1 FOOD AND DRUG ADMINISTRATION

2 CENTER FOR DRUG EVALUATION AND RESEARCH

3

4

5 CARDIOVASCULAR & RENAL DRUGS ADVISORY COMMITTEE

6

7

8 THURSDAY, SEPTEMBER 8, 2011

9 8:00 a.m. to 5:00 p.m.

10

11

12

13 Marriott Inn & Conference Center

14 University of Maryland University College (UMUC)

15 3501 University Boulevard

16 Adelphi, Maryland

17

18

19

20

21

22

Meeting Roster

DESIGNATED FEDERAL OFFICER (Non-Voting)

Kristina A. Toliver, Pharm.D.

Division of Advisory Committee and

Consultant Management

Office of Executive Programs

Center for Drug Evaluation and Research (CDER), FDA

CARDIOVASCULAR AND RENAL DRUGS ADVISORY COMMITTEE

MEMBERS (Voting)

Allan Coukell, B.Sc., Pharm.

(Consumer Representative)

Director, The Pew Prescription Project

Pew Health Group

The Pew Charitable Trusts

Washington, District of Columbia

1 **Sanjay Kaul, M.D.**
2 Director, Cardiovascular Diseases Fellowship
3 Training Program
4 Cedars-Sinai Heart Institute
5 Professor, David Geffen School of Medicine at
6 UCLA Division of Cardiology
7 Cedars- Sinai Medical Center
8 Los Angeles, California
9
10 **Mori Krantz, M.D., F.A.C.C.**
11 Associate Professor, Cardiology
12 University of Colorado
13 Denver Health Medical Center
14 Denver, Colorado
15
16 **Darren K. McGuire, M.D., M.H.Sc.,F.A.C.C.**
17 Associate Professor of Medicine
18 University of Texas Southwestern Medical Center
19 Dallas, Texas
20
21
22

1 **CARDIOVASCULAR AND RENAL DRUGS ADVISORY COMMITTEE**
2 **MEMBER (Non-Voting)**
3 **Jonathan Fox, M.D., Ph.D., F.A.C.C.**
4 ***(Industry Representative)***
5 Vice President, Clinical Therapeutic Area
6 Cardiovascular and Gastrointestinal Diseases
7 AstraZeneca LP
8 Wilmington, Delaware
9
10 **TEMPORARY MEMBERS (VOTING)**
11 **Scott Emerson, M.D., Ph.D.**
12 Professor of Biostatistics
13 Department of Biostatistics
14 University of Washington
15 Seattle, Washington
16
17 **Thomas Fleming, Ph.D.**
18 Professor, Department of Biostatistics
19 University of Washington
20 Seattle, Washington
21
22

1 **Andrei Kindzelski, M.D.**
2 Medical Officer, Program Director
3 Thrombosis and Hemostasis Branch
4 Division of Blood Diseases and Resources
5 National Heart, Lung, and Blood Institute, NIH
6 National Institutes of Health
7 Bethesda, Maryland
8
9 **A. Michael Lincoff, M.D.** ***(Acting Chair)***
10 Vice Chairman, Department of Cardiovascular
11 Medicine
12 Director, C5Research (Cleveland Clinic
13 Coordinating Center for Clinical Research)
14 Professor of Medicine
15 Cleveland Clinic
16 Cleveland, Ohio
17
18 **Debra McCall, B.S., MAM**
19 ***(Patient Representative)***
20 Murrieta, California
21
22

Steven Nissen, M.D.

Chair, Department of Cardiovascular Medicine

Cleveland Clinic Foundation

Heart and Vascular Institute

Cleveland, Ohio

Vasilios Papademetriou, M.D.

Staff Cardiologist

Veterans Affairs Medical Center

Washington, District of Columbia

Philip Sager, M.D.

Pharmaceutical Consultant and Chair, Scientific Programs Committee

Executive Committee Member, Cardiac Safety Research Consortium

San Francisco, California

FDA PARTICIPANTS (Non-Voting)

Norman Stockbridge, M.D., Ph.D.

Director, Division of Cardiovascular and Renal Drug Products

Office of Drug Evaluation I, CDER

Robert Temple, M.D.

Director, Office of Drug Evaluation I, CDER

C O N T E N T S

AGENDA ITEM	PAGE
Call to Order and Opening Remarks	
Introduction of Committee	
Michael Lincoff, M.D.	10
Conflict of Interest Statement	
Kristina Toliver, Pharm.D.	13
Opening Remarks	
Norman Stockbridge, M.D., Ph.D.	17
Sponsor Presentation - Johnson & Johnson Pharmaceutical Research and Development, LLC	
Introduction	
Gary Peters, M.D.	26
Medical Landscape & Study Design	
Kenneth Mahaffey, M.D. .	31
Efficacy	
Robert Califf, M.D.	45
Safety	
Christopher Nessel, M.D. .	58
Key Issues, Benefit Risk and Conclusions	
Robert Califf, M.D.	64

C O N T E N T S (continued)

AGENDA ITEM	PAGE
Clarifying Questions for Sponsor Presenters	117
FDA Presentation - NDA 202439	
Dose Selection	
Preston Dunnmon, M.D.	178
Issues Affecting Interpretation of the Efficacy Data	
Martin Rose, M.D., J.D.	190
Clarifying Questions for FDA Presenters	227
Questions to the CRDAC and Discussion	279
Adjournment	417

P R O C E E D I N G S

(8:01 a.m.)

Call to Order and Opening Remarks

Introduction of Committee

DR. LINCOFF: Good morning. I would first like to remind everyone present to please silence your cell phones, Blackberries, and other devices if you've not already done so. I'd also like to identify the FDA press contact, Ms. Sandy Walsh. If you're here, please stand. Thank you.

My name is Mike Lincoff. I'm the acting chair of the Cardiovascular and Renal Drugs Advisory Committee. I will now call the meeting of the Cardiovascular and Renal Drugs Advisory Committee to order. We'll go around the room and identify ourselves. We'll start with the FDA, and Dr. Stockbridge to my left, and go around the table.

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

DR. FLEMING: Tom Fleming, University of

Washington.

DR. SAGER: Philip Sager, consultant and chair of the Scientific Programs Committee of the Cardiac Safety Research Consortium.

DR. NISSEN: Steve Nissen, Cleveland Clinic.

DR. MCGUIRE: Good morning, Darren McGuire, UT Southwestern Medical Center in Dallas.

DR. KRANTZ: Good morning, Mori Krantz, a cardiologist, University of Colorado.

DR. TOLIVER: Kristina Toliver, designated federal officer, Cardiovascular and Renal Drugs Advisory Committee.

DR. LINCOFF: Mike Lincoff from the Cleveland Clinic.

DR. KAUL: Sanjay Kaul, cardiologist, Cedars-Sinai Medical Center in UCLA.

MR. COUKELL: Good morning, Allan Coukell. I'm a pharmacist with the Pew Health Group and the consumer representative to the committee.

MS. MCCALL: Good morning. Debbie McCall. I'm an AFIB patient.

DR. EMERSON: Scott Emerson,

biostatistician, University of Washington.

DR. FOX: Jonathan Fox. I'm a cardiologist with AstraZeneca in clinical development. I'm the industry representative to the committee.

DR. KINDZELSKI: Andrei Kindzelski, medical officer, NHLBI, NIH.

DR. LINCOFF: 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 Chair. We look forward to a productive meeting.

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

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

Now, Kristine Toliver will read the conflict of interest statement.

Conflict of Interest Statement

DR. TOLIVER: The Food and Drug Administration is convening today's meeting of the Cardiovascular and Renal Drugs 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, SGEs, or regular federal employees from other agencies, and are subject to federal conflict of interest laws and regulations.

The following information on the status of this committee's compliance with federal ethics and conflict of interest laws, covered by but not

limited to, those found at 18 U.S.C., Section 208 and Section 712 of the federal Food, Drug, and Cosmetic Act 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 the federal ethics and conflict of interest laws. Under 18 U.S.C., Section 208, Congress has authorized FDA to grant waivers to special government employees and regular federal employees who have potential financial conflicts when it is determined that the agency's need for a particular individual's services outweighs his or her potential financial conflict of interest. Under Section 712 of the FD&C Act, Congress has authorized FDA to grant waivers to special government employees and regular federal employees with potential financial conflicts, when necessary, to afford the committee essential expertise.

Related to the discussion of today's meeting, members and temporary voting members of this committee have been screened for potential

financial conflicts of interest of their own, as well as those imputed to them, including those of their spouses or minor children, and, for purposes of 18 U.S.C. Section 208, their employers. These interests may include investments, consulting, expert witness testimony, contracts, grants, CRADAs, teaching, speaking, writing, patents and royalties, and primary employment.

Today's agenda involves the discussion of the new drug application, NDA202439, rivaroxaban tablets, submitted by Johnson & Johnson Pharmaceutical Research and Development, LLC, on behalf of Ortho-McNeil-Janssen Pharmaceuticals for the prevention of stroke and systemic embolism, blood clots other than in the head, in patients with nonvalvular atrial fibrillation, abnormally rapid contractions of the atria, the upper chambers of the heart.

This is a particular matters meeting, during which the specific matters related to NDA202439, rivaroxaban tablets, will be discussed. Based on the agenda for today's meeting and all financial

interests reported by the committee members and temporary voting members, no conflict of interest waivers have been issued in connection with this meeting. To ensure transparency, we encourage all standing committee members and temporary voting members to disclose any public statements they have made concerning the product at issue.

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

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

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

DR. LINCOFF: We'll now proceed with the sponsor presentations. Both the Food and Drug Administration and the public believe in a transparent process for information gathering and decision making. To ensure such transparency of the advisory committee meeting, FDA believes it is important to understand the context of an individual's presentation.

I'm sorry. I've moved ahead.

First, we'll start with Dr. Stockbridge's opening remarks.

Opening Remarks

DR. STOCKBRIDGE: Actually, I think Dr. Temple and I both have a couple of words to say. I'll be very brief.

Welcome back to many of you, and welcome aboard to some of you new to the committee. I wanted to emphasize that no regulatory decision has

1 been made on this application. And I'd
2 characterize our internal discussions so far as
3 being fairly ambivalent on it. So what the
4 committee does today is likely to have considerable
5 influence on us.
6 No one, I believe, doubts that rivaroxaban
7 is effective in preventing stroke in patients with
8 AF; that is, what the result would be of a placebo-
9 controlled trial. Nor are there safety issues that
10 would cause one to question that the stroke-
11 reduction benefit was, on the whole, worthwhile.
12 But a placebo-controlled trial was not possible
13 because, of course, several treatments are
14 available that prevent irreversible harm in this
15 setting, making a placebo control unethical.
16 By at least one analysis, rivaroxaban was
17 superior to warfarin in the ROCKET AF study.
18 Nevertheless, the clinical reviewers, from whom
19 you'll hear later today, questioned whether the
20 study can be interpreted as showing that
21 rivaroxaban is even as good as warfarin.
22 You'll be asked what you think about that

and how good a new drug for stroke prevention needs to be in 2011. You're also going to hear about whether adequate instructions for use exist for rivaroxaban, a concern that stems from the strokes observed in patients transitioning from rivaroxaban to warfarin at the end of the ROCKET study.

Late in the day, you're going to be asked to vote whether you think that rivaroxaban should be approved for the treatment of AF. This is always a dramatic moment, and the drama is heightened by the formal process of voting that we implement to keep you from influencing your colleagues around the table. The press is going to make much of what your vote is.

As those of you who are veterans of the committee know, we don't actually care how you vote. We do care about why you vote the way you do, and how you interpret the findings, and what you value in coming to your decision. So we would ask that as you explain what your thought process is, you make sure that we understand all of that.

So, again, thank you all for coming, and we

1 look forward to an interesting discussion.
2 Bob?
3 DR. TEMPLE: Yes. Let me welcome everyone,
4 too. We care a little bit about your vote, but as
5 Norm says, more about the reasoning behind it.
6 I'll try not to repeat much of what Norm said. I
7 mean, in some sense, the drug we're looking at here
8 and the ROCKET study that supports it is relatively
9 straightforward. They're familiar to us because
10 other studies of other drugs have been carried out
11 with sort of similar results.
12 I wanted to say a little bit about non-
13 inferiority studies, and the designs, and what
14 we've done over the years to say what the standards
15 should show. So ROCKET AF was designed as a non-
16 inferiority study. That means it was intended to
17 exclude a loss of the effect of the control agent,
18 warfarin; loss of too much of the effect of the
19 control agent, warfarin.
20 In this case, because the effect of warfarin
21 is very large, the non-inferiority margin was quite
22 large. We agreed, in advance, that the study would

need to show that it preserved not just any of the effect of warfarin, but 50 percent of it.

People are sometimes curious about the fact that non-inferiority studies allow the loss of as much as 50 percent of the enormously-valued effect of the control agent. But the reason is much practical as anything else. If you try to design studies that preserve all of let's say, 10 percent -- or show that you've ruled out a loss of 10 percent, you get studies in the hundred-thousands and things like that, and so we have done that.

It is worth remembering that when you rule something out statistically, what you've ruled out is not the only measurement you have. When you show the drug is effective compared to placebo at .05, you've merely shown that it has some effect greater than 0. In fact, people look at the point estimate of the effect and other things.

So in a non-inferiority study, when you rule out a loss of 25, 50, or whatever percent it is, you also have the point estimate of the two drugs

1 effects to look at. And as a practical matter,
2 non-inferiority studies frequently give point
3 estimates that are right near 1, or, as in the case
4 at least by some analyses, better than 1 here.
5 The non-inferiority guidance that we wrote
6 and the whole concept of non-inferiority studies
7 is, in some way, a reflection of the Clinton-Gore
8 document that's going to be discussed later. That
9 document, which was actually designed to
10 communicate to the outside world that we were not
11 imposing relative effectiveness requirements on
12 things, made an exception. It said when the value
13 of an existing drug is very great, like it improves
14 survival or something like that, we really do care
15 about how a new drug compares to that drug. And in
16 some sense, the non-inferiority approach, which has
17 greatly expanded since Clinton and Gore were in
18 office because there are so many drugs that clearly
19 work and have major effects, is, in some ways, the
20 manifestation of how we implemented that.
21 So it's worth looking at the standards we
22 impose and the expectations to see how we have

manifested it -- I mean, in a certain sense, proven that a new drug is no worse than previous drugs; that it's equivalent. The only way to absolutely show that is to show that it's superior. That has not, historically, been the requirement. All these are matters that one could realistically discuss.

I guess the only other thing I'd point out is that apart from the magnitude of the effect here -- oh, sorry. I wanted to say one other thing. You can't have all this discussion without being aware, at least as a background matter, that there's a drug recently approved for this same use that was shown to be superior to warfarin, and that is plainly on people's minds.

What I think the main effect of this is -- before dabigatran, if someone had come to us with a drug that was reasonably close to warfarin, even if it wasn't necessarily just as good, we would have said, oh, that's fine, because a lot of people have a lot of trouble taking warfarin. And I don't think there would have been that much worry about exactly how effective it was. I think the

1 existence of a drug that's clearly just as good
2 changes that, so that there is now interest in
3 whether the new agent is reasonably close to
4 warfarin, whether that's been manifested. And as
5 you'll hear -- I'm not going to get into
6 that -- there's concern in this study about whether
7 warfarin was used fully, always a question in a
8 non-inferiority study, because one of the crucial
9 assumptions of a non-inferiority study is the so-
10 called constancy assumption, namely, that the drug
11 is having the effect that you calculated that it
12 had when you set the non-inferiority margin before.
13 So, in this case, as will be discussed at
14 great length, whether warfarin was used reasonably
15 well is going to be an important issue and that
16 should not surprise anybody. The other question
17 you will hear -- this also is related to what the
18 primary endpoint was. The primary endpoint here
19 was the as-treated endpoint. Then on that
20 endpoint, the new drug looks quite good compared to
21 warfarin as used. When you add in events that
22 happened shortly after discontinuation, it doesn't

1 look quite as good, and there was an increase after
2 the drug was stopped in certain events. And that
3 needs to be looked at closely because people do
4 stop therapy.
5 So that, to my eye, apart from the magnitude
6 of the effect, what happened there is the second
7 major one of the two major issues. You will hear
8 great attention to large numbers of other issues,
9 and I don't want to anticipate all that, but we
10 welcome your comments, and thanks for coming.
11 DR. LINCOFF: If my apparent dyslexia is not
12 interfering, it looks like the next point on this
13 agenda now is the sponsor presentations. So both
14 the FDA and the public believe in a transparent
15 process for information gathering and decision
16 making. To ensure such transparency of the
17 advisory committee meeting, FDA believes that it's
18 important to understand the context of an
19 individual's presentation.
20 For this reason, FDA encourages all
21 participants, including the sponsor's non-employee
22 presenters, to advise the committee of any

financial relationships that they have with the firm at issue, such as consulting fees, travel expenses, honoraria, and interest in the sponsor, including equity interests and those based upon the outcome of the meeting. Likewise, FDA encourages you, at the beginning of your presentation, to advise the committee if you do not have any such financial relationships. If you choose not to address this issue of financial relationships at the beginning of your presentation, it will not preclude you from speaking.

So now we'll begin with the presentations from the sponsor.

Sponsor Presentation - Gary Peters

DR. PETERS: Good morning, Mr. Chairman, members of the committee, ladies and gentlemen. My name is Gary Peters. I'm the cardiovascular development vice-president for Johnson & Johnson Pharmaceutical Research & Development. Today, it is my privilege to presentation rivaroxaban, a novel oral anticoagulant.

The schedule for the upcoming presentation

is displayed on this slide. Two of the presenters, Drs. Kenneth Mahaffey and Robert Califf, are external experts in cardiovascular disease from Duke University. Both were on the executive committee for the ROCKET AF study, with Dr. Califf co-chairing this committee. Other external experts in hepatology and hematology are also present today and available to discuss specific questions as they arise.

The proposed indication under consideration today is that rivaroxaban is indicated for the prevention of stroke and systemic embolism in patients with nonvalvular atrial fibrillation. Our presentations will show that rivaroxaban should be approved for this indication for the reasons stated on this slide.

ROCKET AF was a well-designed and executed double-blind, global clinical study with an intentional focus on higher stroke-risk patients. The study results robustly demonstrate the non-inferior efficacy of rivaroxaban compared with warfarin, and there were fewer primary endpoint

events while on rivaroxaban treatment. The safety evaluation of rivaroxaban was also favorable, with no differences in overall discontinuations or bleeding events.

Finally, the clinically most important types of bleeding events, intracranial hemorrhage and those with a fatal outcome, were reduced with rivaroxaban. As you have heard, the sponsor and FDA agree on these efficacy and safety analyses.

The important points we will review today are summarized here. First, the impact of the observed 55-percent warfarin group INR time in therapeutic range, TTR, on the study results is limited for the following reasons. TTR is a biomarker that estimates the daily INR value between two actual measurements for each patient. It has high variability of both the patient and the center level and has not been validated for cross-study comparisons. The TTR differences between ROCKET AF and other recent studies are explainable by the different distributions of patient characteristics and country enrollment.

There is little relationship between TTR and the treatment effect at either the center or the region level. In addition, the observed risk-adjusted event rates for the warfarin group in ROCKET AF were comparable to all recent studies. This indicates well-managed warfarin therapy.

The increase in post-treatment thrombotic events at the end of the study was related to the double-blind study process for transition to open-label therapy and can be avoided in clinical practice by allowing overlap of rivaroxaban with warfarin therapy. The data supporting these statements will be presented in detail by Dr. Califf.

Rivaroxaban is an oral direct factor 10A inhibitor. Factor 10A catalyzes the conversion of prothrombin to thrombin. Rivaroxaban selectively and directly acts on factor 10A without the requirement for any co-factors. Rivaroxaban inhibits not only free factor 10A, but also inhibits clot-associated factor 10A. By virtue of where it acts in the coagulation cascade,

rivaroxaban inhibits thrombin generation, as opposed to blocking the activity of thrombin that is already present.

Rivaroxaban has predictable pharmacokinetic and pharmacodynamic properties that have been well characterized in dedicated phase 1 studies and phase 2, 3 population modeling. Rivaroxaban has high oral bioavailability when administered with food and multiple routes of elimination. It has limited potential for clinically important drug-drug interactions with very few restrictions applied in the phase 3 study. These properties support once-daily oral dosing without routine coagulation monitoring, as studied in ROCKET AF.

Rivaroxaban is being jointly developed by Bayer Healthcare and Johnson & Johnson Pharmaceutical Research and Development in the five indications listed here: the prevention of venous thromboembolism in patients undergoing hip and knee replacement surgery, which received approval in the United States on July 1st of this year; secondary prevention of myocardial infarction, stroke, and

death in patients after an acute coronary syndrome; prevention of venous thromboembolism in patients with acute medical illnesses; treatment and secondary prevention of venous thromboembolism; and for stroke and embolism prevention in atrial fibrillation.

The primary focus of today's discussion will be on the ROCKET development program in patients with atrial fibrillation. Over 60,000 patients will have participated when these studies are finished.

Now, I'd like to introduce Dr. Kenneth Mahaffey. Dr. Mahaffey is a cardiologist at Duke University with extensive experience in cardiovascular study design and conduct. He will present an overview of the current state of thrombosis, prophylaxis, and atrial fibrillation, and the ROCKET AF study design.

Dr. Mahaffey?

Sponsor Presentation - Kenneth Mahaffey

DR. MAHAFFEY: Thank you, Gary.

Mr. Chairman, ladies and gentlemen, good

morning. My full financial disclosures can be found on the website www.dcri.org. Relevant to today, my institution, DCRI, has received research grants from J&J for the ROCKET AF trial, and I am being compensated for my time. I have no equity in the companies, nor do I have any benefit based on the outcome of these proceedings.

It's going to be my pleasure this morning to briefly review for you the medical landscape of atrial fibrillation, and then describe for you the ROCKET AF study design on behalf of the ROCKET AF investigators.

Millions of patients in the United States and around the world have atrial fibrillation. And as our populations continue to age, the prevalence will increase. There is a clear association with increased risk of stroke in atrial fibrillation. And anticoagulant therapy has shown to be able to lower the risk of stroke. Unfortunately, many patients are not treated with this agent or treated optimally. Novel oral anticoagulants may offer stroke protection comparable to the vitamin K

antagonists, but may also have additional clinical benefits.

Twenty percent of all strokes occur in patients with atrial fibrillation. And, unfortunately, these strokes are associated with substantial morbidity and mortality, both acutely and in long-term follow-up. It is believed to be due, in part, to the fact that strokes associated with atrial fibrillation are larger and more frequently have hemorrhagic transformation.

Cardiovascular societies have published guidelines with recommendations about therapies for thromboembolic prophylaxis in patients with atrial fibrillation. For patients without risk factors for stroke, aspirin is recommended. For patients with at least one moderate risk factor, aspirin, warfarin, or now, alternatively dabigatran is recommended. In patients at higher risk for stroke, warfarin or dabigatran is recommended in these guidelines.

The benefits of warfarin are well recognized. These data from Singer and colleagues

in the ATRIA study show, very importantly, however, that the net clinical benefit associated with warfarin is particularly prominent in patients with higher risk for stroke, here shown by the CHADS2 score, with CHADS2 scores of 2, 3, or higher.

As I've mentioned, unfortunately, patients are not often treated with warfarin when they are candidates to be treated. And when warfarin is used, it's suboptimally managed. There is much literature published in this regard. These data, from a systematic overview of U.S. experiences, published by Baker and colleagues two years ago, show that when warfarin is managed by anticoagulation clinics, in the top part of this slide, only 63 percent of the time is the INR in the therapeutic range of 2.0 to 3.0.

More importantly, in community-based practices that manage warfarin without the use of anticoagulation clinics, the time in therapeutic range is only 51 percent, with an overall experience reported in the United States from these data of 55 percent, with the 95-percent confidence

1 intervals ranging from 51 percent to 58 percent.
2 On this backdrop, the ROCKET AF trial was
3 designed with the primary hypothesis of showing
4 that rivaroxaban was non-inferior to warfarin in
5 the prevention of the composite endpoint of stroke
6 and non-CNS embolism in subjects with nonvalvular
7 atrial fibrillation. By design, it was intended to
8 recruit patients at high risk for subsequent
9 strokes. Patients were recruited if they had had a
10 prior history of stroke, TIA, or non-CNS embolism.
11 If they had not had one of those prior events
12 previously, they needed to have at least two of the
13 other four risk factors shown: heart failure,
14 hypertension, older age, or diabetes.
15 The enrollment of patients without a prior
16 stroke, TIA, or non-CNS embolism and only two of
17 these risk factors was capped at 10 percent in each
18 region, again, to enhance the risk of the patient
19 population. The patients that were recruited were
20 then randomly assigned double-blind and
21 double-dummy therapy. They were randomized to
22 rivaroxaban at a dose of 20 milligrams once a day

with a down titration of the dose in patients with moderate renal insufficiency, defined as a creatinine clearance of 30 to 49 milliliters per minute. Patients assigned warfarin were managed with target INR of 2.5 with a range of 2.0 to 3.0.

These patients were all followed monthly. Recommendations for therapies associated with atrial fibrillation were not provided in the protocol, but sites were recommended to adhere to standard practice guidelines and local practice patterns. The primary endpoint was stroke and non-CNS embolism, and the principal safety endpoint was major and non-major clinically relevant bleeding. All of these were adjudicated by a clinical events committee.

By design, patients who are randomized and began receiving double-blind study drug therapy were to be followed to the end of the trial. During the period of the study drug therapy, patients were seen every four weeks for mandatory INR monitoring, assessment of adverse events, and ascertainment of clinical endpoints. This was to

1 continue until the point of site notification.
2 This date was the date that the sites were
3 notified that the requisite number of events had
4 been accrued in the trial, and the sites were told
5 to bring the patients back for an end-of-study
6 visit. At this end-of-study visit, patients were
7 then stopped on the study drug, transitioned over
8 standard of care for anticoagulant therapy for
9 their atrial fibrillation, based on institutional
10 policies and guidelines. Following this
11 transition, patients were then seen for a final
12 contact, and adverse events and clinical events
13 were also ascertained at that time, after study
14 drug had been discontinued.
15 Some patients, as you'll see from
16 Dr. Califf's presentation, did discontinue study
17 drug therapy before the end of the trial. If these
18 patients were to discontinue study drug therapy,
19 they came in for a study visit and had a similar
20 transition to standard of care. Importantly, even
21 though these patients stopped study drug before the
22 end of the trial, they continued to be followed

every three months for collection of adverse events and clinical events off study drug. At the data-site notification, they were also contacted for a final contact and ascertainment of events.

