dr. Pocock's really excellent presentation,
I just don't know what that p-value really is. And
I think the comment about multiplicity has been
stated a number of times, and so I have trouble
getting there in any kind of rigorous way

So why did I vote yes? I hope and I think I
alluded to some of it in my comments, it's really
pulling it together with three pieces of
information.

One is that the p-value, nominal as it is,
is extremely small. Two is that -- and this hasn't
been discussed quite enough. I think in a way
there are two trials in here, with two different doses, and each of those populations -- those two populations with the two doses behave pretty much exactly the same.

So that influenced me. And I was influenced by the number of events. I think 300 and whatever we said, over 300 events is a substantive finding, and in my mind weighs a little.

I appreciate Bob Temple bringing up the LIFE trial because I think that's kind of a precedent for what you're dealing with here because you have a primary endpoint that's there that's positive, but with a soft p-value and without another trial, and then you went ahead and approved a component of the primary endpoint -- so there's some analogy here at least. Maybe you had no idea what you were doing, but you set a precedent.

(Laughter.)

DR. KONSTAM: So anyway, that was my rationale, and I wound up voting yes.

DR. THOMAS: Abraham Thomas. I voted yes.

I think the data was very convincing. The p-value
for the composite was the one we're all having
doubts about. However, the mortality is very
convincing. And fortunately, the all-cause
mortality is consistent, so this isn't a case where
in some previous trials, we got an improvement in
cardiovascular mortality and all-cause mortality
goes the wrong way because people are driving their
tractors into fjords and random events like that.

I think what's important, though, is -- and
what I was wrestling with is do you require a
second trial or not. And I voted yes thinking that
you probably don't, but there are a lot of
unanswered questions, the mechanism, heart failure,
things that could be studied in additional trials.
And I do echo the point of if you really are a
purist, really in some respects this is a very
enriched population, appropriately so for a
cardiovascular event trial.

How does this translate to the new diabetic?
We really don't have very good tools of changing
mortality in newly diagnosed diabetics. Several
trials have looked at that and saw no difference in
terms of the treatment arm or the placebo arm.

Does this agent add some promise to the newly
diagnosed diabetic? And is this an agent that you
would add even if glucose control is reasonably
under good control?

I mean, these are questions that we'll have
to ask. Is this something that's going to be not
just a diabetes drug, but an add-on for
cardiovascular safety in diabetics? So I think
there's still a lot of unanswered questions that
would require another trial, but not necessarily
for the indication.

It's in the literature, so people know about
this, but we know people don't act on knowledge
even when they're from very well-done large trials.
There's inertia to doing this. However, patients
can only get medications that their insurance
covers.

I have a feeling if there is an indication,
that might drive the coverage of some of these
medications for the insurers, the pharmacy benefit
managers. And if there really is a mortality
benefit like this trial showed, that would be an important thing for our patients, and patients do care if they die or not.

DR. SMITH: I'm Robert Smith. I voted no. It was a very difficult decision. So again, it was for me a narrow split, a very narrow cut between yes and no. Focusing on the mortality data of this trial, I feel that they are very robust. And so as one trial goes, this is a convincing set of data.

But then as I think about not just what do I think about the outcome of this trial, but how confident would I be that a second trial would reproduce these data, I feel enough uncertainty that that's the basis for my no vote.

I think the issues behind that were very well summarized in combination by Drs. Wilson and Palevsky in terms of issues related to the study population, and things about the study itself that we may not appreciate from a single study as important variables, as well as a number of aspects of how the study itself was conducted, which are individual to a study.
So I agree it's a very difficult problem that now goes to the FDA, but I was not comfortable enough that these robust data would be reproduced by subsequent experience to give a yes vote.

DR. EVERETT: Brendan Everett, and I voted yes. I think there was some concern, probably most eloquently expressed by Marv, about the robustness of the win on the primary endpoint, because once you demonstrate that, then you can get through the door and start looking at the components of the primary endpoint.

But I think once you're through that door, which I felt okay that we did get through the door, the effect on cardiovascular mortality is clear and substantial, and as others have mentioned, survives multiple sensitivity analyses.

The thing that really drove this vote for me was the number of events, cardiovascular deaths, and the size of the p-value, and the importance of that particular endpoint. I don't think I would have voted the same way if the endpoint were myocardial infarction, for example.
That said, I struggled, as many others have, in the great tradition of a yes vote meaning no, and a no vote meaning yes for the FDA. I think that, like Dr. Smith mentioned, I think this is one trial. It's not two trials; it's one trial. And that makes me hesitate about the fact that these results really have not been replicated, whether by this agent or by another agent in this class.

As the FDA is turning these comments over, over the course of the next couple of months, I would not object to waiting to give the label until there are a second or third endpoint trial that includes cardiovascular death as one of its key components.