In order to maintain the study blind and rigorously manage INRs during the trial, a modified commercially available point-of-care device was provided to each site, and they were educated in its use.

On the left-hand side of the diagram is a subject that was randomized to warfarin, and on the right-hand side of the slide is a subject that was randomized to rivaroxaban. For both patients, a blood sample would be obtained. The point-of-care device would be used, and instead of an actual INR being provided, an encoded value would be given to the clinician or clinical investigator.

This value would then be called into a central computer by the telephone. Additional information would be provided and then the computer would give the investigator either a true INR if the patient was on warfarin, or a sham INR if the

patient was on rivaroxaban. Obviously, the investigator did not know if it was a true INR or a sham INR, so they adjusted the warfarin or the warfarin placebo appropriately, based on that INR value.

The time in therapeutic range is a key issue that Dr. Peters has identified and you've seen in your briefing documents. I'd like to review with you what we did for the sites in terms of educating them about the importance of quality INR management.

Again, the target INR was 2.5. No specific warfarin dosing algorithm or protocol was provided to the sites for recommendations about warfarin dosing adjustment based on specific INRs. However, we provided comprehensive education and instruction to the sites about the need for quality INR control. Each site was instructed to manage the INR according to their local practice patterns and standards, both for dose adjustment of warfarin and scheduling follow-up visits to re-check the INR after a dose changes.

Throughout the conduct of the study, we continued to educate investigators at investigator meetings, through newsletters, and study materials about the importance of the target range of 2.0 to 3.0. As you've heard, we had monthly visits, and we only gave 35-day drug supply at those visits to keep the patients coming back to the clinics. Importantly, there was unblinded assistance that was available if patient safety was an issue.

For the ROCKET AF trial, the warfarin event rate was estimated to be 2.3 events per 100-patient years based on trials that were available at the time ROCKET AF was planned, as well as an adjustment for the expected high-risk patient population. The non-inferiority margins are shown. This was based on preservation of 50 percent of the warfarin effect used in the heart overview. There is modest differences in the margins because of different analytic techniques. The FDA used a log hazard ratio scale evaluation. We will show the 1.38 margin in our figures.

A type I error was as shown. It was

1 estimated that 14,000 patients and 405 events in
2 the per-protocol population would be needed. The
3 power was over 95 percent, regardless of the margin
4 used, or 90 percent.
5 As you have seen in your FDA and sponsor
6 briefing documents, there were a number of analyses
7 performed on specific patient populations and
8 observation periods, and I'd like to walk you
9 through these. The intention-to-treat population
10 was obviously all those subjects who were
11 randomized to therapy, 14,264. The safety
12 population, frequently called a modified intention-
13 to-treat population, were all participants in the
14 ITT population who received at least one dose of
15 study drug. Only 28 patients in the overall trial
16 did not receive at least one dose of assigned
17 therapy. Then the per-protocol population was all
18 patients in the safety population who did not have
19 any major pre-defined protocol violations. An
20 additional 182 patients were excluded based on the
21 protocol violations and other factors. And,
22 overall, 98.5 percent of the ROCKET AF population

was included in the per-protocol population.

In terms of observation periods, on-treatment was defined as time on therapy up to the last dose of study drug, plus two days to account for continued pharmacological effect of the study drugs after discontinuation.

Up-to site notification. This included events up to the date the sites were notified that the requisite number of endpoints had been met.

Up-to the follow-up visit. These were events that occurred during study drug and off study-drug administration up to the follow-up visit that was about 30 days after the last dose of study drug.

In the statistical analysis plan for ROCKET AF, there was a prespecified, closed-loop hierarchical testing procedure, as shown here. The primary analyses was non-inferiority. Primary endpoint was stroke and non-CNS embolism in the per-protocol population that I've described while on treatment, a conservative approach for assessing non-inferiority.

If non-inferiority was achieved, superiority testing would proceed in the following sequence: the primary endpoint and the safety on treatment population, followed by a composite endpoint that included vascular death in the same population and observation period, another composite endpoint that now included myocardial infarction as well in the same population and observation period, all-cause mortality in safety on-treatment population, and then all-cause mortality in the classic ITT analyses.

Additional analyses were performed, including the primary efficacy endpoint in the ITT population in the observation periods that I've discussed. Key secondary outcomes were also evaluated in the ITT population as shown, the individual components of the primary and key secondary endpoints, and then I'll note, although not prespecified in the statistical analysis plan, the primary efficacy endpoint in the ITT population while on treatment.

Mr. Chairman, in summary, ROCKET AF was a

large, global, double-blind clinical trial. Practice guidelines and local standards of care drove therapy for the management of these patients with atrial fibrillation. By intent, a high-risk population with other co-morbidities was to be enrolled, and there was rigorous ascertainment of clinical events and adverse events, both while on treatment and off treatment.

Multiple analyses were prespecified across different patient populations and observation periods. We used hierarchical testing to preserve the type I error. We used the per-protocol population as the generally accepted standard for non-inferiority testing.

We did analyses in the ITT population through the end of the study and recognized this as the standard for superiority testing. However, despite the potential analytic limitations and the biases that can be introduced, we did perform on-treatment analyses due to the expected high rate of discontinuations with the long-term follow-up and co-morbidities in these patients.

1 Thank you very much for your time. I'd like
2 to now introduce Dr. Robert Califf, vice-chancellor
3 of clinical research, Duke University, and co-
4 principal investigator of the ROCKET AF trial.
5 Rob?
6 **Sponsor Presentation - Robert Califf**
7 DR. CALIFF: Thanks, Ken.
8 It's a privilege to be here today. I'd also
9 like to acknowledge, as Dr. Mahaffey did, that my
10 relationships with industry can be found at the
11 website www.dcri.org. As with Dr. Mahaffey, our
12 institution has received grants and contracts from
13 Johnson & Johnson to conduct this study. I've also
14 received consulting fees. I'll note that all of my
15 consulting fees are donated to not-for-profits, but
16 still constitute a potential conflict that should
17 be acknowledged.
18 Also, in addition, because of my job, our
19 institution receives numerous grants and contracts
20 from almost every competitor in this field. And so
21 that should also be acknowledged and can be
22 recognized on our website for public display. As

with Dr. Mahaffey, I don't believe the outcome of these proceedings impact me in any financial way, and I own no stock in either of the companies that are directly involved.

So now, let's move to the efficacy results of the trial. And it's a privilege to present these on behalf of the executive committee, the steering committee, the many investigators, and participants who volunteered for this research.

As has already been noted, there is substantial agreement between the FDA and the executive committee of the trial about the computations for the results of the trial. And so this will be a relatively abbreviated presentation compared to what might be presented if the fundamental results in the computations were in dispute. We want to preserve most of the time for the discussion of the key issues that were identified by the FDA.

So the key points I'm going to make in this part of the presentation, first, we found -- and this represents a view of the executive committee,

as presented in the New England Journal, a publication that came out last month, that rivaroxaban is not inferior to warfarin for all populations and observation periods. It's superior to warfarin while on treatment in all populations, as Dr. Mahaffey reviewed with you the method of analysis. There is a substantial reduction in hemorrhaging strokes in particular, an important endpoint that we will review in the last part of the presentation. And the results are consistent across subgroups and for secondary endpoints.

Two points to make on this slide that shows the overall number of sites, over a thousand, and number of countries. I'll refer to the FDA briefing book about the discussion with the sponsors about including a high-risk population. We were not involved in that discussion as an executive committee, but I can assure you, when we were given the opportunity to study a high-risk population that was global in nature, this was something we were very excited about and that we felt was much needed.

So, in designing the study, in thinking about how the results would come out, we carefully considered the need for a global clinical trial that would be relevant in the United States, but also in other parts of the world, and that would truly involve the population in greatest need of treatment for atrial fibrillation to prevent stroke and systemic embolism.

This is a look at the study populations in more detail. Dr. Mahaffey has reviewed this with you already, but when broken down by treatment assignment, there are slight imbalances but nothing significantly different between the two groups in terms of interpretation of the trial. We'll also note here in the footnote that appears in several slides, that one site was disqualified due to violations in the protocol.

We indeed succeeded in enrolling a high-risk population. This gives the key risk factors, according to treatment assignment, in the randomized trial. Over 60 percent had heart failure. The vast majority had a history of

hypertension. Fully 43 plus percent of the patients were over age 75. Many had diabetes, and, importantly, over half the patients already had a history of stroke, TIA, or non-CMS systemic embolism, a very high-risk population, compared to other trials.

Indeed, if we look at the CHADS2 score, which I know you're all familiar with either from previous work or the briefing book, as you remember, we attempted to cap the CHADS2 score of 2 at the 10-percent level. We didn't quite make that; 13 percent of the patients in the trial had a CHADS2 score of 2. But you can see, we had a very large number of CHADS3, 4, and 5 patients in this trial. So we succeeded in getting what we believed was a high-risk population, much in need of an effective treatment. It overlaps with the previous trials but is very much skewed towards higher risk.

If we then look at the patient disposition in more detail in this slide, I'll call your attention to several key points. First, on the two outside boxes, you can see that just over a

1 thousand patients assigned to each treatment
2 discontinued study drug and follow-up. Now, over
3 half of these in each group died as a reason for
4 that. You might argue it's easy to follow patients
5 who have died. That was the reason in over half.
6 Then, if we look, there were a large number
7 that withdrew consent, increasingly a problem in
8 global clinical trials, especially in the United
9 States. We made every effort to follow the
10 patients as much as we possibly could, given the
11 legal issues involved there.
12 If we look at the two inside boxes, just
13 over 1,300 in each group completed the study off of
14 assigned study drug. And Dr. Mahaffey has already
15 reviewed with you, in great detail, the efforts we
16 went to, to make sure we were ascertaining events
17 in this population.
18 Now, what about adherence to study
19 medication? Now, I'll make several points with
20 this slide. The first is that, as expected in
21 chronic trials, there was an early larger rate of
22 study drug discontinuation followed by a relatively

constant rate. There was absolutely no difference between the two treatments. One might argue -- and as an executive committee, we do believe this -- the disadvantage of rivaroxaban somewhat -- since in a double-dummy trial, there are many things patients had to do, that they would not have to do with a non-monitored once a day drug.

On the other hand, the key point is that there was no difference in discontinuations, and if we look at the first year, it was about 22 percent in each group. And if we look at the second year, it's summated at about 35 percent by Kaplan-Meier plot in each group.

Now, the primary result of the trial, as specified in the protocol as shown here, is a hazard ratio of .79, with a very small p value for non-inferiority. If we then look at the multiple ways of analyzing these data for sensitivity that were done and in the closed hierarchical testing procedure, I would call your attention to several key points. I'm somewhat hampered by not having a

pointer for this presentation podium, so I'll have to describe it to you and hope that you can follow it.

Most importantly, the per-protocol on treatment at the top was a prespecified primary analysis. As we go down the list of analysis populations and durations of follow-up, one can see the point estimate gets closer to 1, but never crosses the 1 line, and the outer estimate of the confidence interval never even comes close to the 1.38 margin specified by the FDA.

Perhaps the most important second point to look at on the entire slide is ITT, the site notification, the next to the last analysis. This most closely approximates what has been done in the other trials, who essentially stopped follow-up at the point of the last -- at the taking of study drug by the patient population when the study was completed. The very final point, ITT, regardless of treatment exposure, includes those 30 days.

The only problem with the pointer, it was pointed out to me, is some people are looking

backwards, but I'll do my best here. If you look at this last point, this includes the 30 days that we'll discuss in detail later. But I'll point out that none of the other trials included this time period in their analysis, nor did they actually do anything other than continue patients on their study drug at this time.

The second analysis in this specified hierarchy was a superiority analysis on treatment. And this, again, showed a hazard ratio of .79, hardly surprising, considering the small difference in the two patient populations.

As Dr. Mahaffey has pointed out, we fully recognize that the standard for a superiority analysis would be the full population up until the end of the study, essentially being similar to what we showed you as the next to last point in the previous slide, but this was a prespecified analysis. And we think in looking at the entire texture of the data, as shown in detail in the New England Journal of Medicine, that while patients were taking assigned study medication, there was a

benefit of rivaroxaban.

This shows a Kaplan-Meier graph for the on-treatment analysis. No surprises here. The curves diverge early and stay apart throughout the duration of the time period when people were on treatment.

This slide, then, again, talks about the ITT population with the different methods of counting people in follow-up. At the top, ITT on-treatment, we've already reviewed. The next, the ITT, the site notification, roughly equivalent to what the other trials have looked at, and then we continue to follow patients during this difficult transition that we'll review in detail in the final part of the presentation. In none of these time points did the non-inferiority margin become close to the outer confidence interval of the point estimate of treatment.

Now, what about the primary endpoint component subgroups and secondary endpoints? First of all, with regard to the primary efficacy endpoint components, as pointed out in the briefing

book, the effect was driven by the total stroke rate, which was a major component. But within that, the primary hemorrhagic stroke rate was the primary driver of the difference between the two treatments. We would also point out, however, that with regard to primary ischemic stroke, the point estimate is just to the left of 1, indicating that the benefits of warfarin were preserved in this population for ischemic stroke, in addition to the reduction in primary hemorrhagic stroke. You'll also note the robust difference in non-CNS systemic embolization, shown here, although there are a small number of events in this part of the primary component of the endpoint.

If we look at the many prespecified endpoints, the most notable finding is that there was really no significant heterogeneity when controlling for the number of comparisons that were done. To make it easier for you to follow, we put the orange dotted line in, which is the point estimate for the overall trial for the treatment effect. And you'll notice all of the confidence

intervals for all the subgroups crossed that orange line, indicating there really is no significant heterogeneity.

If we go to the next set of prespecified subgroups, the results are essentially the same. And I'll just call out one group here, because it was brought up in the FDA briefing book. The patients with a prior stroke, TIA, or non-CNS systemic embolization, there was a nominally significant .037 interaction here. But I'll point out, this is a quantitative interaction, not a qualitative interaction. The main difference here is really that, in the people without prior strokes or embolism, there is a point estimate of .59, much better than the trial as a whole. And then for people with a history of these events, a point estimate is to the left of 1. So this is quantitative, not qualitative; and we think, given the number of comparisons as well within the play of chance.

Now, what about the case secondary endpoints? And we added these in to look at

arterial thrombotic events. One would like to know about these if one is giving an anticoagulant. And the reassuring message here is that in the first secondary efficacy endpoint at the top, you can see that when we added in vascular death, it really had no substantial effect on the hazard ratio and went further in favor of rivaroxaban versus warfarin. And then when we added in non-fatal myocardial infarction, a very similar finding, numerically fewer non-fatal infarctions in the rivaroxaban group compared to warfarin.

If we look at the components again, we see there's really no heterogeneity of treatment effects. With regard to every component, numerically, rivaroxaban had fewer of these important endpoints than warfarin.

So, in summary, then, for efficacy, ROCKET AF was a double-blind study in a high-stroke-risk population. Rivaroxaban was not inferior to warfarin in all populations and observation periods, and was superior to warfarin while on treatment in all populations. There is a

substantial reduction in hemorrhagic strokes, and the results were consistent across subgroups and for secondary endpoints.

I'd now like to turn the podium over to Dr. Christopher Nessel, senior director of clinical research at Johnson & Johnson.

Sponsor Presentation - Christopher Nessel

DR. NESSEL: Thank you, Dr. Califf.

Good morning, ladies and gentlemen, Mr. Chairman, committee members, colleagues, and distinguished guests. I am Dr. Christopher Nessel from Johnson & Johnson Pharmaceutical Research and Development, and I am responsible for the safety section of the sponsor's presentation. In line with Mr. Chairman's requests, I am an employee of Johnson & Johnson.

All bleeding events in the principal safety endpoint were adjudicated by an independent committee blinded to treatment assignment, which was comprised of physicians at the Duke Clinical Research Institute. The principal safety endpoint was the composite of major and non-major clinically

relevant bleeding events.

Per the clinical study protocol, major bleeding was defined as that which was clinically overt and associated with one of the following. Non-major clinically relevant bleeding was defined as that which did not meet the criteria for major bleeding, but was clinically overt and associated with one of the following.

For the principal safety endpoint and its components, there was comparability between the treatment groups, as shown here. The hazard ratios for the composite, the major bleeding component, and the non-major, clinically relevant bleeding component, are 1.03, 1.04, and 1.04, respectively.

With respect to the major bleeding events specifically, an interesting dichotomy emerges when examining the elements comprising the endpoint. The hazard ratios for the occurrence of hemoglobin drop and transfusion are 1.22 and 1.25, respectively. Conversely, however, for the occurrence of the more severe events of bleeding into a critical organ or site and fatal bleeding,

1 the hazard ratios are 0.69 and 0.50, respectively.
2 Each of these four comparisons reached
3 nominal statistical significance. Of note, data on
4 the occurrence of fatal bleeding events presented
5 in this slide are taken directly from the
6 adjudication of the clinical events committee; that
7 is, the CEC determined that the event satisfied the
8 definition of a major bleed because it resulted in
9 a fatal outcome. While there were more hemoglobin
10 drops and transfusions in the rivaroxaban arm, the
11 most severe of these events, as measured by the
12 need for four or more units of blood, was balanced
13 between the treatment groups.
14 There were 64 subjects in each treatment
15 group who met this criterion. Over 70 percent of
16 these events were gastrointestinal hemorrhage with
17 upper GI hemorrhage contributing the majority. The
18 CEC adjudicated intracranial hemorrhages as
19 bleeding events. Importantly, however, the
20 committee also determined if the event met criteria
21 for a primary hemorrhagic stroke. Assuming the
22 remaining protocol-defined criteria were met, the

event would be included in the efficacy analyses. As such, primary hemorrhagic strokes are included in both the primary efficacy and the primary safety endpoints. The CEC also assigned all relevant intracranial hemorrhage subtypes to each event.

This is a Kaplan-Meier plot of the time from the first dose of study medication to the occurrence of the first intracranial hemorrhage. This is the safety population while on treatment. The hazard ratio is 0.67, with the p value as noted. The curves appear to separate before day 60 and remain so while patients are taking study medication.

This is the data on the occurrence of intracranial hemorrhage. More than half of the events were hemorrhagic strokes, and all elements of intracranial hemorrhage shown here were lower in the rivaroxaban group.

Here, I present data on the instance of fatal bleeding events in the safety population on treatment, by bleeding site as adjudicated by the clinical events committee. Major bleeding events

occurred more frequently in the GI and GU tract for patients randomly assigned to receive rivaroxaban. However, these events rarely resulted in death. Contrast this with the outcome associated with intracranial hemorrhage. Intracranial hemorrhage events occurred more frequently in patients receiving warfarin, and in these patients, the mortality approached 50 percent.

This is a Kaplan-Meier plot of the time from the first dose of study medication to the occurrence of a fatal bleeding event. The curves appear to separate within 30 days and remain so while patients are taking study medication for a hazard ratio of 0.61.

I presented the results of fatal bleeding defined as an event adjudicated by the CEC as a major bleed because it resulted in a fatal outcome. To gain a different perspective about these important events, we analyzed major bleeding events followed by any death in the subsequent 30 days or the broad definition, and major bleeding events followed by a CEC-adjudicated vascular death in the

subsequent 30 days or narrow definition. By any of these definitions, the hazard ratios ranged from 0.50 to 0.61 with the p value as noted.

The number of subjects who experienced adverse events, serious adverse events, and adverse events leading to permanent study drug discontinuation was essentially balanced between the treatment groups. Congruent with the data I just presented, there were fewer patients in the rivaroxaban arm who experienced an adverse event with the outcome of death.

In summary, Mr. Chairman, there was comparability between the treatment groups for the principal safety endpoint and its components. Rivaroxaban was associated with more hemoglobin drops and transfusions. Importantly, however, rivaroxaban was also associated with fewer critical organ bleeds, fewer intracranial hemorrhages, and fewer fatal bleeding events. There were similar rates of AEs, SAEs, and premature discontinuations, as noted by Dr. Califf. But rivaroxaban was associated with fewer AEs with an outcome of death.

Thank you for your kind attention, and at this point, I would like to reintroduce Dr. Robert Califf, who will conclude the sponsor's presentation.

Sponsor Presentation - Robert Califf

DR. CALIFF: Thanks, Christopher.

This may be one of the most difficult presentations I've ever had to give publicly, and that should be quite interesting. As Dr. Temple noted, I believe that if you look at the results of the ROCKET AF trial in the context in which it was designed, there would be very little to argue about.

I want to give the FDA quite a bit of credit because the issues that they have raised are exactly the same issues that the executive committee and steering committee wrestled with almost from the moment of unblinding. And we've been wrestling with these issues for more than a few months, doing a lot of work to try to understand them and to try to make sure we get this right because of the global implications of

1 decisions that will be made about this therapy for
2 a disease which is affecting a growing number of
3 people. And I would urge, again, that everyone
4 consider not just in the United States. Most of
5 these people do live in other parts of the world.
6 And it's a consideration that I believe has to be
7 part, although not the dominant consideration, in
8 these proceedings.
9 Also, we'll deviate from the usual standard
10 data presentation because the issues that have been
11 raised really require an understanding of the
12 entire picture of the treatment of atrial
13 fibrillation and the issues that are involved in
14 order to have the best discussion. So we will show
15 you a lot of data that partly represents an
16 education on my part about these really complex and
17 difficult issues.
18 If you get a chance, you can look under my
19 chair and see there's a stack of papers that
20 actually comes up to the bottom of the chair to
21 represent the amazingly diverse and interesting
22 literature on this topic.

1 So I'm going to cover four issues, but
2 first, the very first issue is just to give a
3 perspective of the executive committee about the
4 benefit-risk balance, always an important
5 consideration.
6 As Dr. Mahaffey pointed out eloquently, we
7 don't really have to belabor the issue of atrial
8 fib. It's common. It's a serious problem. The
9 negative outcomes, although not terribly frequent
10 per 100 patient-years, are devastating and critical
11 to the patient population involved.
12 Anticoagulation works. However, many
13 patients don't receive effective or optimal
14 management. And the relevant issues for patients
15 and families are, am I going to live longer, will I
16 have a better quality of life, and can I avoid
17 catastrophic and negative events? And this is what
18 we designed ROCKET AF to measure in all these
19 countries around the world.
20 Warfarin is a highly effective drug, as
21 Dr. Temple pointed out in the beginning. It has
22 proven benefits, but it also has proven risks,

1 particularly bleeding and especially intracranial
2 bleeding, which has caused a lot of people not to
3 take the drug and has scared a lot of physicians
4 from prescribing it.
5 There are other miscellaneous toxicities
6 that are important to consider, and there's quite a
7 bit of inconvenience in taking warfarin if you're
8 on the patient end of this situation. This results
9 in failure to use the treatment in situations in
10 which it's known to be effective, and this has been
11 well documented, so I won't show you primary data
12 here. When doctors are asked about this, it's
13 primarily a concern about bleeding. When patients
14 are asked, it's primarily a concern about the
15 complexity of the regimen and the bleeding that
16 might occur.
17 So ROCKET AF was a double-blind, global
18 trial in a high-risk population, primarily intended
19 to find an alternative for warfarin, with an intent
20 to move onto superiority testing if non-inferiority
21 was demonstrated. But make no mistake about it, we
22 were focused on non-inferiority to find an

alternative to warfarin.

We understood from the beginning -- and this is another critical point -- that the high-risk population would lead to significant discontinuation with its attendant analytical issues. Just to give you a metric of the complexity of the population, our average ROCKET AF patient was on nine concomitant medications, with an upper quartile well into the double digits. This indicates that there were many difficult issues in managing not only warfarin but the general patient care of this population. And I would hate it if the message from these proceedings were don't study high-risk populations; only stick with the easier-to-study low-risk populations.

Then careful measurement was made of a broad range of adverse events because in benefit-risk, one wants to consider the whole story that patients and doctors may be concerned about. We really found nothing about rivaroxaban -- and the FDA I think is in substantial agreement -- other than this tradeoff of prevention of ischemic events and

1 bleeding.
2 So as we look at this, this is an array of
3 all the key endpoints that were relevant to the
4 benefit-risk tradeoff that we came up with, and
5 I'll call your attention to several key points.
6 First, the point estimate -- and this is expressed
7 in risk difference per 10,000 patient-years, so
8 these are absolute numbers of events. The point
9 estimate is reflected by the diamond. The
10 confidence intervals for the efficacy endpoints are
11 reflected by the orange bars. The confidence
12 intervals for the safety endpoints are projected in
13 the green bars. Anything to the left of zero
14 indicates fewer events of rivaroxaban. Anything to
15 the right indicates fewer events with warfarin.
16 Since these results are for the trial as a
17 whole and there was no significant heterogeneity,
18 we believe that these are the most important
19 results for doctors and patients to consider.
20 When one looks, then, at the efficacy
21 results, you'll notice they are all on the side of
22 rivaroxaban, although very often, the confidence

1 intervals cross 0. I'll also point out here, for
2 reference, that we are showing the data in terms of
3 the safety on-treatment population. Your briefing
4 book has all the data by intention to treat, up to
5 the end of study also, and it's available.
6 So the principal safety endpoint goes to the
7 side of warfarin, but you'll notice a pattern here
8 that Dr. Nessel reviewed. Fatal and critical organ
9 bleeding, intracranial hemorrhage, favoring
10 rivaroxaban; transfusion and hemoglobin decreased,
11 favoring warfarin.
12 So what do doctors and patients actually
13 want? And we're benefitted here by a lot of
14 previous studies that have been done in the context
15 of warfarin versus no warfarin. And although the
16 details come out slightly different in different
17 studies, in general, across multiple studies,
18 patients and doctors agree that the most important
19 endpoint actually is death, not terribly
20 surprising. But it's also notable that when
21 patients are asked, particularly those who have had
22 a stroke are asked, over 30 percent will rate a

disabling stroke as worse than death. And so these two endpoints do cluster at the top end of the scale of what people would like to avoid. Bleeding is rated at the lower end of this continuum, with non-disabling stroke and myocardial infarction in the middle.

So if we look at these absolute numbers in terms of the most severe at the top and the least severe at the bottom, now couched in terms that patients and doctors would think about, rather than MI stroke, talking more about death, disabling stroke, and embolism to the other part of your body besides the brain, arrayed in this order, you can see the most important endpoints tend to favor rivaroxaban. The least important endpoints tend to favor warfarin.

We're not discounting the importance of bleeding. It is an important endpoint. We're just saying that as the executive committee looks at these data, it's important for us to consider what is most valued by those who will be prescribing and much more important to consider what will be

1 considered by those who are contemplating taking
2 the treatment.
3 So the benefit-risk summary, from the point
4 of view of the executive committee, is that
5 rivaroxaban has a favorable benefit-risk balance.
6 This is not because of declaring superiority by
7 classical intention to treat. This is because we
8 do believe that it's clear that rivaroxaban is non-
9 inferior for the primary endpoint, but has a lower
10 rate of intracranial and fatal bleeding. And that
11 is a very important endpoint for the decision to
12 treat and to end up with better outcomes over all.
13 There are also other benefits in terms of
14 the once-a-day dose and the monitoring and the
15 limited potential for drug and food interactions
16 that are important in daily practice. There is an
17 increased risk of blood transfusion and greater
18 than 2-gram per-deciliter falls in hemoglobin
19 concentration, but no significant difference in
20 bleeding overall.
21 Now, we get to the really interesting and I
22 think difficult issues that I know the discussion

will focus on today, and we look forward to the panel's insight into these issues. But we do have a point of view that's been molded by months of review from a global executive committee that includes a number of experts in cardiology and in coagulation, and specifically on the management of warfarin.

These four issues that were raised by the FDA are dose and regimen selection, events after study drug discontinuation, what about the time in therapeutic range, and this approval standard in 2011, a new issue that was not at all contemplated at the time that the study was designed and conducted.

So first of all, what about the dose selection? Now, I would point out here that dose selection was made in the context of many studies in venous thrombosis that looked at dosage ranging from 20 to 60 milligrams, and looked at once- or twice-daily dosing. And across the range of these studies, when looked at critically by the sponsor, one could pick one study here or one study there

that might imply that once-a-day or twice-a-day dosing is better. But the fact is, across the range of studies, what was seen was a flat dose response for efficacy and a shallow dose response for bleeding. And similar results were once-a-day or twice-daily dosing.

So having to make a decision, it was felt that the convenience factor and the likely improvement in adherence should dominate, and once-daily dosing was picked. In addition, there was modeling of the coagulation system, based on what was thought to be the type of consistent anticoagulation needed to get the benefits from warfarin, and pharmacodynamic studies were done in normal populations for population PK/PD modeling.

The next slide shows what the population modeling showed. And it's a complicated slide. I'll just point out that what it shows is a steady state, whether you measure a dynamic measure like prothrombin time ratio or prothrombin-induced clotting time ratio at the top, or you measure the concentration of the drug.

If you look at the right-hand side of the slide at 24 hours, we're in a range where there still is anticoagulation present. Now, anyone who says they know exactly what the right parameters for these measurements should be, to be sure that you have the best anticoagulation I think would have insight that only some of our politicians today actually seem to have, not that we would claim to have the scientists. But the best, as we could judge, by looking at this, we were reassured as an executive committee that once-a-day dosing was rational. We are not saying that it would be irrational to pick twice-a-day dosing. That's also a viable alternative.

This is just a look at one slice of the data. These are the venous thrombosis studies that looked at the total dose chosen for this trial, 20 milligrams a day, comparing 20 milligrams once a day to 10 milligrams twice a day. This is a cross-study comparison. It's not the ideal way to look at data or to have data, but as we looked at it, there was really no consistent pattern, either with

regard to bleeding or with regard to efficacy.

So we believe the once-a-day choice was rational. It's not the only choice that could have been made. Other choices could have been made and rationally defended. And so, in summary, the modeling was consistent with good choice of a dose. The phase 2 data support it, either once- or twice-daily dosing. We had reasons to pick the once-a-day dose and to put it to the test. And, after all, we're here today, presenting results of a trial with the dose that was chosen, and we think that should be the determinant of a decision about whether the patient should have access to this treatment.

Now, what about thrombotic events after discontinuation of the study drug? In here, out of necessity, I'm going to have to do something I'm a bit uncomfortable with, and I want to lay out the general precept that if you catch me going beyond this precept, please regard it as a mistake.

The FDA, in its briefing book, and as Dr. Temple pointed out, brought up the issue of

1 cross-study comparisons. We don't typically do
2 that in this kind of a setting, but because it has
3 been raised and it is a key issue, I will have to
4 point out differences in studies. But what I will
5 avoid, unless I slip up and make a mistake, is a
6 cross-study indirect comparison comparing the
7 treatment effects of the different drugs. As we'll
8 say at the end, we think that is not only invalid.
9 We think it is treacherous and dangerous to the
10 public, to doing those kinds of comparisons.
11 In discussions with the FDA, they reasserted
12 that the purpose of the briefing book was not to
13 engender illegitimate cross-study treatment
14 comparisons, but to talk about the qualities of
15 anticoagulation, the designs of the trials, and the
16 other issues in the trials, that you can get a
17 rough estimate on by looking across the studies.
18 So I want to be clear about what I'm trying to do
19 when I bring up other trials.
20 So we had a prespecified method of
21 transition at the end of the trial. This was based
22 on the fact that we had very little, other than

phase 1 data, data about giving both rivaroxaban and warfarin together. It was felt it would be better to err on the side of not causing bleeding by over-anticoagulating. And so the prespecified protocols start VKA at the expected maintenance dose. Remember that globally, warfarin is not the only vitamin K antagonist -- there are many other forms of vitamin K antagonists that are used -- no overlap with the blinded study drug; no INRs for three days; heparin bridging therapy to be allowed but infrequently used. And the impact, as we now know, is an imbalance in anticoagulation between the two treatment groups, leading to the results that have been reviewed in the briefing book.

Now, you might ask, how could you possibly do that? Shouldn't you have known better? And in fact, as the FDA pointed out and we reflected in the briefing book that was sent from Johnson & Johnson, we did have access just at the end of the trial to the J ROCKET data, which seemed to show this pattern of excess events.

As an executive committee, we needed to stay

1 blinded. We had a protocol. We asked the data
2 monitoring committee, do you see a signal, looking
3 back at the ROCKET data, that we should change our
4 protocol to make it safer for the patients? This
5 was delicate because we could not become unblinded.
6 The data monitoring committee could not give us
7 evidence that would show that would lead us to
8 become unblinded. They sent us a note, which is in
9 the record, that said we don't see anything that
10 should cause you to have to change your protocol,
11 but of course, the decision is yours, not ours.
12 That was a good thing for them to say and perfectly
13 correct.
14 Now, as we look back at those data, just to
15 report it to you verbally, what they saw from
16 discontinuations in the midst of the trial was
17 that, actually, there were more events in the
18 warfarin group than the rivaroxaban group at the
19 time they looked, right near the end of the study.
20 So they were actually giving us an evidence-based
21 recommendation that turned out to be wrong for the
22 context in which it was used.

The result is shown here. And what I think is an elegant but simple slide, remember, the context of this slide is, we're at the end of the study. We told all the sites the study's over. We want you to make a transition to whatever the vitamin K antagonists that you use in your country or region.

So what they did in the warfarin group is they were transitioning from blinded warfarin to a vitamin K antagonist, and this is a transition from, essentially, one drug to the same drug. And there's a very high rate of INRs greater than 2 in the population right from the start.

But in rivaroxaban, we were making a transition with a three-day hiatus to a vitamin K antagonist, and you can see that there were very few INRs that were in the therapeutic range of greater than 2, all the way out to a month, where we only reached 52 percent out of those who reported an INR.

So the executive committee believes that the excess events at the end of the trial are a result

of the transition strategy that we take accountability for, but we believe it was done in good faith and with a lot of discussion and thought.

But what about the data overall? And I think it's important, as one looks at this, to keep in mind several things about the slide I'm showing you now. Obviously, during the trial, many people discontinued study drug either temporarily or permanently. And if we look at the top row of this slide, we can see that, indeed, there was an imbalance in events overall between rivaroxaban and warfarin after discontinuation with a hazard ratio of 1.5. However, the pattern is quite clear. If one looks in the second row at those who completed study medication, you can see an event rate of 6.4 per 100 patient-years on rivaroxaban, 1.7 in warfarin, clearly a problem, with a hazard ratio of 3.7. But I'd point out a couple of things here. The 6.4 rate, as I'll show you, is exactly what you would expect in this population if they were not anticoagulated. It's not higher than you would

expect. The 1.73 is what you would expect would go to warfarin management, which is what we had at the end of the trial.

Now, look at rows 2 and 3. And here, there's a little trick that I need to tell you, that could be misleading if you didn't realize it. We're obviously concerned about people that need to stop anticoagulation for a procedure or because something has happened in their lives. Will there be a recurrence of events? And that would be called, at the bottom, a temporary interruption of treatment. And you have those who just simply say, I can't take this anymore, I'm going to stop the study drug permanently. That would be column 3.

But the trick is that if you intended to have a temporary interruption and then a patient had an event, they would stop it permanently, they end up in column 3 instead of column 4 when the intent was a temporary interruption.

So these numbers can be a little bit misleading, but when you look at the two groups together, I think what you see is that there's

really no significant trend. There is a slight excess, numerically. You might argue we don't have the ideal power, but there are a lot of people discontinuing in this group, and a fair number of events.

We then looked at the expanded endpoint because you would think if there is a problem with rebound as an issue, that you would see it on their arterial side as well as the venous side. So when we add in vascular death and non-fatal myocardial infarction, you'll see now we have for all participants who discontinued treatment, a hazard ratio of 1.00. It's harder to get closer to a quality than 1.00 in a clinical trial.

Now, there is a pattern here, with a 2.24 hazard ratio at the end of the study. So we do think there was a problem with transition at the end of the study. But let me point out, again, this is not a lot of events. It's a small event rate, overall, in the population, and consistent with what you would expect. If we look, though, at early study medication discontinuation and

temporary interruptions, we see really not a scintilla of a trend towards a negative outcome for rivaroxaban-treated patients, compared with warfarin.

Then the next piece of evidence is what would you have expected? And as I've already told you, if you do a literature review in a population with this degree of risk, if you had an adequate anticoagulation, you would expect an event rate on the order of 6 or so, which is exactly what we observed in the rivaroxaban patients who discontinued treatment and where the transition was not ideal.

In the ischemic stroke after transition to VKA, in that population, when we looked at the INRs in the population that had an ischemic stroke, 73 percent had a last INR that was less than 2. Now, these are kind of messy data, as all of you know that have dealt with clinical trials in this kind of situation, but we've looked at this very carefully and this is what we've found.

So the recommendation is to maintain

1 continuous anticoagulation, applying the same
2 principles as are used for low molecular weight
3 heparin around the world and in the clinical
4 practice guidelines when one is interrupting
5 vitamin K antagonists and to overlap the two drugs
6 until the INR is greater than 2, measured at the
7 trough of rivaroxaban. And the reason for this is
8 there is an interaction between rivaroxaban and
9 warfarin, so you want to measure the INRs at the
10 trough.
11 This is only supported by clinical
12 pharmacology data -- we want to be clear about
13 that -- not by a clinical trial in this population.
14 We do have experience now, considerable experience,
15 with the overlap of the two drugs from the
16 beginning of this study. But we also want to be
17 careful to not say the two situations are
18 equivalent. Starting rivaroxaban from a patient on
19 warfarin is not the same as stopping rivaroxaban
20 and starting warfarin. But, nevertheless, it does
21 give us data about the two drugs given together.
22 Now here I have to pause for a second and

remind the panel and everyone in the audience that we actually don't know what to do for either warfarin or for dabigatran when it's discontinued at this point. As you well know, most of you on the panel, there are currently two NIH trials looking at bridging with warfarin -- and we're the coordinating center in one of those trials -- because it's unclear whether bridging is an effective strategy in modern times.

For dabigatran, patients were followed up to the last study visit, and then they were transitioned from blinded drug to open label drug, exactly the same treatment that they were on throughout the trial. So there was no strategy for transitioning in the RE-LY trial. I'll also point out, just as a way of anecdote, when I was a resident, I used to round with an intern. When I asked about the lab values, they kept saying WNL, and I thought that meant within normal limits, but I found out at the end of the month, he was saying, we never looked.

[Laughter.]

DR. CALIFF: So we're not unique in this regard of basing our transition strategy on pharmacodynamics. As clinical investigator, we obviously think there needs to be more data, but we don't necessarily think that this drug should be held to a different standard than other treatments.

So, in summary, in events after discontinuation, there are no excess events in the midst of the trial, either with early permanent discontinuation or temporary treatment interruptions. And the excess events at the end of the trial we think were due to an imbalance in anticoagulation due to the protocol that we asked the investigators to adhere to. And we have a transition plan that we think will take care of this, which is at least as good as the transition plan basis for other drugs that are currently on the market.

Now, what about time in therapeutic range? And this, I've got to say, is a pretty amazing topic. I commend the FDA for bringing it up and for raising it in quite a provocative way in the

briefing book. I think most of you know I love the FDA, so please don't interpret anything I say as bashing the FDA. I think that the questions were raised in a very provocative way, and this compels us to stand up for our investigators and tell you that we gave warfarin not only in an acceptable way, we gave it in a commendable way during this trial, considering the complexity of the trial design that we had. And we think it's important to study these kinds of patients around the world.

So the key points I want to make is that TTR is a useful biomarker. And it is a very useful quality measure. And once you've made the decision that you're going to anticoagulate a patient or a group of patients at your center, given their characteristics, it is a very useful quality measure, but it is not a surrogate for anticoagulant benefit-risk balance.

The ROCKET AF TTR we're going to show you is consistent with standards for use of warfarin, considering the more complex patients and the regional variations that are well documented and

reproducible in other studies. And most importantly I think for this discussion, there's increasing evidence that TTR has no effect on the benefit for novel anticoagulants versus warfarin.

Now, first of all, I want to reiterate again, we are not here to bash TTR as a quality measure. If you decide that you're going to start a patient on warfarin, achieving 2 to 3 INR is a critical thing to do, and you should make every effort to achieve it within the balance of culture, reimbursement policies, and other issues that you have to deal with in medical practice. And when you fall below 2, you're going to have more ischemic strokes and less bleeding. When you go above 3, you're going to have more intracranial bleeding and maybe, perhaps, even a slightly better effect on ischemic strokes.

So I'm going to hit you with the key points first. I'm going to then go through some detail, and then I'm going to come back to the final points. So what about quartile's center of INR? And we'll go through this in more detail later.

1 But as you know, in our prespecified method,
2 without knowing how it was going to come out, we
3 simply found no relationship between center TTR and
4 the treatment effect of rivaroxaban versus
5 warfarin.
6 We're now bolstered in our thinking by the
7 actual results of the RE-LY trial, which was
8 referred to in the briefing book. And here, the
9 FDA used a sleight of hand that we don't typically
10 recommend for our fellows. What was done in your
11 briefing book was to compare quartile 1 with
12 quartile 4, without showing the data that was in
13 the middle, our mentioning that the test for
14 interaction had a p value that was not significant.
15 The interpretation of the RE-LY
16 investigators themselves is that there is no
17 interaction between center TTR and the impact of
18 rivaroxaban on stroke, compared with warfarin. And
19 then finally, in data available on a public website
20 at the European Society of Cardiology, our
21 colleagues at the DCRI, together with their
22 sponsors, completed trial of another novel

anticoagulant, with even more patients than we had, that showed absolutely no relationship between center TTR and the estimate of the treatment effect.

So the three most contemporary trials of novel anticoagulant simply don't support the concept that TTR should be used as an intermediate variable to somehow decide whether or not you should believe the results of the trial.

Now, the second key point I want to raise is that, as was implied in the briefing book, we chose unskillful investigators in order to make the results come out better than they should have been. You would have thought that the results in North America, where the TTR was 64 percent, would have looked worse for rivaroxaban than the results in eastern Europe, where the TTR was 50 percent.

Well, the data just don't support that conclusion, and it was certainly never anything that we considered as we designed this global trial, remembering that most people with atrial fib don't live in the United States. In fact, the

estimate of the treatment effect is more positive for rivaroxaban in North America, and the least positive two regions were Latin America and eastern Europe.

The third point, despite the fact that the FDA didn't seem to think this was important, as we talked to multiple experts in anticoagulation therapy, they've all told us that the event rates that you get is the outcome of interest. TTR is what we typically call a process measure or a biomarker. What we're interested in is preventing strokes. And if you look at ROCKET AF, where the TTR was lower but the CHADS score was significantly higher, as you can see in the second column, and then look for people with a CHADS score of 2, or a CHADS score of 3, or greater, or you look at those with a prior stroke, you'll see the event rates in ROCKET were absolutely comparable to the event rates that were achieved in the two other most recent trials of novel anticoagulants versus warfarin, RE-LY and ARISTOTLE.

Now, I just want to point out that we did

measure TTR using conventional methods. And this has been reviewed and agreed between the FDA and the sponsor. This Rosendaal method is used, and as a standard for the field, it's important for everyone to recognize, this involves an interpolation. The more often you measure it, if you're in range, the better your TTR will look. It's a linear interpolation.

We used a relatively conservative approach, we believe, for the individual patient TTR, the average for the trial as a whole, because we included the start-up phase of the trial and only excluded interruptions that were greater than seven days. Many other studies have excluded the start-up of the trial, where you would expect the TTR to be low. I have to say, though, no matter how we change the definition within definitions that others have used, it had very little impact on the result that we got.

So by our prespecified method, we had a mean TTR of 55 and a median TTR of 58. I would be lying to you if I didn't tell you that when we unblinded

the trial and looked at this as an executive committee, we went into high gear to try to understand it because we wanted to be sure that we had administered warfarin in an excellent manner.

Now, there's another measure which is different, called the center TTR, which is the subject of discussion, and I know will be reviewed in detail by the FDA. This can be done in several ways. We thought we had introduced, in advance in the field in our prespecified method, because our belief that when one looks at the amount of time in therapeutic range, it would be better to count the total time that one is on warfarin, so that one has a total denominator and the total numerator, rather than the way it's been done traditionally, in the method of Connolly, where a patient on warfarin for one week gets equal credit to a patient who's on warfarin for five years.

We recognize the counter argument that, since people who stay on warfarin tend to be those who do better, that this could be regarded as maybe not the best way. We'd love to have a debate about

the best way, but it's not the main point today. We just want to recognize that there were differences. We don't think they made a major difference in terms of the interpretation of the trial, and we'll show you why.

A slight caveat here; centers that have no events and centers that have very few participants, since you're averaging for patients in a center, it's a little hard to come up with exactly how to measure that. It's just a problem with the method.

Finally, just to point out, when we look at the quartile analysis across multiple ways of measurement, with our way at the top, the FDA way second from the top, you can see it really did not have a significant impact on the multiple different ways the quartiles end up being grouped. So you can imagine, if you change the definition slightly, the centers end up in slightly different rank order.

Now, I promise you, I did not come up with this slide after we learned that Dr. Fleming was going to be on the panel. This only happened the

day before yesterday. We got the roster. But I think it's not the most important argument, but I'm worried that some people who have not worked in this field may be misconstruing TTR for a surrogate endpoint. It simply does not make the cut as a surrogate, and I refer you to the classic paper of Fleming and Demets of 1997.

Now, you might believe, based on some of the literature, that if you have atrial fib with a potential for thrombus there, and you give warfarin, and you measure time in therapeutic range, you can explain all of the effect of warfarin on the prevention of stroke and embolism. But as Fleming and Demets showed, drugs don't work that way. Drugs have protean effects, and the outcomes are affected by protean impacts from things that are true to the drug and not true to the drug. Just as an example, atherosclerosis itself, independent of thrombosis, can cause a stroke. That's not going to be impacted by the TTR that you get. Warfarin has effects on multiple proteins that are not well quantified. I can go on

and on about this, but I won't in the interest of time.

The second key point to make in this section is that, when one looks at the reference trials for constancy, one can see that there is a broad range of estimated time in therapeutic range. Now, it's critical to point out here, that these are old trials. Most of them didn't even have INR as a measurement. They had prothrombin ratio, the old measurement that us old folks are used to.

But we did this in a think tank in 2008 and published it. It was our best estimate at the time, not knowing how our data would come out. And you can see, there's really no very direct correlation between the risk reduction ratio compared to no warfarin and the TTR that was observed in these trials. And while it's imprecise, if you thought the TTR was a great predictor of what was going to happen with prevention of stroke, you would expect to see the studies line up in terms of risk reduction ratio as a function of TTR.

1 So TTR just doesn't accurately predict risk-
2 benefit balance. It is a predictor, and it is an
3 important quality measure in an individual site.
4 And the reasons are shown here. I won't go into
5 them in detail. I think you can all accept that
6 this is the case, but I'll show you more detail
7 from studies themselves that have recently been
8 done. So we believe that TTR is a useful measure
9 for quality improvement, but it is not a valid
10 surrogate and we're happy to entertain any
11 discussion about that point.
12 So what about TTR in the context of this
13 study? So first, let's think about what determines
14 TTR. Here, I'm not giving references because there
15 are over 50, but there is really a consistent
16 result here when studies have looked at the
17 determinants of TTR, and the skill of the physician
18 is only one. And we would argue that these
19 components outweigh the skill of the physician when
20 you look at the weights from the studies that have
21 been done.
22 Prior VKA use is a big factor; the

characteristics of the patient or the patients that you're seeing in your clinic. Comorbidities are especially important. The frequency of INR testing itself is a measure for reasons I'll show you in a minute. Patient adherence is critical, and the idea that a doctor can control that I think is a little far-flung. A doctor can influence it, but not control it.

Structural factors in practice, in region, and country are dominant factors. There are social mores by which people give warfarin for reasons that we think we understand much better now. But even within the United States, a critical factor is how much poverty there is in close proximity to your clinic and how far away from the clinic your people live and whether there is public transportation available.

This is one of my favorite plots, but it's a little hard to absorb. Those of you who know Frank Carroll recognize that we brought Frank in to analyze this data from stem to stern, using modeling techniques. What this shows is if you go

1 out to predict TTR in the ROCKET population, what
2 predicted it? If you start at the left-hand bottom
3 sign, you'll see that, starting with atrial
4 fibrillation type, there are a host of biological
5 clinical factors that have an impact on the TTR
6 that's achieved within the ROCKET trial when we
7 look at each of the sites and each of the
8 individual patients. But the sum of all these is
9 relatively modest compared to whether or not the
10 patient was on a vitamin K antagonist and is also
11 modest, very modest, compared to the country that
12 the warfarin is being administered in. In fact,
13 the country is three to four times more important
14 than all the other factors in aggregate.
15 Now, what about the trials themselves, not
16 only ROCKET AF, but the other trials, in terms of
17 the severity of illness of the population? For
18 display, we picked out three of the baseline
19 characteristics that we thought were important.
20 And you'll notice that our centers with the worst
21 TTR also enrolled the highest-risk patients,
22 89 percent which adds 3 or greater, compared to

82 percent in those in the best quartile. The same holds for prior stroke and for heart failure.

If we then look at the ARISTOTLE trial, what you see is the very same pattern, but a much lower rate of these high-risk characteristics, a much lower rate of these high-risk characteristics. But in ARISTOTLE, 32 percent of the patients in the worst quartile were CHADS2, 3 or greater; 27 percent in the best quartile were CHADS2 or greater.

Now, the best study to look in detail at this and to quantify the impact of patient factors on TTR is of area study (ph). And this actually shows the direct impact on the numerical percent TTR for a population. And you can see that being a woman compared to a man, after adjusting for other factors, has a 3 percent impact on TTR. And we had a very high number of women in our trial. The number of medications, an important factor, hospitalization, socioeconomic status, all of these are factors that were more heavily represented in our trial than the others, and certainly had an

1 impact on our TTR.

2 Now, how does this stack up for U.S.

3 anticoagulation? Dr. Mahaffey has already shown

4 you this slide. The mean for the U.S. is .55,

5 exactly what we got for the trial as a whole. But

6 some very recent data just shown last December from

7 the Quest Diagnostics, in practice in the United

8 States, when people check an INR, how does it look?

9 We have the modest sample size of 3,493,000 samples

10 in this population-based study. And what you'll

11 see is that only 50.6 percent of them were in the

12 therapeutic range.

13 You'll also notice a pattern here that we

14 saw in our trial and every other study. For those

15 who have not looked at these kinds of data, it's

16 important to recognize that, by and large, the

17 problem we see is that under-anticoagulation, not

18 over-anticoagulation, in the populations. But when

19 we look at our trial, we find not only were we

20 better than U.S. practice in North America, we were

21 also comparable to the other studies in North

22 America. Here, you see RE-LY, ACTIVE W, and

SPORTIF. This is hardly a result that should be attributed to unskillful clinicians. And I have to say, on behalf of the executive committee and the over 1,000 investigators, we don't think we're unskillful. We think we're pretty good.

Now, if we then go and say, what about the other regions, you'll find, when we match up with RE-LY, the reference trial referred to by the FDA, our results, taking into account the severity of illness factors that I've just shown you, we think, are directly comparable to the results that were seen in RE-LY. The sum is because we had more patients that were enrolled in these other parts of the world. And, in fact, specifically, we had more patients enrolled in eastern Europe, a point that the FDA makes. It's true. People in eastern Europe have atrial fibrillation also.

So when we take these regional differences and the fact that we were the same as the other studies, given the regions that were included in our trial, and adjust for that, we find that we're pretty close to RE-LY, and that if you take into

account the severity of illness, we believe we did as well, or perhaps even better, than the other trials in terms of management of anticoagulation.

So our conclusion in this part is that TTR was similar to recent trials. It was similar to other trials in North America, despite the higher risk. It was better than global practice. It was similar to the standard U.S. practice with the trial overall, and specifically in North America, it was better than U.S. practice.

So do we have direct evidence of a cultural effect? And this is a fascinating sub-analysis, and maybe even discussion we can go into more, but in the interests of time, I'll do it quickly.