I suspect -- I don't know this for sure -- that the sponsor is planning trials in other select populations that might have a significant benefit from this drug. I certainly hope so because one of the reasons I voted yes is because this is enormously promising, and I think potentially offers a benefit to a huge number of patients who have not, as of yet, had any agent for
managing glucose that reduces their cardiovascular mortality.

DR. MCBRYDE: Kevin McBryde. I voted no. I'd reiterate a lot of the same sentiments that have been discussed earlier. I think what struck me -- I think the MACE 3 endpoint was pretty clear it did not meet the statistical significance that was predefined.

The cardiovascular mortality bothers me a little bit. This was a population of 76 percent at baseline who had evidence of coronary artery disease and yet myocardial infarction didn't pan out as cause of cardiovascular death. And that just set off a little red flag in my mind. Ten percent of the patients were diagnosed with having heart failure at baseline, and yet hospitalizations were much greater for heart failure.

There were some inconsistencies in terms of hazard ratios across regions. Looking at a population -- and again this is only -- I'm speaking purely from the U.S. perspective where 1 in 6 Americans with diabetes are black or
African-American and 5 percent of the study population represent them -- and the data suggested in that small body that there might be an increased risk for that population.

I just think that it was a little bit too much to overlook under the criteria that FDA gave for acceptance of a single trial versus requesting a second. And I think I would feel a little bit better for the U.S. population specifically with a second trial that could support these findings.

DR. ROSENBERG: Yves Rosenberg. I voted no, but also, as many others have mentioned, it was a very difficult decision. Maybe the FDA should have a third category that's maybe.

(Laughter.)

DR. ROSENBERG: Second, I want to thank Dr. Konstam. I'm glad I helped Dr. Konstam reach a decision that's opposite to me by having raised issue of naming the LIFE study that Dr. Temple reported also got to an opposite decision.

Again, the same reasoning that's been mentioned already; it's a secondary endpoint among
many. And although the results are probably true, the p-value is not what's really important. It's the confidence interval. And what am I going to say to a patient? Yes, it will probably reduce your mortality, but maybe it's a little bit and maybe it's a lot. I think your patients deserve better. They deserve more evidence, and that's why a second trial is required.

Again, among the other reasons that were mentioned, it's the first component in class and if we say if it's approved it's probably the end of the story in this class, and really, we need many evidence.

Also, yes, it may not be reimbursed for all its drugs that's still already on the market, but it's available if people really think it will benefit them. But we really need to make sure of the magnitude of benefit, on which patients it really works. The issue of African-American is very important in this country, so we need to answer that question.

Yes, probably the sponsor is not very happy
now, but already probably they're doing studies
there, so hopefully it's not too much of an added
cost for them. And maybe they can talk to the new
leadership of the FDA or talk about ways they can
do their study a little more efficiently than have
been so far. I think that's all.

DR. FRADKIN: I'm Judy Fradkin. I voted no.
I also really struggled because my answer to the
question as to whether there's substantial evidence
would have been yes. I just don't think it's
absolutely proven to the extent that it should get
a label.

I was very impressed by the number of
events, the magnitude of the reduction, the
sustained effect over time, the sensitivity
analyses. I thought all of that was really
compelling.

What I just couldn't get past was sort of my
longstanding belief that a positive outcome to a
secondary outcome is hypothesis generating. And I
wasn't convinced that the primary outcome was
clearly positive given some of the methodologic
issues in the study.

That said, even if some of the other agents in this class didn't show a cardiovascular mortality benefit, I wouldn't say that that excludes this one showing it. I think these drugs differ in terms of their selectivity. And I thought that there was a lot of really promising evidence for this. And I am reassured by the fact that doctors will see this data and make their own decisions, but I did feel that we needed a second study before there's a label.

DR. CHO: My name is Leslie Cho. I voted yes. I voted yes for many reasons. One is that since 2008, the FDA guidance has really led the way from all these diabetic medications to undergo these cardiovascular trials that we've seen time and time again. They've been negative. We've seen that for GLP-1. I mean, we've seen that for DPP-4. We've seen it in 16,000-patient trials, such as SAVOR, where there was absolutely no cardiovascular death benefit on the active arm.

So for this trial to be so positive in terms
of overwhelming cardiovascular death numbers, and
even with the sensitivity analysis to be positive,
that was very convincing for me. And I think as a
cardiologist, it's interesting to note that all the
cardiologists on the panel voted yes, which is kind
of interesting as well. It's just a side note
there.

I don't think we have to -- we've
traditionally thought about cardiovascular endpoint
as athero endpoints, but perhaps it's now time to
move away from athero endpoints to something
different. And yes, it's a mechanism we don't
know, but it's still the number of deaths, the
actual bodies on the ground. It's pretty
convincing.

DR. NEATON: Jim Neaton. I voted yes. An
extreme difference and a very clinically relevant
outcome, with a relatively narrow confidence
interval based on the large number of deaths. An
endpoint, which is very easily ascertainable, less
than 1 percent, were missing. The findings held up
to a worst case analysis very easily.
In this population, one would expect a large number of cardiovascular events. So I thought the two analyses, one where you excluded the non-assessable events, and then the second where you just looked at all-cause mortality, both kind of were supportive and very convincing.