If one asked the question -- in our study, if someone came in with a low INR -- and remember the circumstance. A patient comes in, once a month. There's a blinded INR that's done. The physician gets that result back immediately and is instructed to make a dose adjustment either for the blinded rivaroxaban dummy or for the real warfarin. And you ask the question, when was the patient

1 brought back for the next measurement? As you
2 would expect, in North America, if the INR was
3 really low, less than 1.5, within eight days on
4 average, the patient was brought back, within 14
5 days for a slightly less low INR; western Europe,
6 similar. But as we get to Latin America, eastern
7 Europe, and Asia Pacific, you can see that what
8 they did was to bring the patient back at 30 days.
9 But remember, they adjusted the dose based on the
10 blinded INR they had at that time.
11 It's important to point out here, we
12 excluded the first three months because, during
13 that time, we had specified protocol study visits
14 at one week, two weeks, and four weeks. And that
15 would make it look like people were coming back for
16 clinic, based on the clinician's recommendation,
17 when that was really what the protocol was asking
18 them to do.
19 We do have evidence, then, and you say, did
20 they make the right adjustment? And here, we find
21 there's no material difference, according to
22 region, other than a slight underdosing in eastern

1 Europe and Asia Pacific. And, frankly, what we
2 found in discussing this with colleagues all around
3 the world is that, in general, if a patient has
4 trouble getting back to clinic, you will tend to
5 underdose the anticoagulation, under the precept
6 that we're all taught to do no harm. An
7 intracranial hemorrhage to a clinician, as you
8 know, is a devastating outcome that may cause them
9 not to prescribe anticoagulation in the next batch
10 of patients.
11 So, in conclusion, there is a variation by
12 region. This induces a somewhat artificial
13 lowering of TTR, because with a linear
14 extrapolation, if you adjust the dose but you wait
15 30 days to bring the patient in, you get no credit
16 for that time, whereas if you bring them in
17 quickly, you get credit for that time in the linear
18 extrapolation.
19 For those of you who haven't caught onto
20 this yet, those of you who are looking at practice
21 measures in the U.S., this is a trick that all
22 American hospitals know now. You can look really

good in your TTR to report for public purposes if you bring people back alive and measure them when they're within the therapeutic range already.

So now, finally, the really critical point I think to this discussion, we just don't see evidence that center TTR had an impact on the estimate of the treatment effect. Within our own trial, regardless of the quartile that we looked at, you can see that the confidence intervals, even within each quartile, don't come close to the prespecified non-inferiority margin. I've already shown you that the other two contemporary trials show the same thing.

Now, this is the last part, and bear with me. This is the most complicated part, but I think it's really important. Those of you who are accustomed to looking at data know that when one changes definitions and moves people from one quartile to another, you can hide some really interesting and important data.

So together with Frank Carroll and the statisticians at DCRI and J&J, we worked on some

modeling to look at every single center. We'd gotten used to using the term open kimono because people understand exactly what that means. This is not hiding anything within any arbitrary quartile. This is showing you all the data.

So this is using our prespecified definition, and bear with me here. In the spirit of a global trial, I want you to start at the right-hand side of the slide, not the left-hand. So start on the right. The first dot represents centers with excellent anticoagulation if you believe that TTR is all there is to it. And you'll see the estimate of the treatment effect is well below the line of 1, indicating favoring rivaroxaban in that group of sites. You'll also notice the confidence intervals go to the sky and the ceiling, because this is the first small batch of sites.

As we go from right to left, we enter one site at a time or one group of sites at the time, according to rank order, center, time in therapeutic range. And as you'll notice, as we get

to 65 and 60, the line begins to stabilize. The confidence intervals come together.

As the FDA rightly points out, we do not have voluminous data in the greater-than-75 category, but when we look at this pattern, what we see is really something that shows pretty much constancy across the range and some random variation, because you wouldn't expect the estimate to go up at 70 and then come back down at greater than 80 if this were a true trend.

This is looking at the primary efficacy endpoint. And this is looking at the FDA definition for a greater-than threshold of safety on-treatment population. And you can see that, with the FDA's definition, which defined the outcome somewhat differently, and resulted in grouping the centers somewhat differently, the line looks modestly different. But we would argue that the pattern really looks the same. There's up and down action in the line, but nothing that I would regard as something that's a true effect.

I think, very importantly, if we take this

really critical endpoint of intracranial bleeding, one might think that if you have the most skillful centers, that they would have a very low rate of intracranial bleeding with warfarin. And this is where rivaroxaban would really not look better for this endpoint. But, if anything, rivaroxaban looks even better than warfarin, the better the skill of the center, up in that upper range. And so there's no evidence here that the center TTR should really be impacting our overall thinking about either the quality of the trial or the result that was observed.

Now, we did some fancy modeling because, since we did not see heterogeneity, we did not see breakpoints, and swine plots, and other ways of looking at it, that the best estimate of the event rate on warfarin would be as a function of center TTR is a linear relationship. But look. When we add in rivaroxaban the same way, there is a much lower slope, much smaller slope, but the very same relationship. And you'll notice that the lines do not cross. For completeness, we add in the

95 percent confidence intervals here.

Now, if we take the ratio of these two plots, approximately the ratio of these two plots, and look at the estimated treatment effect as a function of center TTR, you'll notice that the line never crosses 1 for the primary analysis. And, again, we're only claiming non-inferiority in this analysis. There is a slope to the line. It's not statistically significant.

Now, if we add in the quartiles that the FDA uses, you'll see the capriciousness where the third quartile drops below the line, and the fourth quartile goes above the line. And the straight line would bring them into something that I think would be a more reasonable estimate of what you're actually seeing in the data.

Now, obviously, we get into opinion here, and we have to fall back on the other recently completed trials for validation of what we believe to be the case with these data. So we conclude that when all the results are displayed, open kimono, you can see everything that we have, there

is no evidence that center TTR is a significant predictor of the estimate of the comparison of rivaroxaban or warfarin.

Then finally, I think it's important to bring up North America because it has been the topic of much discussion in these committee meetings recently, some of these involving our coordinating center with fairly well-known clinical trials. And recent trials have highlighted these differences. It is important to show that the trial was well done in North America and that there's no reason to think there should be heterogeneity in interpretation.

I've already shown you this slide. I'll move through it quickly, but just to point out, North America had the highest time in therapeutic range, comparable to other trials with a North American result. And a point estimate for the rivaroxaban/warfarin comparison is most favorable in this region, with the best time in therapeutic range. So we see no reason to exclude North America from access to this treatment.

The very last point, event rates as a measure of quality of anticoagulation, I only have one slide on this. I've already shown you this slide. Contrary to the FDA point of view, we spent a lot of time trying to estimate what the event rates would be. There was numerous prior data. We pretty much hit it on the head. And our event rates for the same populations in other contemporary trials show that our doctors were just as good at using warfarin. They may have used slightly different methods because they're from different countries, and we did not artificially tell them to use the U.S. method, but they got results that were just as good.

I also think it's important to realize that these event rates are really representing a very excellent level of anticoagulation in general. And you'll notice that, at the end of the trial where we had under-anticoagulation with either treatment, the event rates went back up to the basal rate that you would expect. So we think this is evidence of excellent anticoagulation.

1 So, in summary, then, about TTR, there's no
2 relationship between treatment effect and center
3 TTR in our trial or the other two contemporary
4 trials that have recently been reported. The best
5 TTR observed was in North America, with the most
6 favorable estimate of the effect, and the risk-
7 adjusted warfarin event rates are comparable to
8 other contemporary trials, indicating well-managed
9 warfarin.
10 Now, the final three slides, the approval
11 standard in 2011. This is a tough one I think.
12 Representing the sponsor, it's not appropriate for
13 me to get into some of the things I would say in a
14 more academic environment about this. But we think
15 that we designed a trial in good faith. I intended
16 to find a substitute for warfarin. We think, in
17 the spirit of non-inferiority trials, we preserved
18 the benefit of warfarin, and that there are
19 numerous other advantages that we have elucidated
20 that are important to patients and doctors and that
21 are critical to anticoagulation around the world.
22 We believe that indirect comparisons for treatment

effects are not reliable.

So what about these cross-study comparisons that may be alluded to in a discussion? We think comparing studies to understand how the studies were done, as the FDA has emphasized, are the right thing to do. But if we fall prey to a situation where indirect comparisons across clinical trials are used to rank drugs in terms of their safety and effectiveness, we will be doing a disservice to the public health.

Specifically, with regard to ROCKET and RE-LY, we'll point out that nowhere in the FDA briefing book does it point out that RE-LY was not a blinded study. This is a major factor in the way clinical trials are done. And if the point of the briefing book is that we should do not-blinded studies and hold them in the same regard as complex, blinded studies and then to do away with investigator bias, I think we're taking a step back in terms of the history of clinical trials, and I would hate to see that happen.

Let me be clear here. I am not bashing

1 RE-LY. It was a great trial. I'm on record of
2 having said the ideal design would be two trials,
3 one blinded, one unblinded. But at today's cost
4 and the complexity of doing these trials, it's not
5 necessarily the case that you can always do that in
6 one study program. These different designs,
7 different drugs, different patient populations, and
8 different regions for a global problem simply do
9 not support using indirect comparisons to draw
10 conclusions about ranking one drug compared to
11 another on behalf of the public health.
12 Final slide. I know you're worn out; I am,
13 too. ROCKET AF was a definitive double-blind trial
14 comparing rivaroxaban with warfarin. It shows
15 robust, non-inferiority efficacy for all analyses.
16 It shows superior efficacy while on treatment, but
17 a classical intention to treat does not reach the
18 mark for superiority in that regard.
19 It has a favorable safety profile, the
20 reduction, major clinically important but also very
21 psychologically important to practitioners' outcome
22 of intracranial hemorrhage. And, therefore, it's a

1 proven alternative to warfarin for the prevention
2 of stroke and systemic embolism in patients with
3 nonvalvular atrial fibrillation.
4 Thank you very much for your attention
5 through this presentation.
6 **Clarifying Questions for the Sponsor Presenters**
7 DR. LINCOFF: For about the next half-hour,
8 we'll move onto clarifying questions from the
9 committee for the sponsor. I'll remind you that,
10 at this point, we'd really like to confine the
11 discussion to questions for the committee. We'll
12 have time later -- for the sponsor. We'll have
13 time later for discussions on other issues.
14 Dr. Kaul?
15 DR. KAUL: Thank you, Mike. I have three
16 questions that I would like to ask. But I'd like
17 to start off with, the efficacy of warfarin as an
18 active control is established on the basis of
19 reduction in ischemic stroke, or hemorrhagic
20 stroke, or both?
21 DR. PETERS: The non-inferiority margin
22 calculation was based on ischemic -- on stroke and

non-systemic embolism, so the same endpoint we used in this trial.

DR. KAUL: So the question I have is, does warfarin reduce ischemic stroke or hemorrhagic stroke?

DR. PETERS: It reduces ischemic stroke.

DR. KAUL: So why are we adding hemorrhagic stroke in the assessment of efficacy?

DR. CALIFF: This is a good point of discussion and consideration. You're well aware that many people die from a stroke, and no one ever knows whether it was ischemic or hemorrhagic. We're in an era where, if one does an MRI, for example, the classification itself has significant bias.

So we thought the best thing to do, and what is the standard in the field, for better or worse -- I actually think for better -- is to say, we got a big endpoint that's critical to patients called stroke. The summative effects of the treatment should be beneficial in that regard. I fully understand the argument that one could look

at it differently, and it would not be totally illogical, but it was not a prespecified plan.

DR. KAUL: So the follow-up question is that if you're going to count hemorrhagic stroke towards an efficacy assessment, is it fair to also count that towards a safety assessment?

DR. CALIFF: Fair? Some quotes come to mind about fair, but I want to be clear with everyone, and I think we were clear in our presentation. Right at the beginning, we pointed out, we're double-counting. And it comes up in every presentation and discussion that we have.

From my perspective, the most important thing to know is what's the total effect on stroke. And for a patient and a doctor, it's weighing this array of ischemic and bleeding outcomes. So I think when you talk about efficacy and you talk about safety, I would argue it's totally fair as long as you're transparent about what you're doing. When it comes to the risk-benefit balance question, I think it's pretty irrelevant.

DR. KAUL: Do you have data where you don't

count the hemorrhagic stroke as a bleeding endpoint? And so are the differences in bleeding that don't count hemorrhagic stroke statistically persuasive, in favor of rivaroxaban? And conversely, I also would like to see -- I mean, I like the data that you showed, the graph that you showed, slide 72. But what I would like to see there is also data for disabling stroke that excludes hemorrhage.

DR. CALIFF: I think we ought to start with the first question. Probably, Dr. Nessel can answer this the best, but let me try to reinterpret your question to make sure I got it right. You want to know about the bleeding outcomes with the two treatments if you don't count hemorrhagic stroke as part of the bleeding outcome.

So let's start with that. We don't have that? Anything that you ask for, we can get before the end of the day, and we certainly will.

DR. NESSEL: Just a comment to the committee. Since hemorrhagic stroke was a prespecified composite, or part of the composite

primary safety endpoint, it is included.

DR. PETERS: The hemorrhagic stroke of the intracranial hemorrhages -- hemorrhagic stroke was the largest component of the intracranial hemorrhages, but also subdural hemorrhages, and there were a few subarachnoid hemorrhages, were also numerically fewer with rivaroxaban compared with warfarin, but in the intracranial hemorrhage endpoint.

DR. CALIFF: But we will get you the data specifically that you asked for.

But if I could see the slide that had the disabling strokes, the clinical outcomes.

DR. PETERS: We could do the primary endpoint with the stroke outcomes for the primary endpoint. We do a slide of that for the outcomes. So this is stroke type and subtype.

If we could have the outcome for the primary endpoint.

DR. CALIFF: Disabling, not disabling. We saw it up here just a minute ago.

It'll take a second. There are a few slides

in backup.

DR. PETERS: While we're waiting for that, there was a reduction both in death and disabling, and the non-disabling was more equal.

DR. CALIFF: Yes. I think there is a story here that I think you all recognize from all the novel anticoagulants, which is a reduction in intracranial hemorrhage. And so you can see there's a hefty reduction in disabling stroke, really a quality in terms of non-disabling stroke. So the worst outcomes are favoring rivaroxaban here.

Is that what you wanted to see?

DR. KAUL: No. I wanted to see the disabling stroke data, excluding hemorrhage.

DR. CALIFF: Excluding hemorrhage. Okay.

DR. PETERS: If we could look, we do have ischemic stroke outcome data, if we could look for that slide. We're looking for outcome data by the subtype of stroke, ischemic versus hemorrhagic.

What that will show is similar to what you might expect, is that for the hemorrhagic strokes,

there's a clear reduction in the mortality and the disabling; actually mostly in the mortality because there are fewer of them, and they have a high mortality rate.

Slide on. But for the ischemic strokes, what you can see here is that the -- so there are three populations. There's the on-treatment population, which was most of the data you saw up to last dose, plus seven days up to last dose, plus 30 days, the left side, the rivaroxaban, those three on warfarin, the right side.

So the total number of strokes, you can see there, are 149 versus 161 for the on-treatment; for death, 26 versus 28, fairly comparable; for disabling, numerically fewer, 37 versus 50; and for the non-disabling, 79 versus 75. So, clearly, overall, some numerical reduction in the ischemic strokes and no adverse impact in terms of the outcome; if anything, a trend for a little bit fewer disabling.

Is that what you were looking for?

DR. CALIFF: Let's get the other data he

asked for and show it later.

DR. LINCOFF: Dr. Nissen?

DR. NISSEN: I have two quick questions, one for Dr. Mahaffey. You made the case that doctors don't use warfarin very well, and, particularly, I think in some of the third-world countries. But I was very puzzled by what you said, is that no education of sites on INR management was performed, nor was there any protocol-mandated approach to this.

So my concern is whether you, in fact, made an effort to get the sites to use warfarin as well as they might. Now, maybe I misunderstood what you said.

DR. PETERS: Dr. Mahaffey?

DR. MAHAFFEY: Thanks, Dr. Nissen. Sorry if I wasn't clear. We did comprehensively educate all sites across the world in multiple investigator meetings about the importance of INR management and how to get to a target INR of 2.0 to 3.0. What we did not do, as some other investigators have done, was to provide a cookbook algorithm for the sites

about what to do with a specific INR in terms of a dose change. That, we decided as an investigative group, we weren't going to do, but we did educate people at the beginning of the trial and then, as I said, throughout the trial with subsequent investigator meetings, newsletters about the importance of INR management, and with study materials. And some of these specific details are shown here.

DR. NISSEN: Right. But given the fact that I think you guys knew that the TTR was going to be a critical factor in the approvability of the drug, I'm just puzzled as to why you didn't make an effort to give sites, particularly third-world sites, some guidance about what to do. I mean, basically, you sort of left them on their own, and my concern is that that's why you got such a low TTR.

DR. CALIFF: So, Dr. Nissen, I'm sure you didn't mean to use the term third world, because I don't think that's an appropriate term anymore. And, in fact, if you look at the news today, you

would find what you're calling third world actually owns most of our national debt.

[Laughter.]

DR. CALIFF: So I think investigators around the world would be highly offended by this sort of American-exceptional tone of some of the discussion, not you personally, but it's part of the background here we need to acknowledge.

DR. NISSEN: But Rob, you --

DR. LINCOFF: Let's focus on -- let's focus on the science.

DR. NISSEN: Robert, you showed us that those sites in those countries didn't do a very good job of managing warfarin.

DR. CALIFF: Yes. That's where I want to go.

DR. NISSEN: That's where this all comes from.

DR. CALIFF: But these two concepts are related. I don't think we need to show a slide here, because I think the critical point is -- and I think Dr. Mahaffey misspoke slightly. It's not

1 that doctors do a lousy job. This is a really hard
2 drug to use, and it's especially really hard to use
3 in various socioeconomic personality characteristic
4 circumstances. You know that as well as I do.
5 I have an engineer patient that comes in
6 with his Excel spreadsheet. He's right on the
7 numbers every time. My mother does home INR
8 testing. She's right on the numbers every time.
9 Most people don't live in that world. So what we
10 felt obligated to do was to do better than standard
11 of practice. Given the heterogeneity of the
12 practices in those centers, we spent a lot of time
13 with every site saying how do you do it? Let's
14 make sure what you're doing is okay, but let's not
15 force you to use one specified algorithm. Let's
16 make sure you make the appropriate dosing changes.
17 In fact, I showed you that it wasn't that they
18 weren't making the appropriate dosing changes.
19 They just brought the patients back when it was
20 feasible to do it.
21 Now, one could argue that one should spend a
22 fortune sending a limousine out to pick up patients

in countries that need anticoagulation and artificially make the TTR come out better. We chose to really focus on the highest quality care that we could get at these sites in a global trial.

DR. NISSEN: But you could have told them that if the INR was less than 1.5, that the patient should be brought back quickly. You could have instructed them to do that.

DR. CALIFF: We did instruct them to make every effort to get the patient back. I mean, there were detailed investigator meetings. We had numerous monitoring visits to the sites, but there are limitations on what a site can do in practice in many places.

But, again, I want to emphasize, the TTR is a somewhat artificial measure if you've already adjusted the dose but you measured 30 days later.

DR. NISSEN: My second quick question, if I could, if you could bring up slide CC-47.

DR. PETERS: While that is coming up, the other point we showed is that, within those regions that were poor, in ROCKET, they did as well as in

1 studies that did provide such advice, and they were
2 open label, and they had all the advantages
3 of -- or approximately the same given patient
4 characteristics.
5 DR. CALIFF: Mr. Chair, could I -- I mean, I
6 think that last point is critical, and I'm sorry
7 that I neglected, but I want to pile on, on that
8 issue.
9 If you think that these extraordinary
10 measures make a difference, why is it that in these
11 same regions, we got exactly -- accounting for
12 slight differences in patient populations, exactly
13 the same TTR as the other studies did in the same
14 region? So I think our methods were just as good
15 as the others. We just chose a different route.
16 MR. COUKELL: I'd like to see this slide,
17 exactly the same format, but not for the
18 on-treatment population, but for the full analysis
19 set, the full ITT analysis set through the end of
20 the study. So this exact format.
21 DR. PETERS: So this we have for ITT to site
22 notification for the primary endpoint?

1 MR. COUKELL: Not site notification, but
2 through the end of the trial, not to site -- the
3 full ITT, full observation period.
4 DR. PETERS: I'll have to check to see if we
5 have that. Okay. Slide on. We do.
6 So here's the data, here to look at. I
7 don't know, Rob, if you'd like to --
8 DR. CALIFF: Yes. So this is it. Again,
9 I'd point out that no other trial has carried on
10 for 30 days after the final follow-up visit, so
11 we're including this transition period that we've
12 discussed in detail, but at least to executive
13 committee's satisfaction, the non-inferiority
14 criteria are met. And we've shown you that in the
15 point estimate and confidence interval.
16 DR. NISSEN: The reason I wanted to see this
17 is I wanted to see whether, in fact, there were
18 actually more ischemic strokes in the rivaroxaban
19 arm than the warfarin arm with the full analysis
20 set through the end of the trial. And there are
21 slightly more, 226 versus 220. So the point
22 estimate, in fact, is above 1 for rivaroxaban, if

you look through the end of the trial.

DR. CALIFF: That's true, but I don't think, in practice, there would be the entire population taken off of a treatment and making a transition, so it's a somewhat artificial last month. But, again, we looked. The comparable data that are being discussed today, nobody looked.

DR. PETERS: So the other point, so this analysis with this primary ischemic hazard ratio of 1.03, the upper bound is 1.24, which is still well below the non-inferiority margin. And this does include the ischemic strokes in the post-treatment period, where you saw there was clearly inadequate anticoagulation in those who were on rivaroxaban compared to those who were on warfarin. So even with that included, the hazard ratio was very close to 1.

DR. NISSEN: Later, I'd like to see this, then, in the patients in whom the TTR was in the upper range, up in the mid- and higher 60s. In other words, how does that, a point estimate, drift -- if you use the full ITT analysis -- and

I'd be interested in later hearing from Dr. Fleming who's, I think, been an advocate of that sort of an approach -- if you looked at the patients where the TTR was up in the mid- to upper 60s. That's the question.

DR. PETERS: So you are asking for a quartile analysis?

DR. NISSEN: I think, at some point, I think we are going to see that in the analysis from the agency, but I'm just -- it's a rhetorical question.

DR. PETERS: I'm not sure. Okay.

DR. LINCOFF: Dr. Sager?

DR. SAGER: A few clarifying questions. One is, did the executive committee have a target TTR for this study? I mean, was there a number that you would have liked to have seen, actually, when you started the study, that you thought was appropriate?

DR. CALIFF: Well, of course, we'd like to see as high a TTR as possible, and so shooting for around 60 would be -- but at the time the study was designed, we didn't have all these data about

1 regional differences and severity-of-illness
2 differences.
3 DR. PETERS: I would say that SPORTIF III
4 and V were kind of the trials that these were being
5 designed, following on, an ACTIVE W a little bit
6 just before we started. So we were looking at mid
7 60s, but we didn't set a specific target that I'm
8 aware of for the trial.
9 DR. SAGER: Thank you. And we looked at a
10 lot of data in the presentation, but I just want to
11 clarify that the sponsor would agree that higher
12 levels of TTR are associated with reductions in the
13 primary endpoint, stroke and embolic events, in
14 just the warfarin group.
15 DR. PETERS: If you look within the warfarin
16 subjects, if you look at that center-based
17 quartile, or the line, with the model, you will see
18 that the event rates decrease as TTR gets better.
19 But also, event rates in the rivaroxaban group were
20 lowest in the best quartile because there are many
21 other factors. Patient differences, intensity of
22 concomitant medications, other things. So, again,

as we showed in the presentation, the relative difference -- rivaroxaban line was always lower than the warfarin line, and the relative difference narrowed slightly, but not statistically significantly differently.

DR. SAGER: Thank you. And then the last question is -- and Tom Fleming may want to make a comment here also. With this whole question about 55 percent versus 65 percent, has the sponsor attempted to do any kind of modeling? And, obviously, this is more hypothetical of course, but what might have happened had the TTRs and the warfarin group actually been a bit higher?

DR. PETERS: I think, if we could put up again the slide with the Connolly model, just to review that, the one before this, since you've presented this the first time.

DR. CALIFF: I guess I went through this too fast. But this is not just the sponsor's modeling. It's the executive committee's modeling and independent consultants that also looked at it, most notably Frank Carroll. And the major point I

was trying to make is that, as best we can tell, for the comparison of rivaroxaban and warfarin, there's no significant interaction with center TTR that we can detect. And it's also true of the other contemporary studies. So the best model of the data that you can do would say TTR should not be a consideration when looking at a comparison of the two drugs.

DR. LINCOFF: Mori Krantz?

DR. KRANTZ: Sorry. I just had a quick follow-on on the whole TTR thing, but it's more of an operational question. I think we alluded to, in the briefing documents, that we're using Coumadin clinics quite a bit now. They're very common, these programmatic approaches. So I was curious. What proportion of your sites, globally, had systematic clinics or programs in place? And then what proportion actually used that operationally to dictate changes in therapy?

DR. PETERS: I'll ask Dr. Nessel to respond to that question.

DR. NESSEL: Thank you. Patients in

ROCKET AF, as Dr. Califf indicated, were garnered from 45 different countries. The spread of physicians included in this study included internists, cardiologists, neurologists, and occasionally hematologists. While there was no proscription against using an anticoagulant clinic, they were included. But I'm not certain that we could possibly discern how many sites precisely had a Coumadin clinic versus not.

DR. KRANTZ: And I think in the U.S. certainly we were aware of that, and there's data in the literature that you cited, that Coumadin clinics, irrespective of TTR, do better in terms of outcome. So is there any way we can just get a feel for whether that was completely disregarded or whether that was used in certain cases?

DR. PETERS: To my knowledge -- we can check. But to my knowledge, we didn't record in our database, either at baseline or during the study, what form of management was being done locally. Again, the sites were free. They had to use the blinded device, but that could have been

done out of their clinic or out of their local office. And I don't believe we recorded that information and which type of setting it was used in. But we will check into that.

DR. KRANTZ: Let me just make one final comment. The reason I'm pursuing this is because I think what's somewhat nice about your study is it was designed to be a mimicking of what's happening in real-life practice rather than something that's artificial protocol. If you accept that argument, though, it would be useful to know if indeed we followed what the standard of practice was. That's the reason I'm being persistent on that question.

DR. PETERS: Right. So I do know that, within North America -- you saw the data, and we were 64 percent in range for North America as a whole, across the spread of sites. For the U.S. only, it was 63 percent. Canada was 66, southern North America.

So that's better than what was shown kind of in the -- that's equivalent to the anticoagulation clinic data that was shown and better than the

usual care data. And we did have, obviously, a range of sites, which some were on the lower end and some were on the higher. But on average, in the U.S., the time in therapeutic range was 63 percent, which we think is very representative of a high level of management of warfarin.

DR. CALIFF: If I could just add briefly to that, obviously, these were mostly high-volume sites. People that participate in clinical trials have to have a certain number of patients who are eligible to participate. So I don't think we've formally collected whether they're something called a Coumadin clinic. They were high-volume sites, and we tried to show you that, actually, in the dose adjustment, they did what they were supposed to do, even in the sites where they didn't bring the people back as quickly.

So we think the anticoagulation management was better than standard practice. I would call it real world plus an adjustment for enhancing practice to be closer to clinical practice guidelines.

DR. PETERS: I can make one additional comment that came to mind. In terms of that time to return for an INR less than 1.5, there is literature from a large group of coagulation clinic data in the U.S., and the average time there was right at seven to eight days, which very closely matched what we saw in our study in the North American population.

DR. LINCOFF: Dr. Fox?

DR. FOX: Yes. This is quite apropos to the line of questions in the last couple of minutes. So you showed us the very nice observational data in the United States, kind of giving the context of a comparison between the best you can get, so to speak, and a rigidly controlled clinical trial atmosphere in the mid- to high 60s of TTR, versus Coumadin clinics kind of getting close and usual care, kind of falling down to about what was seen in the trial overall, or maybe lower.

Are there similar data for other regions in the world that would place some context around the TTRs that were achieved in those regions?

DR. CALIFF: The short answer is yes. I had a slide on this that we took out, and it probably didn't make it to the deck, but I can show it to you. At the European Society of Cardiology that just occurred, there was a very large registry that was shown with the INR data from each region.

Do you think we have that? If not, we can get it and show it to you. But on average in the rest of the world, there's a tendency to under-anticoagulate, very much mimicking what we saw in the trial, that we had to overcome with a dose adjustment approach that we used with the blinded INRs.

DR. FOX: That's pretty much what I would have expected you to say, and I think it's relevant to Dr. Krantz's comment about, which do you want. Do you want a somewhat artifactual situation of a clinical trial, or do you want it to resemble more real-world medical practice? And I think that's an important distinction that we should talk about later today.

DR. LINCOFF: Dr. Emerson?

DR. EMERSON: This is in reference to slide 82, and I can ask the question better if you bring that up.

DR. PETERS: So slide 82 from the core? Slide on.

DR. EMERSON: So my question is, have you done an analysis to look at these sorts of things, within both the completed study medication, separately from this early discontinuation, to see what the event rate is during the first 30 days after stopping, as a function of cumulative drug exposure?

DR. PETERS: We have Kaplan-Meiers of those time periods over time, but not based on cumulative -- nothing by cumulative drug exposure.

DR. EMERSON: Is it possible to get that?

DR. PETERS: Cumulative drug exposure to their stopping to -

DR. EMERSON: Yes, prior to their stopping. So the difference between this --

DR. PETERS: How far they were into the trial before this stopped?

DR. EMERSON: That's correct. And I realize that you have staggered entry, so even at completed, there's quite a variation --

DR. PETERS: We don't have analyses like that currently. The completers, obviously, went all the way to the end, to the site --

DR. EMERSON: But they didn't all start at the same time?

DR. PETERS: No. Correct. But enrollment finished about June of -- they were in at least a year from the last enrollment to the end. So we would have to look into that. We don't have anything that looks at the exposure before they discontinued.

DR. CALIFF: Just to make -- I mean, we have the capability of doing analysis, I'm told, if we need to. So just to understand exactly what you're asking for, and maybe why, so we can --

DR. EMERSON: So if there is any sort of a rebound or an aspect like this, the disparity -- I mean, now you do actually have a slight excess, even in the early study medication; but what this

disparity is, that if we were to buy that you had different results from the early study medication than you have when everybody's stopped, as you notified the sites, that raises the question of is this an aspect of cumulative dose exposure?

DR. CALIFF: Yes. So I think we could come as close as possible, but the one caveat is, often having an event is a reason to stop study medication.

DR. EMERSON: No question. And you have double the rate among the early discontinuation or, actually -- yes, quartiles.

DR. PETERS: The other thing I could comment on this is that at the end of the study, 90 percent plus of people did transition to another VKA, which was mostly warfarin. So it was 80 percent warfarin to warfarin, which you would expect. They were doing well. They weren't having events. They were on anticoagulation for the years of the study that they should be on anticoagulation to continue.

The ones that discontinued in the middle of the study, both groups discontinued, only about

50 percent went on a vitamin K antagonist. So it was a very different situation, discontinuing in the middle of the study. And at the end -- and there didn't appear to be -- there wasn't a gap in the anticoagulation in the middle as there were at the end because most people stopped, or the majority of people stopped both drugs and didn't resume it.

DR. LINCOFF: Just to comment on time, we're running a little bit behind because one of the presentations ran a little long, but there are a number of additional questions here that I have. So we don't have a public hearing, anyone who has signed up in the afternoon, so we'll have time to make up.

What I'm going to do is take three people now, and then we'll take the break that was scheduled for five minutes ago, and then we'll come back and continue these questions to the sponsor. So we'll move everything back a little bit.

So Dr. Kaul?

DR. KAUL: Two quick questions. Were there

any centers that participated both in the RE-LY trial as well as the ROCKET AF trial?

DR. PETERS: I don't know the answer to that. Out of a thousand centers, I would guess there might be some.

Dr. Nessel?

DR. KAUL: The question I have is that if there were, what was the TTR in the centers that participated both in the RE-LY as well as the ROCKET afib trial? And the reason for asking that question is that, the difference in the trial design, how the maintenance warfarin dosing was managed, one centralized through protocol algorithm and the other based on the local standard practice, did that impact anticoagulation management?

DR. NESSEL: That's a superb question, Dr. Kaul. I wish we had had the access to that data, but the sites, of course, are required to file FDA Form 1572. But other than that, we did not collect all the other clinical trials in which they were participating, so we only have half of the data that you're looking for.

DR. KAUL: Second quick question, with a quick answer if possible. How many of the subjects with a major bleed either resumed treatment or had no interruption? And of those who resumed treatment, what was the percentage with another recurrent major bleed?

DR. PETERS: Dr. Nessel will respond to the bleeding question. And then, Dr. Califf, you wanted to maybe make a follow-up comment?

DR. NESSEL: So I'll ask my team to look for a slide of multiple bleeds. Other than the event of intracranial hemorrhage, the protocol did not require that patients come off study medication for a major bleeding event, or even for a myocardial infarction, for that matter. So it was, of course, at the treating physician's discretion, but they could be kept on study medication for after a major bleeding event. I'm thinking of the slide of bleeding events that occurred more than once per subject.

DR. PETERS: The other thing to say, while we're waiting for that slide, is for the

interruptions, they obviously were for a variety of reasons. Some of them -- for the interruptions greater than or equal to three days, the largest reason was interruption for surgery or invasive procedure, and then also adverse events, either bleeding, or non-bleeding, or another fairly common cause. Between those three was about 75 percent of the interruptions.

Slide on.

DR. NESSEL: So Dr. Kaul, in this slide, you will see the total number of subjects with one, two, or three major bleeding events. And, of course, since this is the safety on-treatment populations, all these patients remained on study medication for each subsequent major bleeding event.

I'll further remind you that the major bleeding component was hemoglobin drop requirement for a transfusion of two or more units, leading into a critical organ or site, or a fatal. Obviously, they wouldn't be counted here if they don't bleed.

DR. CALIFF: Just a quick comment. I wouldn't want the panel to leave your first question thinking the only difference between RE-LY and ROCKET was giving an algorithm for INR. Doing a blinded, double-dummy trial is very different than an unblinded trial. In an unblinded trial, everyone knows the INR, the treating physician, the patient, the study center. It's a whole different set of logistics that one has to go through in a double-blind, double-dummy study, which is the reason why I sort of favor having one of each, if you could possibly do it, because I think there are big differences in trial logistics in those circumstances.

DR. LINCOFF: Mr. Coukell?

MR. COUKELL: Thank you. Two quick questions. So the FDA has asked us to think about rivaroxaban in the context of having dabigatran on the market. We haven't finished or maybe even started that discussion yet, but is there any data that suggests there's a subpopulation, pharmacodynamic or clinical data, that would

respond better to this drug than the other, for which this drug might be particularly suited?

My second question is, in ROCKET, you've deliberately selected a higher-risk population, where, presumably, the effect of both drugs would be greater. As I understand it, the indication you're seeking is in the general afib population. And what's the basis for that?

DR. PETERS: Dr. Califf?

DR. CALIFF: So let me try to respond to both questions. The first question, as I've said, we are fundamentally opposed to cross-trial, indirect comparisons, but there is one circumstance that's relatively obvious I think to everyone. And that is someone who can't tolerate dabigatran, can't take it, and so we have an alternative non-warfarin opportunity for those people.

There's a significant rate of GI side effects, for example, with dabigatran. And that's a minority of the population, but it's an obvious group where, if you can't take a drug, having an alternative, in our view, could be important.

Your second question again?

MR. COUKELL: The second question was that you've studied a higher-risk subset of the afib population, but you're seeking an indication in, as I understand it, the general population.

DR. CALIFF: I'm sorry. I forgot that. Two key points here I think that the committee should think about. The first is that while the two populations -- one is high risk and the others are more general -- there is a lot of overlap in those.

Remember that, per CHADS1, the patients that we excluded right now, either anticoagulation or aspirin is in the guidelines. That may change in the near future, but that's the current recommendation. So all the trials are dealing pretty much with CHADS2 and above as the primary population. We just excluded many of the CHADS2 people on purpose to get this higher-risk population.

So we think there's enough overlap that the general indication holds, and there is no heterogeneity in treatment effect across any of the

major subgroups that we looked at. So we think the result is generalizable.

Second point, just in terms of policy, we routinely do trials in namby-pamby populations of low-risk patients for a lot of reasons. It's easier to do the studies. And then we extrapolate to the 80-year-old person on 15 other medications. We think it's time to begin to think about doing trials in the more sick people and having enough overlap that you can talk about the general population. We're somewhat responding to what practitioners are saying, which is, you're doing these trials. They're not relevant to the patients that are of biggest concern to us.

DR. PETERS: I just have a follow-up point. I think your first question was also about were there any subpopulations within ROCKET that might look particularly better or worse. And as Dr. Califf showed for the efficacy, the efficacy was very consistent across all the subgroups. The safety subgroups are in the briefing book as an appendix, and those were also very consistent. So,

overall, our efficacy and safety was very consistent across all the baseline factors that we looked at.

DR. LINCOFF: We'll take one last question now, before the break, from Dr. Temple, and then I have Drs. McGuire, Papademetriou, Fox, and Fleming, and many others for afterward.

DR. TEMPLE: It's possible this is a matter for further discussion later. This is somewhat following up on a question Steve raised and Sanjay hinted at. The primary endpoint -- and we don't disagree with this -- was both kinds of strokes together, so I have a couple of questions related to that.

Are you reasonably convinced you can tell the difference, with available approaches, between hemorrhagic and thrombotic strokes? That's one question.

Second, there's been not that much attention to the implications of, for example, time in therapeutic range on the two kinds of strokes. If you were too low, if you had a low INR, that might

affect the comparison for thrombotic strokes. You wouldn't expect it to advantage rivaroxaban for hemorrhagic strokes. And yet, almost all of the effect on stroke is on hemorrhagic stroke.

Steve pointed out that when you add the people after withdrawal, the effect goes slightly the wrong way, but it was never very strong in the first place. The hazard ratio is only .95 for thrombotic strokes, which is not very impressive.

In other situations, for example in dabi, dabi had an effect on both thrombotic and hemorrhagic strokes. But even though hemorrhagic strokes were less frequent, half of the effect of that drug was on the hemorrhagic strokes. And the lower dose, which many people in our committee thought should be approved, actually didn't have any effect on thrombotic strokes at all. It went the wrong way. So all of the effect of the low dose was on that.

I also glanced briefly at the apixaban data. We haven't reviewed it, but most of that effect is on hemorrhagic strokes also. So the considerations

of time and range, that all seems to be a very important part of all of this. I don't know if any of these differences are real, but it's not being discussed enough in my opinion. I just want a preliminary view on that, and we can talk about it later.

DR. PETERS: So that sounds like a long answer, but I'll have Dr. Mahaffey come. He led the CEC efforts, so I think in terms of the confidence of assigning a diagnosis of ischemic versus hemorrhagic or unknown, we have confidence. It's hard to do, but we did it.

Dr. Mahaffey?

DR. MAHAFFEY: Dr. Temple, obviously, an important issue across a variety of different endpoints and adjudication efforts and different disease states in patient populations. We did go to rigorous lengths to collect both clinical information about stroke events, but also radiographic imaging information about the strokes as well. And I think we did a fairly good job, and we did a very rigorous job in this trial. All of

the strokes were looked at by a neurologist who had experience in event adjudication, as well as understanding CEC issues and anticoagulant therapy in patients with atrial fibrillation.

The second thing, then, is to point out that the number of unknown strokes -- slide up, please -- in this trial is also very low. And I think it reflects the rigor that we went after this. I think, though, you're right.

One of the subtle challenges that we have is differentiating between ischemic strokes with early hemorrhagic conversion or large hemorrhagic conversions and primary intracranial hemorrhages. And we spent significant time going over these events to try and differentiate those, probably erred on the side of calling them intracranial hemorrhages if there wasn't clear evidence of an infarct zone.

DR. PETERS: We could actually show the breakdown of the primary efficacy endpoint, shows the unknown. That was trying to classify different types into mechanism, but the number of unknown was

the minority, compared to hemorrhagic or ischemic, being able to show that.

So while we're waiting for that, if we could -- so the later part, we did do some analyses looking at, by quartiles, hemorrhagic events above INR, above time, above 3, and for ischemic stroke, INR below 2.

So this slide on , just quickly. In the primary endpoint safety on-treatment, there are only 18 events that the CEC didn't assign to either ischemic or hemorrhagic.

DR. CALIFF: Just two points. We want you to believe that we went to the end of the earth to try to classify these and that we had great people doing it. And we did come up with a small number of unknowns. But anyone that does this kind of work knows this gray zone between hemorrhagic conversion, particularly with better imaging methods. If you do reliability studies, it's far from perfect. So there is crossover.

I would like to go back to the slide of the long plots, the one on intracranial hemorrhage that

I showed, because it may have gone by too fast. I think it's very pertinent to the point you made, Bob. One might suppose that if you pushed anticoagulation, that you would cause more intracranial bleeds.

What we show in this slide, just to review it again, for intracranial hemorrhage, is that trying to achieve a time in therapeutic range, our best sites are the ones that look the worst in the comparison of rivaroxaban.

So your point that this is complicated I think is quite good. And understanding what the implications are, as with many other drugs, of trying to hit an intermediate biomarker by pounding the patient, perhaps in circumstances where it may not be beneficial, can be complicated.

So the overall impact on the patient may be a few more intracranial hemorrhages, relative to what a once-a-day drug can do.

DR. PETERS: So this slide shows -- and if we could have the hemorrhagic stroke equivalent. Slide on.

So this is a quartile analysis of ischemic strokes, as diagnosed by CEC, by quartile time below the range of 2. You can see, the ischemic stroke in the overall was pretty close to the line of 1, and it's pretty close to the line of 1 here.

The most time below 2 is the bottom, so the event rate was higher, 1.7, and it's a bit lower, but you can see rivaroxaban versus warfarin, not much difference. Then the next slide is a little bit different for format, but the same similar analysis, where you can see in the gray. And this is now hemorrhagic stroke by quartile time above 3, so above 3.0. And the warfarin rate is not strikingly different. And actually may be even a little bit lower with the quartile that had the most time above 3,. and the rivaroxaban rates low and comparable.

But these analyses, we did look at some of these and didn't do a lot of because there's relatively little time at the two extremes. So I actually think, as in the literature, the time in the range, 2 to 3, is probably the most reliable

for looking at how things spread out by TTR.

Slide off.

DR. LINCOFF: So we'll now take a 15-minute break. Panel members, please remember, there should be no discussion of the meeting topic during the break, amongst yourselves or with any member of the audience. We'll resume at 10:55 sharp, and we'll have time for four additional questions. We'll confine it to that so we can move on with the FDA afterwards.

(Whereupon, a recess was taken.)

DR. LINCOFF: We'll continue with the sponsor on the hot seat for the next four questioners, and then we'll move onto the FDA presentations.

So Dr. McGuire?

DR. MCGUIRE: Thanks, Mike. I am going on with this question about TTR. What I've been convinced, both in the briefing documents and the presentations, is that this measure will join the graveyard of other biomarkers proven invalidated as surrogates. And before we throw it away, one of

the reasons surrogates fail is because the concept fails. And I'm not sure many of us believe that. So the question is, why does this fail?

I think there are two important considerations that might be addressed. One is, the treatment of all patients outside of therapeutic range, similarly, is a little bit confusing to me, that people below therapeutic range should bias in favor of the study drug for efficacy, and above for safety. And we saw a little bit of that data at the end here.

But what I'd like to see is the overall efficacy analyses and safety analyses stratified by quartile of center performance for above and below therapeutic range. We have one table in the FDA document below therapeutic range. We saw some above-therapeutic range data here just recently.

The second issue is the method of interpolation. And would it be more reasonable to use the last observation carried backward, as the interval data represents the dosing the patient is on during the interval, being interpolated? And I

wonder if anyone would have a comment if that strategy has ever been employed.

DR. PETERS: So if you could just clarify the last part again, and then I'll have to clarify further.

DR. MCGUIRE: Yes. As Dr. Califf brought up the Rosendaal method biases against time in therapeutic range for patients who undergo a dose adjustment, especially when the interval between evaluations is long. So it's chasing a ghost to think that we'll ever get anywhere close to 100 percent using that method of assessment of therapeutic adherence.

DR. CALIFF: Yes. The second question is a softball, because it's interesting, and I would agree with it that that would actually be a way to get around the bias. What you're saying is, if you have a dose adjustment, and you come back in 30 days, and you're in range at 30 days, the best extrapolation would be looking backwards, not the one looking forward. But we haven't done that, and I don't think we can do it in time for this.

Let's see. Then your first question is -- I think, a really good concept. It's not one that we -- despite having a lot of people come in and look at our data in the executive committee, it's not one that we formally addressed.

DR. MCGUIRE: I would just add, in this study, unlike several others, the ratio of time below to time above is 2 to 1, where we've commonly seen 3 to 1 and 4 to 1. So for most prior studies, time out of therapeutic range almost equated with time below therapeutic range.

In this study, you have probably more power than most to partition that out. A comment earlier that there weren't enough events, about 15 percent of the time -- by my back-of-the-envelope math, about 15 percent of the time, out of therapeutic range was actually above.

DR. CALIFF: We have an interesting plot here, which I believe we should put up, but with the caveat. I'm sure you all understand. We have something like 4,000 slides, and having spent so much time on it, we're anxious to show all the

data. But some of these are very complicated analyses that would be very hard to explain.

But let's put this slide up. The caveat here is that once you begin to slice by these multiple different things, it takes a long time to even understand what you're looking at in an analysis. So I think you're actually asking for even slicing it one more time into safety and efficacy. But for the primary efficacy endpoint, what you see here -- and I credit J&J for having done this. I didn't even know we had done it. What you can see is that, for time in INR less than 2, by quartiles of centers, you can see that rivaroxaban, the light blue line, is always better than the gray.

Now, this is on-treatment analysis, so I'll point that out. And even for INR greater than 3 times by quartiles, rivaroxaban comes out ahead again. But this is integrating the safety and efficacy into a common endpoint.

DR. PETERS: The top panel is exactly what I was asking. I don't know if it's possible to do

the same analysis for safety for the bottom panel. And the second, to follow-up, were interaction tests done for that top panel analysis?

DR. MCGUIRE: The top panel, for ischemic stroke, we did show, if we want to show that one again, was the equivalent analysis by quartiles, for ischemic stroke with time below 2, if we could pull that one up again. And I think you were asking for median, which would essentially just combining the --

DR. PETERS: Not necessarily. This top panel is exactly what I'm interested in.

DR. MCGUIRE: So slide up. So this is for ischemic stroke now, not the primary endpoint. So it's just the ischemic events, and it's the time below 2; where, at the bottom is the most time, 34 to 100 percent above -- below 2. And the top row is the best time. And there's really no difference in the treatment effect, relative to warfarin, across those four quartiles.

DR. LINCOFF: Dr. Papademetriou?

DR. PAPADEMETRIOU: Thank you. My question

is related to the period at the end of the study, when the patients were transitioned to long-term warfarin. And it was said that the investigators were allowed to treat patients with heparin during that period, but they were still blinded.

Do we know how much heparin was utilized and if there was any difference between the two groups, rivaroxaban and Coumadin, and if the utilization or the lack of utilization of heparin was associated with any events or hemorrhages?

DR. PETERS: If we could do the completers only.

So the use of heparin was allowed in the protocol for interruptions during the study. The use of bridging therapy at the end was also allowed and even recommended as appropriate in the communications from the executive committee. But in the end -- slide on -- this is all anticoagulants besides VKA for the completer group, and you can see less than 3 percent in the rivaroxaban and less than 2 percent in the warfarin were actually used, some other anticoagulants.

So, in practice, even though it was allowed and even recommended to follow guidelines, like ACCP, it was considered at the investigator level that they didn't implement that recommendation.

DR. CALIFF: So in retrospect, this is obviously very disappointing, but I think it's important for panel members not intimately involved in this field to understand that even in the case of warfarin prolonged interruption, bridging is an unproven strategy, with two randomized trials currently ongoing. And so what we're seeing in clinical practice is that, for a variety of reasons, people are not going to the trouble to do it. We now -- I think seeing this, we think the best strategy for rivaroxaban is going to be bridging.

DR. LINCOFF: Dr. Fleming and then Dr. Fox?

DR. FLEMING: Could I have slide 39? And while that's coming up, just a couple of quick thoughts. And that is, I think, Rob, you pointed out cross-study comparisons are treacherous. That's certainly true. Of course, that fact

undermines the integrity of non-inferiority analyses.

The second fact is, bridging or the blinding of the trial could impact TTR, you were saying, as you were noting that, to your credit, you have a blinded trial. Interestingly, though, for ximelagatran, SPORTIF V that was blinded had a higher TTR than SPORTIF III, that was open label.

The question that I want to get at here is, Dr. Marciniak were here, as he has stated publicly in some previous meetings, one of the biggest issues that can undermine integrity of research is missing data, missing follow-up.

The withdrawal of consent rates are very high here. I assume these 700 plus people did not have additional follow-up after they withdrew consent, based on your slide.

Is that correct?

DR. CALIFF: Only in the case where the study participant was the patient of a doctor who talked them into coming back for follow-up, because what we did was, of course, contact the sites once

we realized the withdrawal of consent; by the way, let me just mention, commonly seen in North America now and increasing in Europe due to laws that have been passed, making it easy to withdraw in the midst of a study.

So there were some cases, we made every effort to reach the doctors and have them talk to the patients, but by law, we can't follow --

DR. FLEMING: So, in fact, as I'm understanding, these people were not followed?

DR. CALIFF: Right.

DR. FLEMING: So my question is, what was the nature of the instructions that were given to investigators about what the exact proper context is for use of withdrawal of consent?

DR. CALIFF: Instructions were very clear, that withdrawal of consent should only be an option if the patient demands to stop the study drug and have nothing to do with the trial for one reason or another. I think all of us involved in clinical trials are aware that, particularly in long-term trials now, particularly in North America, one

thing you'll notice on this slide is there are 89 patients from closed sites and 6 from retired sites. We're in an environment where investigators feel under assault, and sometimes they take the easy way out of telling a patient, if you just sign this, then I don't have to follow you anymore, and my work is done. And we did our best, as we are in all studies, to keep that from happening, but people should be aware that this is an issue.

DR. FLEMING: So I am focusing only on the withdrawal of consent, not the closed sites. So you're saying there were consistent efforts made to ensure all investigators knew this had to be patient initiated and that this did not just mean they didn't want to continue treatment or follow-up, but they did not authorize you to follow them, and that was the basis for withdrawal of consent.

Were all investigators instructed about that?

DR. CALIFF: Yes. All investigators were instructed multiple times, and even after

withdrawal of consent, investigators or their doctors were re-contacted if they were having other reasons to see the patients, to try to do their best to get them back into the study, realizing that as a coordinating center or a sponsor, we're not legally allowed to contact the patient at that point.

DR. FLEMING: In the interest of time, could I have slide CC-50?

DR. PETERS: If I can make -- while we're getting that --

DR. FLEMING: Very brief.

DR. PETERS: There were a number of patients this large or a little bit larger that withdrew consent to study drug, but did continue in the study. So what's reflected there is those who actually refused continued participation in the study.

DR. FLEMING: I saw that. Yes.

DR. CALIFF: I'm sorry, Tom. But you made a comment, and then didn't ask a question that I think deserves a response from us about the two

SPORTIF trials, because it was frankly an unfair debating tactic, I felt.

Yes. The blinded trial had better TTR than the unblinded trial, but when you look at the composition of the sites that were involved in that trial, in those two trials, it follows exactly the pattern that we showed here. We've done pretty extensive analysis on this. So there, the blinding and unblinding was overwhelmed by the regional distribution of the sites.

DR. FLEMING: On slide CC-50, you provide results, again, for the safety on-treatment. Dr. Nissen has saved me some time here because I had the same question he did about CC-47, which is to look at the results in stroke, not only by on-treatment per protocol, but ITT.