I thought the discussion that we had around some of the non-fatal outcomes, maybe the lesson from this is that we should be doing more trials with cardiovascular all-cause mortality outcomes that are easily ascertainable where the results are very clear-cut when you get them.

DR. COOKE: David Cooke. I voted no. And for me, it also was a difficult decision. And I did go back and forth throughout the day in terms of what I thought made sense.

But it came down to this question of whether this trial should stand alone as the evidence for benefit on cardiovascular mortality for this medication. And I really think the data to support a single trial as evidence of efficacy needs to be very high. And I think just this uncertainty to me
almost defines the fact that that really says we need another trial to clarify this.

Yes, the data on cardiovascular mortality in this trial is very suggestive of a benefit, but it is just a single trial. And kind of folded into that is this issue of mechanism. And I don't feel that we absolutely have to understand a mechanism to accept a benefit, but we do have to have at least an understanding in some ways of who is going to benefit from a medication like this.

The most obvious question with this medication is, what is this impact on heart failure. I think we definitely don't have our hands around that because the study wasn't designed to investigate that sufficiently. So I think we need to understand that.

Or if that's not the answer to how this is benefiting patients, then understand how we can use this to benefit patients if it's supported by a second trial. But I do feel that use of a secondary outcome as the final answer to an efficacy trial is risky.
DR. BUDNITZ: Dan Budnitz. I voted no.

Like David, I had some concerns about this being the single trial in the end. I think there probably is likely a mortality benefit, but had enough concerns that I wanted to go over those four characteristics that were outlined at the beginning of the day about when FDA has relied on a single trial. Included excellent study design, highly reliable, statistically strong, and evidence of important clinical benefit.

Just going through those, I think it was excellent design as a safety study, but I think we heard about -- there's enough concerns that this might not be the ideal design for looking at CV death as a primary endpoint. Is it highly reliable? We heard about issues of multiplicity, the other CV death that was not further described, the numerous changes to protocol that might question the reliability.

The question of statistically strong, I think on the surface obviously it's a very small p-value. But digging into issues of multiplicity
and the methodologic, the statistical methodology
of looking at this secondary endpoint after there
was a barely statistical MACE endpoint, raises
concerns.

Finally, is this an important clinical
benefit? Of course, cardiovascular mortality is,
but for me the question was, is it generalizable to
my patients here in the U.S., because that's what
the indication is for. And I just don't know
because didn't know enough about the patients'
baseline drugs, what usual care entailed over the
course of the study. So I would very much look
forward to a second study with, if not all in the
U.S., with a strong U.S. component.

DR. SCHAMBELAN: I'm Morrie Schambelan. I
voted yes. I think Dr. Neaton expressed most of
the thoughts that I had about this. I did go back
and forth during the day. I wish the FDA good luck
in trying to decide what they're going to do with
our input. concerned, though, that we make sure
that the population who

I am is treated with this drug, if this
indication is approved, are the people that were 
studied in this trial. And I don't think we're 
ready to give it to the 19-year-old diabetic, the 
concerns that Peter Wilson raised.

So I was very, very close to wanting that 
second trial, but I'm so impressed with this impact 
one, the most important outcome, and the fact that 
all-cause mortality was the same, I would hate to 
see this not available to people in the attempt to 
try to keep them alive.

DR. SMITH: Dr. Kewalramani, as the industry 
representative, you don't vote, but do you have any 
comments you'd like to make?

DR. KEWALRAMANI: Thank you. I have no 
additional comments.

DR. SMITH: How about the FDA, any further 
comments from you? Any closing remarks you'd like 
to make?

DR. GUETTIER: I think you've just made our 
job easier. Thank you.

(Laughter.)

DR. GUETTIER: I'd like to extend my thanks
to the DMEP review team for their hard work in preparing for this; for our colleagues from the Division of Cardiorenal Products for their help in preparing for this advisory committee meeting; for the applicant in the weeks leading up to this AC. We did issue a lot of information requests to get to the bottom of some of the issues that were raised in the review, and they were very timely in their responses.

Then for the members of the advisory committee for all your input today, I think it's helpful to us and we'll take everything that you've said today into consideration. So thank you.

Adjournment

DR. SMITH: I, too, would like to thank the committee members for your very thoughtful comments and discussion today. This obviously was a difficult discussion and so special thanks also to the sponsor and to the FDA for all the data you provided to us and all the assistance in our questions. We very much appreciate it and it's really helped with the process, so we appreciate
that.

So with that, this meeting is adjourned.
The committee members, please remember to drop off your name tags at the desk so they can be recycled. And thanks everyone for participating, including the open public hearing speakers. We appreciate your input.

(Whereupon, at 4:43 p.m., the meeting was adjourned.)