My understanding here, when we're looking at all-cause mortality, is there are 458 events. When I look at your briefing document, on page 94, in terms of deaths, there are almost triple that number of deaths. So it's perplexing to me. There are about 25 percent -- from Dr. Nissen's answer,

about 25 percent more strokes when you look at the ITT. There are 200 percent more deaths. I'm perplexed about why there are so many more deaths. And then the actual additional evidence that I want to get is, can you give us the ITT results for MI? And then I want to come to the last part of the question, which is I'd like to see those results by U.S. only.

So there's really probably three parts to the question. Why are there triple the number of deaths when you look at ITT? Secondly, what is the MI ITT? And thirdly, what are death and MI in the U.S. sites?

DR. CALIFF: That's a trifecta, as we say in basketball. But if we can put up the first slide that's up here, I think the main point to make is, in a sick population, we have non-fatal events as a primary endpoint. A lot of non-fatal events occur. Let's take stroke as a good example. If you don't die from your stroke, there's a good chance you're going to come off of study drug, not necessarily because anybody wants to do it, but you may no

longer be a candidate for anticoagulation. You may be in a rehab center that can't manage the trial. There are multiple reasons.

So what we had was a lot of deaths that occurred either as subsequent events to the primary event of the trial. Or, remember, these were people on nine medications, on average, with multiple other diseases who died from a variety of causes.

DR. FLEMING: So what percent of person years were occurring in the ITT? So what you referred to in the briefing document as the ITT analysis used for efficacy, the ITT population used for efficacy analysis, that has three times the number of deaths than the number of deaths that are on treatment, per protocol. What's the difference in the person years of follow-up, in total, for the ITT versus for the per protocol?

DR. CALIFF: We'd have to calculate that. That's an interesting question.

DR. FLEMING: So if you could get it for me later on.

DR. CALIFF: We will.

DR. FLEMING: So, specifically, for MIs, what is the MI distribution by treatment arm for the ITT for efficacy analysis population?

DR. CALIFF: Regardless of treatment exposure is what you're saying. But just to make sure people absorb this slide, when taking into account all deaths, there is, numerically, a difference, and you see it there. Let's put this slide up.

I want to point out, again, when we show regardless of treatment exposure, we're including that 30 days, which is artificial compared to what would happen in practice. But here you see the results for myocardial infarction. Hazard ratio is 0.89; upper confidence interval 1.13.

DR. FLEMING: And then the last question that I had was, when you look at your on-study per protocol, within the U.S., the mortality differences that you see at .85 shrink to .97. The MI from .81 shrinks to .95 when you look within the U.S. by the on-study.

What are the ITT results for U.S. for survival and for MI?

DR. CALIFF: We're going to put the slide up, but I want to be clear again. When you're asking for ITT to end of study plus 30 days, you're asking for something different than any other trial has included in this field. So we have an artificial additional 30 days, but we do have the data, and go ahead and put it up.

So you can see that -- and this is for North America, not U.S. These are the data that you have. And as you would expect, there were events in those additional 30 days that moved the point estimate.

DR. FLEMING: So in the U.S., MI is in the wrong direction, actually. That's 41 -- that's excess, more on rivaroxaban, 41 and 36.

DR. CALIFF: Yes. It's 41 versus 36.

If we could, put up the slide that's up now. So this would be the data, as other trials have looked at it, which you can see here. The MI is now 30 versus 29, not counting those additional

30 days after discontinuation.

DR. FLEMING: But even as this analysis says, in the U.S., MI and all-cause mortality are neutral.

DR. CALIFF: Yes.

DR. PETERS: But, again, this is a subgroup. About a thousand patients in each group were from the U.S. So compared to the overall trials, it's much smaller numbers.

DR. LINCOFF: Finally, Dr. Fox?

DR. FOX: Yes. Thanks. A follow-on to one of the points that Dr. Fleming made, I was afraid I was going to be throwing this into the middle of the TTR conversation, so I'm glad he brought up the issue of missing data as it relates to patient retention in these trials.

Speaking on behalf of industry, I think it's important for everybody to recognize that keeping patients in these long-term trials is extremely difficult. And I take it, as a demonstration of the investigator's effort, that they had as few patients withdrawing consent and refusing to be

followed up as they did.

Having said that, it does amount to about 5 percent with the simple math I did from the flow diagram. I wonder if the sponsor can reassure the panel, if you've done any sensitivity analyses on those patients that withdrew consent, if you assume that they were all on treatment and you assume the worst case scenario, that all the patients on riva had a primary endpoint, and all the patients on warfarin didn't. As artificial as that might sound, if it still comes out that your non-inferiority analysis is intact, I think the panel would be reassured.

DR. PETERS: So slide on. So this is an analysis, as you describe, and there's a very similar one in the FDA briefing book that gives a very similar result. So this takes the people with the missing follow-up time that we were talking about previously, and the observed event rates are across the first row, so 2.2 percent, rivaroxaban, 2.4 percent, warfarin. This is the ITT to site notification, so the hazard ratio is .88.

The second row assumes that in those -- is the number 77. You would have to have 77 events in the rivaroxaban group with that event rate, zero events in the warfarin group to actually just hit the point where you're inferior, the lower bound hits 1, which is just a little bit before you would hit the non-inferiority bound of 1.38. So the FDA came up with a number of, I think, 80, if I recall. We came up with 77 because we stopped when the upper bound was 136. So that's a pretty extreme result, to think that we would be missing almost 80 events out of -- we collected only about a little over -- between 2 [200] and 300.

Slide off.

DR. LINCOFF: Thank you. We will now proceed with our presentations from the FDA.

FDA Presentation - Preston Dunnmon

DR. DUNNMON: Good afternoon, Mr. Chairman, committee members, ladies and gentlemen. My name is Preston Dunnmon. I was the medical safety reviewer for this NDA. And I'd like to talk to you today and review with you the data that either was

available or has become available regarding dose selection in the ROCKET trial, dose selection, obviously, being something important, potentially to all outcomes in clinical trials.

We have four sources of information for dose selection in this trial. And those four I'll cover with you include the PK/PD profile of the drug, the two small venous thromboembolism studies that were essentially the basis for dose selection in ROCKET. Subsequently, the ATLAS-ACS I TIMI 46 dose ranging study for the Acute Coronary Syndrome Development Program became available. And that's something that was of interest because this is a high-aspirin use population, and there was also a high-aspirin use within ROCKET. And then, finally, data comes from within ROCKET itself, in that the prothrombin time was actually confirmed to be a surrogate for exposure within ROCKET, as far as the range of exposures within ROCKET. Because of that, exposure-driven efficacy and safety analyses were performed, and I will show you those.

To begin, the half-life of rivaroxaban is

1 fairly short, 5 to 9 hours in healthy people. And
2 arising from that characteristic, what you see here
3 is a prothrombin time-course of an average of
4 prothrombin times over time. And this is a
5 simulation that was based on a combination of
6 matched PK/PD data when it was available from
7 ROCKET and then also the previously reported PK/PD
8 model from phase 2 studies that was confirmed
9 within the ROCKET trial itself.
10 The once-daily dose of 20 milligrams per day
11 is demonstrated here in the red hatched line, the
12 simulated 10-milligram BID dose in the blue solid
13 line. And what you see here is the expected
14 greater fluctuation between peak prothrombin times
15 versus trough prothrombin times over the
16 interdosing interval.
17 To put this into perspective, dabigatran has
18 a longer half-life than rivaroxaban and is dosed
19 BID. The other thing that you'll notice from this
20 simulation is the fact that, during the second half
21 of the interdosing interval, either in this slide
22 from hours 12 to 24, 36 to 48, or 60 to 72, the

1 drug's impact on the PT itself is actually fairly
2 minimal with the daily dose administration.
3 Next, we have data from the two small venous
4 thromboembolism studies that were used to select
5 the 20-milligram per-day once-daily dose that got
6 taken forward into ROCKET. There was no dose
7 ranging done in atrial fibrillation patients, and
8 so these two studies were the basis for the dose
9 selection.
10 With respect to those two studies, it's
11 important to understand that the 20-milligram per-
12 day dose was actually the lowest dose that was
13 tested in either of them. There were a small
14 number of outcome events, and neither of these
15 studies demonstrated an efficacy dose response.
16 What you saw this morning in the slide where
17 you saw 10-milligram BID data together with
18 20-milligram per day data was actually data from
19 two different studies. One of the doses was in
20 study 11223, the 20-milligram per-day dose in
21 11528. And so when you actually confine your
22 analysis to looking at a single dose of the drug

1 that was given as either daily dose or BID dose,
2 that only occurred in one dose in study 11223. And
3 when you look at those results, which I'll show you
4 here, you get better efficacy with split BID dosing
5 and numerically fewer non-major bleeding events
6 with BID dosing.
7 This is the trial study 11223, the efficacy
8 results from that trial that were based on
9 resolution of thrombis by ultrasound. And what you
10 see here is that the 40-milligram per-day total
11 daily dose was the only dose given both split and
12 as a daily dose. And when you look at those
13 patients who were improved, by the definition of
14 this study, 59 percent improved with the split BID
15 dose versus 44 percent with the daily dose. And
16 that, retrospectively analyzed by Fisher's exact
17 test, was statistically significant.
18 When you look at the bleeding from study
19 11223 for the 20-milligram BID versus 40-milligram
20 per-day dosage, using definitions of bleeding that
21 were very similar, if not identical, to ROCKET,
22 there were numerically fewer bleeding events in the

split dosing group, 11 versus 14. But here, again, the number of events was exceptionally small.

Proceeding forward, then, in time, this day was not available at the time ROCKET was initially designed. It became available subsequently. This is the data from ATLAS-ACS I TIMI 46, which was from the Acute Coronary Syndrome Development Program with rivaroxaban. And it was of interest because when you look at the ROCKET population, over 4,000 people in ROCKET were taking aspirin concomitantly with their rivaroxaban, more than 2,000 people per treatment arm. And it somewhat makes sense because if you look at the CHADS scoring system, encompassing age, diabetes, and hypertension, those are things that not only portend a bad outcome as far as strokes are concerned with atrial fibrillation, but those are also things that are certainly risk factors for the development of atherosclerosis and acute coronary syndrome. So it's not surprising that there would be overlap between these populations and if there would be high aspirin use in ROCKET.

There was less triple therapy used in ROCKET with aspirin, clopidogrel or thienopyridine and rivaroxaban. And so of interest was this data from stratum 1, from TIMI 46, where this was concomitant aspirin only. And what you see here is a total daily dose analyzed of 5 milligrams per day, 10 milligrams per day, or 20 milligrams per day, given as either a once-daily dose, as you see in this first set of columns, or as split dosing, as here in this set of columns to your right.

What you see is no matter how you give this drug, whether you give it once daily or BID, when you increase the total daily dose from 5, to 10, to 20 milligrams per day, the hazard ratio for bleeding increases in the daily dose from 1.67 up to 6.69, going from the minimal dose to the maximal dose tested. And you see that very same directional trend in the twice-daily dosing, going from a hazard ratio of .81 for bleeding with twice-daily dosing of the total daily dose, up to 6.43 when 20 milligrams per day is split as 10 BID.

The thing of note, though, is to compare

1 these two shaded columns horizontally, because
2 that's comparing the same total daily dose, split
3 to be given either as a single dose together or
4 split as BID dosing. And what you can see here is
5 that 5 milligrams per day, total daily dose, has a
6 lower hazard ratio for bleeding when the dose is
7 split as 2.5 BID, as compared to when it's given
8 once daily. That is also true for 10 milligrams,
9 given as 5 milligrams BID, with a lower hazard
10 ratio for bleeding than when the dose is given all
11 at once. And when you get up to the high dose,
12 20 milligrams per day, there's not a lot of
13 difference.
14 It was on the basis, then, of this safety
15 data, as well as maintained efficacy from this same
16 trial, that 2.5 milligrams BID and 5 milligrams BID
17 were carried forward into the acute coronary
18 syndrome phase 3 trial, the results of which have
19 not been reported.
20 Then finally, the data that we can bring to
21 bear on the dose selection for ROCKET comes from
22 the ROCKET trial itself. And I agree with the

sponsor's statement that the established close-to-linear PT rivaroxaban plasma concentration relationship in phase 2 supported the use of the prothrombin time exposure-driven safety and efficacy analyses for the phase 3 trials.

Of note, the PK/PD relationship that you're seeing here, or the PT rivaroxaban concentration relationship, is actually -- this is ROCKET data. And so this very tight relationship between rivaroxaban plasma concentration versus the prothrombin time was actually reproduced again within ROCKET.

Based on that relationship, we proceeded forward with looking at outcomes from ROCKET, either ischemic stroke or major bleeding, based on the prothrombin time initially or the INR for warfarin relationship. What you see here -- and we'll start with warfarin -- this is a logistic regression, unadjusted, for the warfarin data in ROCKET, looking at the percent of patients with either ischemic stroke or major bleeding as a function of the last observed INR in the study

before one of those two events. And this solid line is the logistic regression. The shaded lines are the 95th-percent confidence interval, and, really, for informational purposes, we put the medians of the various INR quartiles here in the red diamonds, with standard errors around them.

What you can see here is what you've gotten used to seeing with warfarin, that as INR increases, there is a progressive decrease in the incidence of ischemic stroke, and as INR increases, there's a progressive increase in the incidence of major bleeding, with the balance being what you're used to seeing and what we target clinically with the INR in the range of 2 to 3.

However, what you see with rivaroxaban is quite a different pattern. Again, these are logistic regressions, unadjusted, looking at the percent of these patients having ischemic strokes or major bleeding, this time with the last observed prothrombin time, approximate to the event. And what you see is a flat response here in ischemic stroke, with respect to what's considered to be a

1 surrogate for drug exposure within ROCKET, as far
2 as the prothrombin times. But you see a
3 progressive increase in the incidence of major
4 bleeding as a function of increased prothrombin
5 time.
6 This was a very robust finding, and we got
7 this same pattern of result, no matter whether you
8 looked at prothrombin time, prothrombinase-induced
9 clotting time, factor 10A activity, or the decoded
10 rivaroxaban INRs, which were available to us in the
11 dataset.
12 You also get this same pattern if you change
13 the definition of major bleeding. We got the same
14 result if we used TIMI major bleeding, as we did
15 with the international society definitions of major
16 bleeding, or a modification of the international
17 society definitions, where we required a five-unit
18 blood transfusion to count as a major bleed.
19 The pattern was the same, and it also held
20 up when covariants that were -- covariants to these
21 clinical outcomes were adjusted for by Cox
22 proportional-hazards models regression. Once those

variables were adjusted for, this pattern was still present.

Therefore, Mr. Chairman, this reviewer concludes that the clinical pharmacology attributes suggest BID dosing of rivaroxaban. The VT studies do not provide a firm basis for the dosing regimen that was studied in ROCKET. The 20-milligram per-day dose of rivaroxaban was tested in ROCKET, is twice the approved dose, both in the U.S. and globally, for the VT indication. And lower-in-split dosing was incorporated into the Acute Coronary Syndrome Development program, as 2.5 milligrams BID or 5 milligrams BID.

Then, finally, when one uses pharmacodynamic parameters to assess as a surrogate for exposure, to assess ischemic stroke and major bleeding in ROCKET, for warfarin, you see the expected decrease in ischemic strokes and increase in major bleeding as INR increases. And then for rivaroxaban, there is a flat PT ischemic stroke relationship, but increasing the PT was associated with increasing incidence of major bleeding. And there again, this

finding was true no matter which pharmacodynamic parameter that we looked at. Thank you.

This is the completion of the dose selection talk. I'd like to now introduce Dr. Marty Rose. Dr. Rose was the efficacy clinical reviewer for rivaroxaban and will now go into that review for you.

FDA Presentation - Martin Rose

DR. ROSE: Thank you, Dr. Dunnmon.

Mr. Chairman, members of the panel, I'm here to discuss issues affecting the interpretation of efficacy, and there are four areas I'd like to discuss. I'll start with assessing efficacy in non-inferiority and superiority trials. I'll note that Dr. Cahill indicated that superiority was off the table, and I confirmed that with a representative of the sponsor. So I won't be spending much time talking about superiority.

Next, I'll go onto adequacy of warfarin management in ROCKET, which is where I'll be spending most of my time, then post-treatment primary endpoint events in ROCKET and J-ROCKET, and

finally finish up with the approval standard in 2011 for a drug to prevent stroke in patients with AF.

A couple of caveats and statements. First, I'd like to remind you that the opinions that I will be expressing are my own and do not necessarily reflect those of my management. Second, like everyone else here, I dislike or don't trust cross-study comparisons, and I will not be comparing data from one study to another. However, I will be comparing studies in terms of how they were conducted. At the very end, I will allude to the results of RE-LY in talking about standards.

So in terms of non-inferiority and superiority, I'm going to skip over all of the superiority language and remind you that the findings for non-inferiority were supported by multiple supported analyses. However, they were analyses of the entire study population, and the quality of warfarin management was not considered in any of these analyses.

In terms of ITT analyses, I'd also like to

suggest that ITT refers to a patient population. It does not tell what the observation -- it does not tell you what the observation period is or was, and it does not tell you the event window, as some people call it. And so those are two separate concepts, and you have to keep that in mind in considering any ITT analysis. It's been stated, and we agree, that most of the studies in this area have collected events until a time corresponding to site notification in this study, and we agree with that.

So let's go onto the adequacy of warfarin management in ROCKET. I'll start with this graphic. You've seen it before. This is from the latest update of the guidelines for the management of atrial fibrillation. And it shows that if you keep the patient within the range of 2 to 3, you can minimize both ischemic stroke, which is represented on the solid line here, which is pretty flat all the way across here and then turns sharply up at a level of 2, as well as intracranial hemorrhage, a component of which is hemorrhagic

stroke, which again is pretty flat in this region and then curves up, but its slope is much less steep. And that's something that you should keep in mind and will come into play in a later slide.

This is a slide that Dr. Dunnmon showed you, and it shows the relationship between INR and the ischemic stroke rate in warfarin. And it's quite similar to what you saw in the previous slide. As INR gets closer to the target of 2 to 3, the ischemic stroke rate drops. So this is data within ROCKET.

So briefly, management of INR in ROCKET -- you've heard some of this before. The protocol included a procedure for starting study drug in patients who entered the study on warfarin or some other vitamin K antagonist. However, there was no algorithm for managing VKA dose. However, there were guidelines for bringing the patient back in for INR more frequently at the start of the study than later. Other studies have provided investigators with an algorithm. These include RE-LY and ARISTOTLE. It may not have been

mandatory to use it, but at least they had the paper in front of them at the site. I'll discuss transition to VKA at the end of the study later.

So this slide shows the percent time in various INR ranges in the warfarin arm. The yellow row corresponds to the target range of 2 to 3, and the mean is 55.2. And that is the mean TTR, which we've talked about. And as Dr. Cahill indicated, it is used as a quality control measure and there are many studies that show that it has a relationship with event rates in the warfarin arm. And I will show you some of them in a few minutes.

You'll also note that there was more time spent below range than above range, and, in fact, those numbers are 30 and 15 percent. If patients were in range 55 percent of the time, they were out of range 45 percent of the time.

These are our two studies that underline that first graphic that I showed for the relationship between ischemic stroke and INR. The one on the left is from a Hylek's paper in the New England Journal in 1996, and I can't see that slide

very well, but you can.

So the yellow rows show odds ratio for ischemic stroke versus patients at an INR of 2. The yellow rows show the INRs of 1.9 and 1.8. Both of those have odds ratios that are modestly elevated above 1. The results appear to be statistically significant.

The chart on the right are more recent data from the ATRIA study. The yellow row shows the combined range of INR to 1.8 and 1.9. You'll notice that the odds ratio is in line with Hylek's data. And I emphasize this because some have suggested that one can include that stippled area when considering time in therapeutic range, which adds about 15 percent to it. And I think these data show that that would not be appropriate in terms of INRs below range.

For INRs above range, there's not a lot of stroke risk, between 3 and 3.2, so there's about 5 percent there. But if we added that in here, then you probably should add it in for every other study that's been performed in this area. And I don't

1 have the data to do that, so I won't show you those
2 numbers. So I think it would be wise to stick to 2
3 to 3 as the target range for INR, consistent with
4 the consensus guidelines.
5 So this shows INR in recent studies, and by
6 recent, I mean published since 2004. ROCKET is at
7 the top. They're arrayed in alphabetical order and
8 they range from 62 to 73. ARISTOTLE, EMBRACE AC,
9 and RE-LY -- excuse me. ARISTOTLE, EMBRACE AC,
10 and SPORTIF V were all double-blind studies. The
11 rest, other than ROCKET, were open label.
12 You can see that it doesn't make a lot of
13 difference. And I apologize to the committee for
14 failing to note that RE-LY was a blinded study. I
15 thought I had done that, but it doesn't seem to
16 make a lot of difference. And the largest of all
17 of the high TTR studies was SPORTIF V at
18 68 percent.
19 So from here on, I'm going to show a number
20 of slides about site mean TTR, which we defined
21 simply as the mean of the TTRs for all of the
22 individuals, all the patients who are in the

warfarin arm. And I'll show you that later, but it's very simple.

So there's a lot going on in this slide. The rows represent two fairly recent studies. ACTIVE W was a warfarin-controlled study. The experimental regimen there was aspirin plus clopidogrel. RE-LY was a warfarin-controlled study. The experimental agent there was dabigatran in two doses.

The quartiles are quartiles of site mean TTR, basically constructed by lining up all the sites in order of their site mean TTR and then dividing them into four, not always the same way -- in ACTIVE W, which was the earlier study done of these two, and is the one that most people cite for methodology for site mean TTR. In ACTIVE W, the sites or the quartiles were divided into an equal number of sites. So the number of patients was a little bit different. In RE-LY, in the publication, the quartiles were divided into equal number of patients. And I emphasize that because we did it both ways, and it changes the

results in ROCKET. It doesn't change it much in other studies, but it changes the results a lot in ROCKET. And these are the annualized event rates.

So to the left of the dark vertical line, you see the annualized event rates in the warfarin arm, and you'll see that from Q1 to Q4, there's a consistent drop here. Better warfarin control is associated with a lower event rate in the warfarin arm, and the results are quite similar, obviously, in the two arms.

To the right of the line, you see the hazard ratio for the experimental treatment versus warfarin. For ACTIVE W, it's greater than 1 because warfarin is a better anticoagulant than aspirin plus clopidogrel. For RE-LY, it's less than 1. And, again, I'm not comparing any of these to ROCKET, but I'm trying to show a pattern.

You'll notice that between the first quartile and the fourth quartile, the hazard ratio goes up in both. And it goes up by about 70 percent in each one, which means that, comparatively, warfarin looks better than the

comparator in the fourth quartile than in the first quartile.

There's not a lot of studies that have looked at this, but this is what you would expect. The experimental arm may also be affected. However, the warfarin arm is affected much more by whatever is happening between the first and the fourth quartile. If it was simply better care in quartile 4 sites than in quartile 1 sites, you would expect the two arms to be affected the same way because of the nature of center-based TTR analyses. They preserve the randomization. When you put those patients into that first quartile or that fourth quartile, based on similarities in their TTRs, you take the experimental arm patients with them. So you're comparing like to like in each quartile.

So, again, if it was something else, if it was quality of care, if it was better hypertension treatment, you would expect to see that affect both quartiles, but it doesn't. It affects -- whatever is going on affects the warfarin arm much more.

1 And so, probably, the best explanation for what's
2 going on here is better control of warfarin, more
3 time in the therapeutic range.
4 So let's talk about calculating site TTR.
5 You'll see two ways of doing it. Connolly, in that
6 paper from ACTIVE W that I referenced just a few
7 minutes ago, did it and described it as the average
8 of TTR values for individual patients at the site.
9 I have not seen a single published paper that uses
10 another method for calculating site TTR. I expect
11 there'll be one soon, but it's in no paper that I
12 know of.
13 This is what we used for our analyses. The
14 sponsor did something else, and I've described it
15 up there, but it's easier to understand in the
16 hypothetical example I'm going to show you.
17 The important thing to know about it is that it
18 gives more weight to patients who remain on study
19 drug longer. It gives less weight to those who
20 leave the study early. So a patient who goes in
21 and gets bad treatment, gets bad warfarin
22 treatment, and leaves the study in three weeks for

a stroke or maybe for a bleed is hardly counted at all, basically disappears from the analysis.

So that may or may not be a good thing, but the upshot of this is that the results of the sponsor's method can't be compared to anything else you've seen because no one else that we know of has done it that way.

So here is TTR data for a hypothetical site with four patients in it. And each patient is represented by a row. And patient number 1 was treated for 1,080 days. That's 36 months, and there were patients in ROCKET who were treated for 36 months. That patient had 756 days in the therapeutic range. Some of those represent actual values, and probably most of them would be imputed values.

So we calculate that patient's individual TTR by dividing 756 by 1080, and we get exactly 70 percent. The other three patients, patients 2, 3, and 4, have similar data. Each one in the study was treated for 120 days and had 60 days in range, and, of course, that yields a TTR of 50 percent for

each of those patients.

So here is the same slide I showed, plus a couple of other rows. FDA and Connolly would calculate TTR simply by averaging 70, 50, 50, and 50, and we would get 55 percent for the TTR. The sponsor would add up the days treated, get 1440, add up all the days in range, get 936, divide 936 by 1440, and get 65 percent. So, obviously, this method weights this patient much more than these three patients.

As has been said, one would expect patients who are doing better on warfarin to have higher INRs, higher TTRs, and be in the study longer. So this would tend to inflate site TTR a little bit, and we see that, in fact, in ROCKET, where the cutoffs for the quartiles are all 2 to 3 percent higher using the sponsor's method than using the method of Connolly.

So let's move on and look at a number of TTR analyses. We analyzed TTR in several ways. We looked at it by quartiles, either containing similar numbers of patients or similar numbers of

sites. We used various TTR cut points, and we also looked at the hazard ratio as a continuous function of TTR.

We also looked at several event windows. LD plus two days, or last dose plus two days, is the on-treatment analysis, randomization to the last dose plus two days after VKA was discontinued. And we also looked at last dose plus 30 days. All the analysis you'll see here are in the safety population. The last dose plus 30 day analysis includes a lot of post-treatment events and is in between the on-treatment analysis and the site notification analysis.

So let's start at the top. This is a quartile analysis, and I'll be showing either the best quartile here, which is the fourth quartile, or a cutoff at a fairly high level for TTR. And you'll see that as I go along.

So this is the best quartile in the safety population, last dose plus two day analysis. The sponsor's TTR cutoff here was around 66. We're around 64. But we and the sponsor got very similar

results, a hazard ratio of about .75 and event rate less than 2, 1.75.

But that's not the first analysis we did. The first quartile analysis we did after we saw the sponsor's results was this one, which is, again, the safety population, last dose plus two days, or on treatment. But this time, the quartiles were constructed with similar numbers of sites. And if you look out to the right, you see a hazard ratio that crosses 1 and a very wide confidence interval. And we thought, what's going on here? We looked back and there's what's going on. You have far fewer patients. When you construct the quartiles this way, in this study, you get a much smaller fourth quartile, and the TTR cutoff here is 67.8. So something seems to be happening between a TTR of around 64 and a TTR of around 67.

We did this analysis as well. This is 72.6. It's the TTR cutoff for the fourth quartile in the published paper describing the quartile analysis of RE-LY. But, again, these are not RE-LY data; it's just the cutoff that they used. And you'll see

first that there are few patients left, 623, in the warfarin arm. These are all in the warfarin arm. And there were only 11 events. The event rate is down to 1.05, and the hazard ratio is up to 1.17, and a very wide confidence interval that spans a fourfold range of values. So this is quite different from the other analyses.

The last two rows show last dose plus 30 day analyses. This population corresponds to this one, and you'll notice that the event rate goes up and so does the hazard ratio. And this analysis corresponds to this one. And, again, the event rate is higher, and the hazard ratio is 1.30.

So this suggests that if you look at high levels of TTR, the evidence of superiority certainly goes away. And you begin to wonder whether this drug is as effective as warfarin when it's used well. Certainly, the hazard ratios are quite high.

So this is a continuous function analysis. And I'll describe how we did it. I think it's really not too different from the way it was done

at J&J or Duke. And we're starting -- first, we calculated the site mean TTR for each site. We laid them out from low to high, and started at the left-hand by including all of the sites. And when you do that, you get the hazard ratio for the whole study, which is .79. Then we took off one site at a time from this end, moving across, again, with fewer and fewer patients remaining in the analysis.

You see it's fairly level across here, and then the hazard rate turns up at around a TTR level of 64. You'll also notice that there's this row across the bottom, which is the percent of subjects excluded. And when you get up into the 70s, you start excluding more than 90 percent of the patients. For that analysis with a TTR cutoff of 72.6, it's around 92 percent of the patients excluded. And there's very few events.

When you go above 72 or so, you'll notice the hazard ratio drops down. By the way, this is the hazard ratio. These are the 95-percent confidence intervals. The hazard ratio drops down below 1 and then goes back up. But now you're in a

range where you can count the number of events on your fingers and toes. And, in fact, at the very end, you can count them on one hand.

So you're not going to get a very accurate representation of the hazard ratio with that number of events, and you would expect it to bounce around because a change of one or two events, plus or minus in either arm, could change the hazard ratio markedly. So, again, there's not a lot of patients here. There's not a lot of events in this region. And you have to wonder whether this drug is as good as warfarin when it's used well.

So give me a second while I catch up on my notes with where we are in the slides.

So I've talked about using warfarin well or using it skillfully, and if you've read the review, you'll see that language in there. That's a hard concept to define. Reasonable people could differ. But no matter how you define it, it seems pretty clear that, at the levels of TTR attained in such places as England, where the overall country average was 69, in Sweden, in New Zealand, in

Australia, all of which were around 70 or even higher, or Duke University, where the TTR was 69.6, there's not a lot of data in this study. There were 72 U.S. sites that had a site TTR greater than 70; 72 U.S. sites. Didn't enroll a lot of patients, but they did very well. I looked closely at 15 of them; 14 of them were community cardiology sites with no apparent academic connection whatsoever, so people in the community are getting TTRs of 70.

The little clinical perspective that followed Connolly's paper on TTR in ACTIVE W suggested that practitioners should be aiming for a TTR of at least 70 because of the benefit for their patients. So it's not unattainable. It's certainly not easy to get, but it ought to be a goal. It's an aspirational thing. It ought to be your goal.

So you've seen data suggesting that hazard ratios looked pretty good in the top quartile in ROCKET, so let's look at that a little bit. This is our top quartile. Again, the hazard ratios in

our quartile and the sponsor's quartiles are very similar. The cut points are a little different. Ours are a little lower. It's not a huge difference.

So these are the upper limits of the bottom three quartiles, 46.8, 55.9, and 63.9. And you'll noticed that, again, comparing to the conduct of other studies, not the results, that the first quartile for RE-LY had a higher cutoff than the second quartile for ROCKET. And the second quartile for RE-LY had a higher cutoff than the third quartile for ROCKET.

So that fourth quartile for ROCKET, which was everyone with a TTR above 63.9, included a lot of patients who might have been in the third quartile, in one of these other studies. ACTIVE W had very similar cutoffs to RE-LY. So that fourth ROCKET quartile has a lot of patients that would have been in the third quartile in either of these two studies, which may account for the fact that you don't see much difference in the hazard ratios there.

Finally, here is RE-LY, and, again, there is no efficacy data here. We're talking about study conduct. This is from Wallentin's paper on TTR versus results in RE-LY. And, again, these quartiles were divided by equal numbers of subjects. RE-LY was an 18,000-patient study, so each quartile had one-fourth of 18,000 patients, or about 4500. And in RE-LY, which, by the way, had three arms, not two arms, they enrolled 4500 patients at a TTR greater than 72.6. ROCKET had around 1200 patients at that TTR. They had 25 percent of the population. ROCKET had about 8 percent of the population at really good levels.

So the sponsor suggests that this is geography, and, indeed, it is. It is geography. So let's take a look at this. This shows the percentage of patients from each of the five regions. Let's start in order of regional TTR, which, again, is just the mean of all of the individual TTRs for all of the individual patients in those regions, in the warfarin arm.

So in North America, 18.8 percent of

patients, Europe, 14.8, you add it up, and you get around 34 percent of the regions with TTR better than 60. And, traditionally, these are the regions that do well in terms of warfarin control and in RE-LY. So this was 34 percent. In RE-LY, it was 62 percent of the people that came from regions where the TTR was good. Eastern Europe, in this study, had the worst TTR, 49.7, but it enrolled more than twice as many patients as any other region. It had about 39 percent. In RE-LY, it was 12 percent.

So when the sponsor says that the results of this study are related to geography and cultural patient factors, they're absolutely right. We're not quibbling with that at all. When you look at it in an operational sense, what these data mean is that the ROCKET study enrolled the majority of its patients from regions where physicians don't use warfarin very well. They did a warfarin-controlled study, compared themselves to warfarin in a study where most of the patients were enrolled at sites where warfarin is not used well. And that's our

1 major concern with this study.
2 This is an interesting analysis, and it goes
3 back to that curve that I showed you. And I'll be
4 very fast. This is a plot of time below range, INR
5 less than 2. And what you see -- and I'll just
6 describe what I like about this -- the range here
7 in terms of events per 100 patient-years is wider
8 than in the TTR analysis. This now is the best
9 quartile because it's the least time below range.
10 And the warfarin event rate and the hazard ratios
11 show a wider range, from 1 to 4. And this may be a
12 better way of discerning warfarin control than
13 looking at TTR.
14 A couple words on observed event rate as a
15 measure of adequacy of warfarin administration. So
16 the protocol for ROCKET specified an event rate of
17 2.3 in the warfarin arm. And they did meet that.
18 How reassuring is that? They met their warfarin
19 target, but we know that warfarin event rates can
20 vary greatly. SPORTIF III and V were run under
21 basically an identical protocol. One, of course,
22 was blinded and one was not. And SPORTIF III had

an event rate of 2.3 percent, and the blinded study had an event rate of 1.2 percent, despite identical eligibility criteria. They were run in different parts of the world, but this suggests that you just don't know what kind of event rates you're going to get out of a study.

The protocol for ROCKET specified an event rate of 10 percent per year. That was their target for the primary safety endpoint. And they were about 50 percent above that. And even though TTR was low, you just don't know what you're going to get.

So, in summary, increasing levels of site TTR are associated with a reduced rate of thrombotic events in patients treated with warfarin with atrial fibrillation. Increasing levels of site TTR are associated with changes in the hazard ratio for thrombotic events that favor warfarin in warfarin-controlled trials of novel anticoagulants. And site mean TTR is a useful tool to understand the impact of INR in warfarin-controlled studies.

INR controlled in ROCKET was worse than in

other recent trials, which might have biased the overall results in favor of rivaroxaban. Very few patients were enrolled at sites where warfarin was used skillfully, resulting in insufficient data to assess the efficacy of rivaroxaban to warfarin when the latter is used well.

Let's go to post-treatment events. You've seen a lot of this, so I won't show it to you again. You've seen these data. The J-ROCKET was a much smaller study, and we haven't talked about it much. It was run exclusively in Japan. There were no Japanese patients in ROCKET. This study used a lower dose of warfarin and a different INR target, and also had the same phenomenon in terms of post-treatment events, and completers was seen in this study as well. I'll scroll through these very quickly.

I'd like to make some comments on the sponsor's proposed transition regimen. Starting rivaroxaban in patients taking warfarin is not comparable to starting warfarin in patients taking rivaroxaban. Rivaroxaban, as you saw, may increase

INR. The top quartile was somewhat above 1. Discontinuing rivaroxaban in patients who reach an INR of 2, they'll lose their rivaroxaban support, and they may drop under 2. That could be a problem. They could be under-anticoagulated. We all know that the onset of warfarin is slow, that time to therapeutic range is unpredictable, and that INR can fluctuate in and out of the therapeutic range during the initiation of therapy, which, again, makes just reaching 2 problematic.

Also, concentrations of protein C, which is antithrombotic, decrease faster than procoagulant factors when you start warfarin. So stopping rivaroxaban too early may result in higher risk of pathologic thrombosis. So, again, I think there are some issues with the regimen that the sponsor has proposed.

So, in summary, completing patients in the rivaroxaban arm of ROCKET had a 28-day post-treatment stroke rate that was 3.7 times that of warfarin-arm patients. The J-ROCKET data were consistent. The rivaroxaban arm stroke rate was

not inconsistent with the observed pro-warfarin control. And we will grant that it was not good. However, the data are also not inconsistent with rebound hypercoagulability. And the sponsor has not done much to rule this out. And I think that a clinical study may be needed to evaluate rates of bleeding in thrombotic events using the sponsor's proposed strategy for transitioning patients from rivaroxaban to warfarin, or perhaps a different strategy. This would not necessarily have to be a study counting events, but perhaps a pharmacodynamic study. And this should be done prior to approval.

Finally, moving onto drug standards, in 1995, FDA issued a federal register notice, that was August 1st, 1995, that discusses the agency's standards for demonstrating effectiveness of human drug products, as well as devices. This was an outgrowth of the Clinton-Gore Reinventing Government initiative, which Dr. Temple alluded to earlier.

The policy indicates that, most of the time,

1 we don't require comparative studies -- I'm not
2 reading from the quote -- and that a placebo-
3 controlled study is good enough. However, it goes
4 on to state -- and I will read this -- "in certain
5 circumstances, however, it may be important to
6 consider whether a new product is less effective
7 than available alternative therapies, when less-
8 effectiveness could present a danger to the patient
9 or to the public. For example, it is essential for
10 public health protection that a new therapy be as
11 effective as alternatives that are already approved
12 for marketing when, one, the disease to be treated
13 is life-threatening or capable of causing
14 irreversible morbidity, e.g. stroke or heart
15 attack," and then it goes on to name a few others.
16 So stroke is a classic situation of a drug
17 covered by this policy. It is essential that the
18 drug be as effective as what's on the market. So
19 we've had occasion to apply this policy in
20 situations involving drugs for stroke prevention in
21 afib patients. A dose of dabigatran that was
22 clearly non-inferior to warfarin was not approved

because it was significant inferior, in the same study, to a higher dose of dabigatran in terms of stroke rate. That was RE-LY, of course.

We also have stated that an NDA for a drug to prevent stroke in atrial fibrillation patients that was based solely on a trial in which the active comparator was warfarin would not be approvable. Our rationale for that was that, while a finding of superiority to aspirin clearly indicates that a drug is effective -- i.e. if it's effective, it's better than aspirin, it's got to be better than placebo -- however, warfarin is so much more effective than aspirin as an anticoagulant that aspirin is not an appropriate comparator.

The Federal Register notice, which is in your packages, was a broad policy document, and like most policy documents, and it didn't include a lot of operational detail. There are a number of unanswered questions. How much of the comparator's efficacy has to be maintained? How much uncertainty about the efficacy results is acceptable? What if another treatment for the

indication is approved during the course of development?

So, again, this is my opinion only. I think that the current landscape of anticoagulant therapy for atrial fibrillation patients is an important factor in answering these questions. In 2006, when ROCKET was planned, there was no alternative to warfarin therapy, but dabigatran was approved last year. So there is now an alternative to warfarin without warfarin's troubling interactions and without the need for monitoring. RE-LY demonstrated that dabigatran was robustly non-inferior to warfarin in patients at sites with TTR above a study median of 67 percent.

I think it's reasonable that rivaroxaban should have to meet the same standard, i.e., robust non-inferiority to well-managed warfarin. A lot of our patients are well-managed on warfarin. If they switch to another drug, you should be comfortable that it as good as warfarin when warfarin is used well. The alternative could be that the patient would be less well-anticoagulated and might have a

stroke. Rivaroxaban does not meet this test.

I'd like to go onto my backup slides here because some issues have come up, so let me scroll through. I don't have far to go.

The U.S. data. I've got three slides of data from the U.S. And the last one is a little hard to read, but it is on page 132 in a very easy-to-read form in my review. So if you want to look at it, but we've got a few slides to go before we get there, and I think these are pretty reasonable.

So this is various data scopes, various populations, and endpoint cuts in the U.S. only for the ROCKET study population in the U.S. The thing I'll point out is that in the on-treatment population, rivaroxaban looks very good. The hazard ratio here you'll see on the next slide is .54. There are only 15 events. But as you go out -- this is the on-treatment analysis, last dose plus two days -- five days later, you're up to 20 events; 14 days later, 21; 30 days, 24. And then, if you go all the way to the end of the study, you're at 36 events, and the hazard ratios move

much closer together.

So let's look at some more. So this is global data versus U.S. data, and I just want to show you some patterns. Again, there's a lot going on, on this slide. I'm sure you're quite familiar with the various populations and the observation periods.

At the top, we've got the rivaroxaban event rate and the warfarin event rate in the white rows, the hazard ratio out here. These blue columns are a little bit different. The blue column, say, for example, for ITT notification, is the ITT to site notification event rate divided by the on-treatment event rate. In each case, it's the event rate divided by the on-treatment event rate. So it starts at 1 at the top.

So let's look at warfarin first. What you see as you go down from here to here is that event rate changes modestly, that this ratio in the blue column starts at 1 and gets as high as 1.12. So it increases by 12 percent. The hazard ratio goes up modestly. If you look at rivaroxaban globally, you

see that there's more change as you move through time. And this is quite consistent with an increase in post-treatment events, both in patients who discontinue early, where rivaroxaban was a little bit worse than warfarin, and those who complete and finish the trial, where rivaroxaban was significantly worse than warfarin.

But, at worse -- so here it's 1.12 for the ratio at the ITT, regardless of treatment exposure, and here it's 1.29, so a 29 percent increase in the hazard ratio. And look what happens in the U.S. So the warfarin arm looks very similar to the global results. And these global data include the U.S., and this is the U.S. versus the U.S. plus the rest of the world.

So this ratio goes from 1.00 to 1.19 for warfarin, not too different from 1.12. But look what happens over here. It goes from 1.00. Thirty days later, it's up to 1.55. And for the ITT, regardless, it's 1.93.

So what's going on here? I don't know, but I can speculate. There are a lot of patients who

seem to leave the study and have events soon afterward. It certainly looks like informative censoring. Why would it be different in the rivaroxaban group from the warfarin group? Well, probably, one of the first things that you do when a patient gets into trouble is get an INR. And if the patient's not in your office, you have to get an unblinded INR. All this is speculation. You have to get an unblinded INR. And that unblinded INR, if it's 3, would tell you, probably, this is a patient in the warfarin arm. This is a drug that I know how to use. I've used it for years. Everyone I know has used it for years. I can keep the patient in the study.

If the INR is 1.1 or 1.2, then it's probably a patient in the rivaroxaban arm. I've got a patient in trouble on an experimental drug. What should I do? I'm going to take them out of the study and put them on a drug that I know how to manage.

So maybe that's what's going on here. Maybe this is informative censoring that we're seeing,

1 but at the end of the day, it's hard to trust that
2 on-treatment hazard ratio of .54 and those on-
3 treatment events. And I think you have to look at
4 the ITT data in this study, especially in the U.S.,
5 where the on-treatment analysis is suspect.
6 One more issue that I'll go to. There has
7 been discussion of adjustment --
8 DR. LINCOFF: We're going to have to finish
9 this in about five minutes because of the time
10 constraints, please.
11 DR. ROSE: -- okay -- adjustments of the TTR
12 for various demographic and other factors. And we
13 did that, and that's in section 6.7.2.2, starting
14 on page 132 of the review in table 41. And we did
15 it three ways.
16 I've left something out. Another thing that
17 you need to know is that the dropout rate in the
18 U.S. was very high. Overall, 35 percent of
19 patients in the study dropped out. In the U.S., it
20 was over 40 percent. And another 8 to 9 percent
21 were lost to follow-up in the U.S. So that's
22 consistent with the possibility that there was

informative censoring.

So, going back to the adjustment, we adjusted it three ways, and if you'd like, you can read about it. We did logistic regression. We adjusted for the various factors described in Rose's -- now, no relation to me -- Rose's paper. And we also adjusted to match the population to the RE-LY population based on data from within ROCKET, what would happen if the patients were like RE-LY.

For the first adjustment, we got less than 1 percent. For the second adjustment, we got about 2 percent. We also did a logistic regression, again, based on RE-LY, and got a little bit over 2 percent. So those two matched. The first one was algebraic. I did it. The second one was logistic regression. You can read the details in the review. But the bottom line is that when you adjust the TTR for all these factors, region, sex, everything else, it doesn't change it by more than about 2 and a half points.

So that's it. Thank you.

DR. LINCOFF: We are going to break for

lunch now and come back with the questions for the sponsors. I'd like to shorten the lunch, if we can, a bit, so instead of the full hour, if we could, convene at 1:15, which would be 50 minutes, to catch up some of the time, that would be great. So thank you very much.

(Whereupon, at 12:26 p.m., a lunch recess was taken.)

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

(12:26 p.m.)

Clarifying Questions for FDA Presenters

DR. LINCOFF: We are going to resume now. We're going to start with, now, questions, clarification questions for the FDA, and we'll begin with Dr. Kaul.

DR. KAUL: Thank you, Mike. I have three questions and I'd like to be able to ask all three questions, so the answers should be brief. I was very intrigued by the U.S. subgroup data, and you showed data from the primary efficacy endpoint.

Two questions related to that. I would have liked to have seen the data for just the ischemic endpoints, because I want to see what the upper bound of the 95 percent confidence interval is. With the primary efficacy endpoint, the point estimate is .86, and the upper bound is closer to about 1.33, which approaches the non-inferiority margin. But the non-inferiority margin was based on trials that evaluated ischemic events.

So I want to see if you have done that

analysis primarily for ischemic events, number one. And number two question related to the U.S. data is, the primary efficacy data, you said there were a lot of dropouts in the U.S. subgroup.

How many excess events would need to occur in the rivaroxaban arm to negate non-inferiority? So have you done that type of sensitivity analysis in that subgroup?

DR. ROSE: So your first question, could you be a little more specific about which analysis you're talking about?

DR. KAUL: The U.S. data, the analysis that you showed in your backup slides, 36 versus 42 primary efficacy.

DR. ROSE: So the ITT?

DR. KAUL: The ITT, all --

DR. ROSE: Right. So offhand, I don't recall seeing the ischemic stroke data for that analysis. Those tables were all taken right out of the sponsor's submission. And they may know whether those ischemic stroke data exists for that analysis. I don't remember them.

DR. KAUL: What about the impact of the dropouts?

DR. ROSE: We did a calculation for the number of events that it would take to overturn efficacy in the global analysis. We didn't do that for the U.S.

DR. KAUL: Okay.

DR. PETERS: So we do have a Kaplan-Meier of discontinuations in the U.S., if we could show that first, and then the primary endpoint in the U.S. So slide on. I'm trying to go very quickly. The hazard ratio was 1.06. And you can see the curves are pretty close to each other over the course of the study.

Actually, if we could show the table. This is the primary efficacy endpoint in the U.S. across all of the analyses. So you can see that it actually meets the non-inferiority margin for all of the analyses. And then the next slide on. This is up to site notification, but if we could show safety on treatment. So the ischemic stroke here up to site notification is actually .89. You're

getting to small numbers of events, so the upper bound of the confidence interval is 1.59.

This is the on-treatment data to be comparable to what we presented earlier, and there, it's 8 versus 16 for ischemic stroke in the U.S., so a hazard ratio of .52.

DR. KAUL: The question number two for the FDA is, in your last part of your talk, you focused on the term "as effective as," which is the FDA standard. Can you clarify what you mean by as effective as? Are you trying to suggest equivalence, similarity?

DR. ROSE: That's a difficult question, and, of course, it's not defined in that very broad policy document. And I have never seen a written document that defines how to think about it. It's one of the questions we're asking you.

The way I think about it would be, it was summed up in probably the last sentence, last two sentences that I said, which is robust -- in this case, it would be robust non-inferiority to warfarin when it's used well. And robust non-

inferiority is a vague term.

Let me just say that for dabigatran, in that analysis, the upper limit of the confidence interval was 1.05, and the hazard ratio was less than 1. And so that was in the analysis for sites above the median, which was 67 percent TTR. So the upper limit of the confidence interval in the comparison at those sites was 1.05.

So I would personally consider that robust non-inferiority, and that would be the standard I would expect in a population where warfarin was used well.

DR. KAUL: From the philosophical underpinnings of the non-inferiority design, as best as I understand non-inferiority, it's that you don't design non-inferiority to demonstrate that the new therapy is as effective as the standard treatment. You design it to rule out an unacceptable degree of harm or inferiority.

So how can we take the same construct and arrive at as effective as, when it was never designed to address that? So can you routinely use

a non-inferiority platform to arrive at that standard?

DR. ROSE: Possibly, yes, if you choose your margins carefully and choose your patient population carefully. At least, that's what I'm proposing. This is a difficult question. I mean, that policy didn't come with an instruction set. And you've raised issues that have befuddled us. We've certainly thought about these things, and they're hard to grapple with.

Remember that the policy is to protect public health, to prevent people from having strokes, or other things that can kill them, or leave them irreversibly maimed in some way. And I think behind it, if I may be so bold, the goal is to protect people from the possibility of taking an inferior therapy which won't do the job so well.

I think you have to approach it from that angle. And maybe that calls for a non-inferiority margin that's considerably smaller than what we usually deal with. Maybe there's some other way of doing it. I haven't thought so. Dr. Temple

certainly has thought about these things, and I know Dr. Stockbridge has as well, and they may want to add something to this.

DR. KAUL: Before Dr. Temple answers that, I want to get my last question in. You were concerned that about 40 percent, nearly 40 percent, of the enrollees were enrolled from eastern Europe, and the TTR was under 50 percent there, the mean TTR. But if you evaluate the impact on the outcomes of that TTR, the event rate in the warfarin arm is 3.3 percent, which is actually lower than the 3.69 percent you see in Latin America or the 4.3 percent you see in Asia Pacific.

So, ultimately, we are more concerned about the impact of this less skillful TTR on outcomes, rather than the degree of TTR.

DR. ROSE: So I would remind you that this whole study had about 400 events. So eastern Europe, we've already had around 39 percent of the patients. I forget how many were in Latin America or Asia Pacific. They were less than the U.S. What, 10, 12 percent? Something like that.

1 So their share of those 400 events would be 40
2 events, 20 per arm. That's not a lot of events.
3 You're going to see a lot of swings around the
4 hazard ratio when you're talking about small
5 numbers of events.
6 So you can't expect everything to line up.
7 You can't expect the quartiles to line up when you
8 do a quartile analysis for overall, because you've
9 got eight cells and you're sharing 400 events among
10 eight cells, so that's 50 per arm. That's more
11 than what you've got in Latin America, and they
12 don't line up, but they do show a consistent
13 pattern from first quarter to fourth quarter.
14 DR. KAUL: I'm not talking about hazard
15 ratios. I'm just talking about the event rate.
16 DR. ROSE: Event rates drive hazard ratios.
17 If you're talking about 20 events from this
18 subgroup, and that subgroup, and another subgroup,
19 I don't think we can expect everything to line up
20 when the number of events is that small.
21 DR. TEMPLE: I think the difficulty of the
22 situation has been well described by everybody.

The whole concept of -- I mean, you use non-inferiority studies where it's unethical to use placebo. But more than that, where there's a therapy that really works, you really want to know more than that the new therapy is just better than placebo. That's probably not good enough, so you use the non-inferiority design and you set a non-inferiority margin.

One possibility -- we agreed on a non-inferiority margin to be the one to be ruled out, so-called M2, of about 36 percent. It's not out of the question that, being aware that there was another drug that looked actually at least as good or maybe better, we might have set a smaller margin in the future. I'm not saying that. I'm not predicting that. But it's not out of the question.

Some of the irony is, in some ways, that when a control drug is really, really, really effective, you can have a sort of big margin. I don't know whether that's a satisfactory outcome or not, but it's worth thinking about.

But what I was trying to say when I opened

1 up is that the manifestation of that 1995 policy in
2 some ways is reflected in the non-inferiority
3 guidance that we wrote. And it certainly doesn't
4 say you have to show you're better. It also
5 doesn't say you have to rule out all conceivable
6 possible differences. It's trying to make a
7 balance of feasibility, and public protection, and
8 get all those things right.
9 So you want the study to be robust in one
10 important sense. You want to be pretty sure that
11 the active control had its usual effect, the so-
12 called constancy assumption, and that's one of the
13 issues here, was Coumadin as good as usual? I'd
14 want to point out, again, that much of the effect
15 of these drugs is on hemorrhagic stroke, and the
16 concerns about -- most of the advantage of these
17 drugs over Coumadin is over hemorrhagic stroke, and
18 that probably wouldn't be affected by having INR
19 too low, but that probably deserves more attention.
20 But if you look at the history of the
21 margins we've chosen, they plainly represent a
22 compromise between trying to be sure you're not

1 worse and trying to be sure you can still do the
2 study. So with the margin of 36 percent, they
3 still needed 14,000 people. And we went through
4 the same thing on thrombolytic agents. We went to
5 an advisory meeting. They said rule out a loss of
6 25 percent. That would have required a study of
7 55,000 people, so we bought 50 percent.
8 I don't know. But we also feel you should
9 look at the point estimate, too. It's not only the
10 lower bound or the upper bound of the confidence
11 interval. And in some of those cases, the point
12 estimate was actually favorable to the new one. So
13 that was reassuring, even though all you'd really
14 ruled out rigorously was a loss of 50 percent of
15 the effect. And you have similar considerations
16 here. At least for certain analyses, it actually
17 doesn't look only non-inferior, it actually looks
18 better. Well, that could be reassuring that you
19 haven't lost too much if you believe that the
20 control drug was used properly.
21 So, as Marty said, we didn't say exactly how
22 to do this, partly because we didn't know at the

time. We'd really only begun to develop the non-inferiority stuff. And I can't even blame that on anybody else. I think I wrote that part.

So there it is. It's a major problem of practical enforcement. You don't want to lose anything really valuable, but you don't want to make it impossible to get new drugs because you never know what's going to happen to the older drug. We believe in having therapeutic alternatives. And the more I look at the data and how there are these differences between hemorrhagic stroke and thrombotic stroke, I'm not sure I understand a lot of this stuff, and I think there's still work to do. So that's not as direct and absolute a response as I'd like to be able to give.

DR. KAUL: How do you operationalize as effective as within the construct of non-inferiority? I mean, that's a key question.

DR. TEMPLE: That's how we did it. We said you have to rule out a difference more than the certain thing we're going to specify, and that will be as effective as.

DR. KAUL: But that depends on what the ancillary advantages of what the new therapy are. Is it safer? Rivaroxaban lost on its safety endpoint, didn't meet it. Is it convenient? Possibly. Is it less expensive? Probably not. So how much margin do you -- how much effect you want to preserve depends on the ancillary advantages that this drug brings.

If you have an active control that has a 60 percent risk reduction as a treatment effect compared to placebo, are we just happy with preserving just 50 percent? Shouldn't we be thinking about maybe 80 percent? Why just 50 percent?

DR. TEMPLE: That's what I was trying to address. What the choice of M2 is, is very much a judgment call. And there's no point in not acknowledging that it's based partly on the desire to not lose any useful effectiveness and the practical limitations of doing a trial. I mean, we don't think you can realistically do a 100,000-patient trial. So maybe we want to preserve all

1 but 10 percent, and we want to be sure we haven't
2 lost 10 percent. Probably, I do. I don't -- you
3 don't. But if that makes it impossible to ever
4 develop new drugs in this area, we've got to take
5 that into account, too. And we do, frequently with
6 the help of advisory committees like you guys. And
7 I don't think it's out of the question at all that
8 one could change from 50 percent to a smaller
9 value. But, again, you have to think of whether
10 the study becomes doable.
11 I want to remind everybody, the point
12 estimate counts, too. It isn't only the lower
13 bound.
14 Sorry, Tom.
15 DR. TOLIVER: Just a reminder, please
16 identify yourself.
17 DR. ROSE: Marty Rose, FDA.
18 So the large size of these studies stems
19 from the non-inferiority margin that we've
20 selected, but it also stems from the point estimate
21 that we select. And we see studies based on the
22 assumption that the drug is equivalent to warfarin,

that the true hazard ratio is 1.0, and you want to rule out a confidence interval of 1.38. And you get 16 [000] or 15,000 patients.

If you assume your hazard ratio is .9, it's a lot smaller. If you assume it's .8, it's a lot smaller than that. The study doesn't have to assume a hazard ratio of 1.0. And as we're seeing more and more studies, we're getting the impression that, at least in an on-treatment analysis, the appropriate hazard ratio for powering a study may be less than 1.0, considerably less, and that will change the sample size.

DR. LINCOFF: Dr. Emerson?

DR. EMERSON: So, again, I want to sort of try to understand exactly some of the comments you were making regarding the fact that this study's distribution of TTR was different in a non-inferiority study. And if I can -- I think what I heard Dr. Fleming say -- and if he didn't, he should have. But the aspect of, necessarily, the fact that we're here discussing a non-inferiority study does mean that we're comparing cross-studies.

1 So protestations aside, to say, I'm not
2 comparing across studies is irrelevant.
3 So the issues that I can see with this
4 difference in the TTR populations can be, does it
5 change the margin due to the fact that this study
6 isn't comparable to the placebo-controlled studies
7 that presumably led to some comparison of
8 the -- choice of margin, or is this all the second
9 part of the as-effective-as, that we're just trying
10 to say, we want comparability to the dabigatran, so
11 that we can regard whether we have comparability
12 there?
13 So in this non-constancy question that keeps
14 coming up in any non-inferiority study, in this
15 study, we have the questions of, is this a
16 difference in the patient populations but the
17 treatment's the same -- they got warfarin -- or is
18 it a difference in the treatment, that they got
19 inadequate treatment? And therefore, we don't know
20 how to compare it. But either of those non-
21 constancy situations lead to problems in both of
22 those.

DR. ROSE: That's a very interesting question. So the five placebo-controlled studies that were used as the basis of the non-inferiority margin led to that non-inferiority margin of 1.38, based on an meta-analysis and the effect size. And I won't go through all the details of, but at the end of the day, mathematical operation made its windup at 1.38. We took the geometric half of the effect size, and we put basically the square root of the effect size, and we ended up at 1.38. We were trying to preserve half of the effect size.

Those studies, five of them had a very low percentage of patients at high risk. The fifth had a population that was 100 percent composed of people with a prior history of stroke, or TIA. So that was EAFT. The hazard ratio in EAFT was basically the same as the hazard ratio in the other placebo-controlled trials, suggesting that having at least that one risk factor, which is a very big risk factor, prior history of stroke or TIA, doesn't really change the effect of warfarin vis a vis placebo.

1 So that's one question. Those studies had a
2 broad range of INR targets. The INR targets ranged
3 from something like I think 1.4 to 4.5 in those
4 studies. I'm talking about the lowest bottom and
5 the highest top, so a very broad range of INR
6 targets, some of which were quite low and some of
7 which went to 4.5
8 So the ROCKET target falls within that
9 range. The level of attainment of that broad range
10 of INR was also broad, and ROCKET falls within that
11 broad level of attainment. So you're talking about
12 INR range kind of all over the place, a level of
13 attainment that's all over the place. Hard to say
14 you're outside of that range. I'm not sure what it
15 all means, but it's hard to say you're outside of
16 that range.
17 So in that respect, I think ROCKET met the
18 constancy assumption with respect to those
19 parameters. Recent studies, like ATRIA, also
20 suggest that more patients at higher risk for
21 stroke get a larger benefit from warfarin than
22 patients at lower risk. Now, that's not a

1 controlled study. ATRIA is an observational study,
2 but there's a lot of observations in ATRIA.
3 So that, too, suggests that this
4 study -- it's okay to use the original 1.38 for
5 this study. There may be some other issues with
6 respect to, perhaps, a lowering of the efficacy of
7 warfarin, just because we're doing so many other
8 good things for patients, that adding warfarin
9 doesn't have quite as much of an effect as it used
10 to have when we weren't controlling, say,
11 hypertension so well. Now, we control hypertension
12 really well.
13 But the problem, to me relates not to the
14 previous studies, but to what people do now, what's
15 been done now with respect to getting patients to a
16 good INR. It has nothing to do with the placebo-
17 controlled studies. It has to do with the fact
18 that, if you work at it, a site can get to an INR,
19 a TTR, of 70 if they work at it. And whole
20 countries get to a TTR of 70, and they do that in
21 study after study, not a lot of studies, but a few
22 studies.

So that's where we ought to be. That's what we ought to be striving for. So it's not really based on BAATAF, AFASAK (ph), and all those studies from 20 years ago. It's based on what's been happening in the last 10 years.

Does that answer your question?

DR. EMERSON: So just to clarify, this goes back to the very first statements that you all made, saying that you're really willing to stipulate, relative to placebo, that you have high confidence that if we did a placebo-controlled study, that we'd match that efficacy, so that everything about this TTR is really hinging on your point estimate, which I would agree with. It's this concept of looking at it relative to what else could be done with dabigatran, rather than what's the true non-inferiority just against placebo.

DR. ROSE: That's what I believe, and I think that's what Dr. Temple and Dr. Stockbridge believe. I personally think that this drug would beat placebo in a placebo-controlled trial. But you should ask them what they think.

DR. TEMPLE: You can't avoid the fact that dabi exists. There it is. It did very well. And I think it increases the insistence we have that the findings here seem robustly persuasive as to non-inferiority.

So as Marty says, everybody agrees that this drug is better than nothing. But how sure are you, from these data, that it's reasonably similar to the control? And I'm very leery of cross-study comparisons. They have different populations and all that stuff. But it does, I think, make us want to be quite sure that it looks robustly non-inferior. And, remember, we accepted a 36 percent difference as what had to be ruled out. That's not very aggressive. But based on assumptions about sample size that did not presume an advantage -- maybe it could have -- that was what seemed to be the necessary sample size, and that was pretty big already. But, of course, now you have the data, so you get to look and see whether you're reasonably convinced that non-inferiority of a variety of degrees has been ruled out, and that

all depends on which analysis you look at.

I do want to say, again, that the concern about INR is mostly about embolic or thrombotic strokes. And most of the advantage of the drug -- a little bit maybe. But most of the advantage drugs are on hemorrhagic strokes. So these low INRs shouldn't have any impact on that. I mean, if anything, it should make warfarin look better.

DR. LINCOFF: Dr. Krantz?

DR. KRANTZ: Just real quick to follow up with Dr. Rose, you had mentioned that generally, those at greatest absolute risk derived the greatest benefit. And I was curious. The folks with prior strokes, which is the secondary prevention population that I deal with in the heart, mainly, you would think would do a lot better. And, in fact, their point estimate, their hazard ratio, was diminished somewhat.

So I don't know. Have you looked at data in terms of CHADS score or other things that might suggest, at least, that that principle is still in

place within the ROCKET study?

DR. ROSE: In this study, patients with higher CHADS scores tended to have less robust results than patients with lower CHADS scores. Yes, you would think it would go the other way. I agree with you, and I was surprised that the history of prior stroke, or TIA, or SE subgroup had that hazard ratio that it did.

But, remember, that's a comparison to warfarin, so it was around .9. But it doesn't mean it didn't work. It means it didn't work quite as well, relative to warfarin, as it did in the lower group. Now, why it wasn't the other way, if I had to bet going into it, I would have expected it to go the other way, that the higher-risk subgroup would do better.

Maybe a lot of those patients had sclerotic vessels in their brain, and that's why they had strokes. They didn't have thrombotic strokes. They had strokes that were based on arterial sclerosis of the cerebral vasculature. And that's what we're seeing, and maybe warfarin doesn't do

much for that, and maybe dabigatran doesn't do too much -- or excuse me -- rivaroxaban doesn't do too much for that.

DR. KRANTZ: I'm just a little puzzled. Does that sort of lead to a question about the robustness of the data or is that a function, somehow, secondarily, of a moderator like TTR? I think the sponsor did a nice job showing that age, heart failure, other co-morbidities, con medications, all lead to lower time in therapeutic range.

DR. ROSE: Well, they do a little bit, but not very much. When you adjust the study to match, say, the RE-LY population, the TTR doesn't change all that much. It's not that different, at least when we do it. And we did it three times, three different ways, and it never changed more than about 2 and a half points.

DR. LINCOFF: Dr. Sager?

DR. SAGER: Despite Bob's comment about how lower INRs should reduce the incidence of warfarin-induced hemorrhagic stroke, we're really struggling

with the TTRs and the difference between this study versus other studies.

So the question I want to ask you is, you've done many different types of statistical analyses. Have you looked at the individuals on a center basis who were between, let's say, 60 and 70 percent TTR, and calculated their primary endpoint in the safety sets and the ITT sets, and whether their confidence intervals fell within the non-inferiority range? Because that, again, would kind of be comparing at least a little bit, even though we say we don't want to compare across studies, but it would be looking at the people who at least have higher TTRs, more consistent with previous investigations.

DR. ROSE: So the answer to your question is no. We haven't looked at the subgroup of patients between 60 and 70. I would remind you that, as you move up from 60, the patients start disappearing fast. And so most of the patients, if that subgroup were 2,000 people, most of them would be crowded in the lower end in ROCKET, maybe not in

other studies. But in ROCKET, they would be crowded at the lower end.

DR. SAGER: A follow-up question, if I could just quickly. In RE-LY, I understand that it was fairly equal both above 3 and those below 2. Is that correct?

DR. ROSE: Those above 3 and those below 2?

DR. SAGER: For INRs?

DR. ROSE: Oh, you mean in terms of INR? I will defer to the team that did the RE-LY review. You're speaking of TTR?

DR. SAGER: Yes.

DR. ROSE: I don't know the RE-LY TTR data all that well. I'll defer to them.

DR. THOMPSON: Aliza Thompson. My recollection -- and perhaps someone from the audience can also correct me -- was that it was more common to have times below, but I don't remember what the exact distribution was.

DR. SAGER: Thank you.

DR. LINCOFF: Dr. Fox?

DR. FOX: Thanks. With respect to Sanjay's

1 questions about the 1995 policy statement, as
2 effective as, I think the spirit behind that policy
3 statement was a correct one and a good one, but I
4 think it was intentionally non-operationally
5 detailed for some of the reasons Dr. Temple has
6 stated, because if you take it to its extreme, then
7 it's saying, well, prove equivalence. And I'll
8 leave it to the statisticians around the table to
9 argue how legitimate an idea that might or might
10 not be.
11 Having said that, I think the FDA deserves a
12 lot of credit for choosing a reasonable and
13 pragmatic middle ground. There is nothing magical
14 about preserving 50 percent of the treatment effect
15 of warfarin or ARBs versus ACEs, for that matter,
16 which is another area where a similar approach has
17 been taken in non-inferiority trials in the past.
18 Keeping in mind that there's nothing magical
19 about that 50 percent and also keeping in mind that
20 the FDA has taken a very consistent -- at least
21 this division has taken a very consistent approach
22 with different sponsors in this therapy area to say

we treated this sponsor and their product this way, you've got to meet a 1.38 margin. That's what we like. And they went to the other sponsor, and the other one, and the other one, and it was the same number.

They weren't capricious about it. They didn't change their mind in the middle, because you can't plan and execute a trial like this -- it's a gigantic undertaking that takes years. And let's not forget the 14,000 people who devoted their time and morbidity and mortality to this study. You can't then get six or eight years down the road and say, oh, well, never mind. Let's take a different approach and apply a different standard.

So you ought to give the agency some credit for at least being consistent, even if there is no perfect answer. And as Dr. Temple said, he and the FDA tried to capture this in a detailed way in the guidance that they issued on the topic.

So going back to the policy statement, to me it reads a bit consumeristic. But let's remember, this is not toothpaste or laundry detergent. It's

not whiter than white. We're talking about people's lives and serious morbidity here.

The last point I'll make is, I agree with putting the emphasis on, don't just say, okay, you met the bar, but what's the point estimate? When I asked the sponsor the question in the morning session about the sensitivity analysis, that admittedly took a somewhat ridiculous extreme position of assuming all the events happened, and withdrawn consent in rivaroxaban, and essentially none in the comparator, and how did that look. That point estimate started to creep around to like the 1.0 line, and the upper bound started to creep up towards the 1.38 upper bound limit. And I could tell there was a little bit of discomfort on their part when they got that, but it still was within the bounds. And I think it shows that they've done their due diligence in exploring the data from as many different angles as they could.

One last thing I'd like to say is, I'm not persuaded --

DR. LINCOFF: Comments are later. So if

there's no question here, then I think we should continue with the questions.

DR. FOX: Okay.

DR. LINCOFF: Dr. Nissen?

DR. NISSEN: I'm still struggling, Norman and Bob, with the regulatory standard question here. And let me see if I can pose this some.

So I understand that you set 1.38, and I understand the logic behind that. But there are some other recent examples where you said that wasn't nearly good enough. And the example I'm thinking about are the antiplatelet agents, prasugrel and anticoagular (ph). If I'm not mistaken, didn't you say they had to really be superior to -- or what - did you say 1.38? Did you use the same standard, non-inferiority?

DR. TEMPLE: It wouldn't be the same standard. You'd have to make some estimate of what the effect of clopidogrel was and then preserve half of that. I'm absolutely positive the effect of clopidogrel is not the 75, 80 percent reduction that you get with -

DR. NISSEN: That you set a margin --

DR. TEMPLE: What you do is you try to figure out as best you can -- this is not easy. Our biostatisticians spend a lot of time doing it. You pull data. You figure out what the effect of the drug is. I don't remember. For clopidogrel, it would have been in the neighborhood of 30 percent, or 25 percent, or something like that. And then we would, typically -- although you can debate this, our non-inferiority margin, M2, would generally be half of that. Now, somebody could say, well, geez, I don't want to lose half of the effect. But as a practical matter, the trials were already bordering on impossible. So much more than that, you just wouldn't have any trials.

As people have pointed out, the best way to be non-inferior is to be a little better. That really helps you a lot with the margin. Whether it's safe to assume you're going to be better and calculate your sample size that way is something that companies have to figure out.

Anyway, the margin would have been

1 considerably smaller there, but we did not reach
2 the conclusion at all that it had to be superior.
3 That is not so.
4 DR. NISSEN: So you use the same standard?
5 DR. TEMPLE: Yes.
6 DR. LINCOFF: Dr. Fleming?
7 DR. FLEMING: So much to get into, but I,
8 too, want to wait for the discussion. There's so
9 many issues to consider, such as the fact that the
10 sensitivity analysis wasn't done on ITT and many
11 issues around as effective that I want to come back
12 to.
13 I'd like to go to slide 23, maybe to just
14 address one issue. And while that slide is coming
15 up, just to comment, I don't think the issue around
16 TTR is whether it's a valid surrogate. I think the
17 fundamental issue here is, do we have different
18 warfarin regimens? Is a regimen that would be
19 intending to get to a target of 55 TTR different
20 from one intending to get to, let's say, 68 percent
21 TTR? Where the worry would be, if the non-
22 inferiority margin was based on the 68 but the

clinical trial was based on 55, and if that's less effective, then we'd lose the constancy assumption.

However, we're not actually -- this analysis, as useful as it is, doesn't get at the question I just mentioned because it is a post hoc indirect assessment. When you target 55, you don't get 55 in everybody. When you target 68, you don't get 68 in everybody. So this is, I think, an informative analysis, but one that has to be viewed with caution.

When I look at it, if you look at 67, 68 percent, this result looks to be a little more suggestive than what the sponsored showed, that INR could be an effect modifier here, to a degree, of achievement of INR, it could be an effect modifier. Obviously, again, it's post hoc, small amounts of data. But once you get to about 68 percent, you start approaching 1.0, which actually, though, could be reassuring still. It's 1.0 as a point estimate, although two concerns. One is the missing data concern, and another one is the ITT versus on-study concern in a trial like this, where

what happens after you go off per-protocol on-treatment is not neutral and is a lot of the events.

Therefore, do we have this analysis? I assume this analysis is based on the per-protocol on-treatment. Is that correct?

DR. ROSE: That is correct.

DR. FLEMING: Can you show us this analysis, based on the ITT?

DR. ROSE: I don't have it.

DR. LINCOFF: The sponsor has very similar slides. Do you have that?

DR. FLEMING: But I'd like to have the slide that uses the exact same methodology used, because this tends to show more effect.

DR. LINCOFF: Yes. I think the only difference is, the X axis here is percent of patients instead of TTR cutoff.

DR. ROSE: It doesn't look --

DR. PETERS: A primary endpoint ITT to follow-up with our CSR method. We only did ITT to follow-up for these. We did on-treatment and ITT

to follow-up.

DR. FLEMING: I really want to see the FDA's, and it sounds like the FDA is saying you don't have it.

Okay. We can move on. I want to get to the discussion.

DR. LINCOFF: So are there any other questions for the FDA?

DR. TEMPLE: Can I just ask something? Remember, the advantage that rivaroxaban had was almost entirely, but not quite entirely, on hemorrhagic stroke. Okay? So now you saw that that advantage starts to go away when you look at people who are more time in range. And I just wondered whether you thought that makes any sense, because at time out of range is mostly below. Why should that interfere with the advantage on hemorrhagic stroke?

DR. LINCOFF: So since that is a question maybe I -- mechanistically, first of all, intracranial hemorrhage that were not related to strokes, such as subarachnoid, although a minority,

1 were not part of that endpoint. So, I mean,
2 mechanistically, hemorrhagic strokes can happen
3 from ischemic strokes that then transform, even
4 though they're not obviously hemorrhagic
5 transformations. And maybe the CEC would want to
6 comment on that, but I think that preventing an
7 ischemic stroke that would then be more or less
8 likely to become hemorrhagic is still
9 mechanistically linked to anticoagulation.
10 The sponsor had been asked to -- yes?
11 DR. ROSE: I did have a response to
12 Dr. Fleming. We looked at that ITT to site
13 notification analysis in a couple of our TTR cuts.
14 And what we saw -- and I don't remember the data; I
15 didn't put them in any of our slides. But
16 consistently, in the TTR cuts and in virtually
17 every other analysis, the hazard ratio for the site
18 notification cut or site notification analysis was
19 higher than for the last dose plus 30.
20 I can tell you what that is. When you get
21 up into the 70s, the last dose plus 30 is around
22 1.30, 1.37 for the point estimate, and the hazard

ratios are maybe up to 3 or 4. So it's going to be higher than that. And the last dose plus 30 is around 1.3, in that range.

DR. FLEMING: And the last dose plus 30 is higher than the last dose plus 2, which is what this is based on.

DR. LINCOFF: Dr. Kaul?

DR. KAUL: At the end of your talk, Dr. Dunnmon, you said that clinical pharmacology attributes suggest BID dosing of rivaroxaban. And how predictive are those data? I mean, you utilize the same pharmacometric modeling in the dabigatran trial. And the 150-milligram dose that was approved, based on that pharmacometric modeling, would have yielded a ratio of 1 stroke prevented for 3 bleeding incurred. And yet, the outcome data was the opposite, 4 strokes prevented for 3 bleeding incurred.

So my question to you is, how often have you seen that kind of pharmacometric modeling data be congruent with the outcome data? And, therefore, what is the clinical relevance of that?

DR. DUNNMON: Actually, I think I'll defer to my clinical pharmacology colleagues on the dabigatran modeling, because I was not involved in that.

DR. MADABUSHI: This is Raj Madabushi, clinical pharmacology. So what Dr. Dunnmon presented was, if you believe the exposure of a drug is important, you would want to have as much as exposure without too much fluctuations within the danger dosing interval within that date.

If you believe in that concept, then you would think that a BID regimen, which gives almost reasonable exposure over the entire area with less fluctuation be more effective. That is the type of concept.

Now, what you are trying to bring up is a comparison with what happened in dabigatran. Let us not forget, in dabigatran, the exposures which were collected, that subpopulation. If you were to look at the point estimates when you compare it with warfarin, it was different than the overall population. So we have this limitation of the

data, which has exposures available, and sometimes even aren't expressed fully. So there is a reason why one may have to collect exposures in almost everyone, so that you keep the overall events in tact and not work off a smaller subset.

I don't know if that explains, but here we are talking about if you believe exposure, systemic exposures have any value, then you would want to have them as smooth, which is how you would have it for warfarin when it comes to INR.

DR. LINCOFF: The sponsor had been asked to provide some additional data over the lunch break, and so if they want to make that presentation, then we'll move onto the questions.

DR. PETERS: If I could have Dr. Califf as well, and I think we'll try to go in order. The first question was major bleeding without hemorrhagic strokes. Slide up.

So the overall hazard ratio was about 1.03, and hemorrhagic strokes were in our favor. So when those come out, it's a hazard ratio of 1.11, with an upper bound of 1.29.

I don't know, Rob, if you want to comment.

The next one was there was some discussion of the hemorrhagic and the non-hemorrhagic components of the primary endpoint. So we actually did the hemorrhagic, and combined the unknowns, and the non-ischemic, and non-CNS embolism to see what that split looks like.

Then after that was the time on treatment -- or the time off treatment as a percent of the total follow-up time in the study. So slide up.

This shows, again, the primary hemorrhagic strokes were 25 and 50, with a hazard ratio of .59, upper bound of .93. The rest of the primary efficacy endpoint, which would be, presumably, thrombotic and the few unknowns had a hazard ratio of .84, with an upper bound of 1.04, just to split those out.

In the next slide up is the percent of the time in the study. So the total on-treatment patient-years to our last contact -- so this is ITT, regardless; I'm looking at James -- was a

little over 11,000 patient-years in both groups; or the total was a bit over 13,000 in both groups, and the on-treatment was about 11,000 in both groups, so about 83 percent in both groups. So the off-treatment would be the other 18 percent.

We did have a few clarifying points to the FDA presentation, if we'd be allowed to go through those. First, if we could have the core slide 97.

DR. FLEMING: Before we leave this, could we have one question on this?

DR. LINCOFF: Yes.

DR. FLEMING: This I think may have been in response to the question I asked, which was trying to understand what fraction of the total person-years follow-up that would be in the ITT is also included in the per-protocol on treatment. And these numbers are somewhat consistent with the reflection that we have 20 to 30 percent more events for strokes and MIs in the ITT.

Why do we have fourfold the number of deaths?

DR. CALIFF: Let me try again. We spent a

1 lot of time worrying about this in the midst of the
2 trial. And so I think you have to understand, a
3 sick population, multiple co-morbidities, non-fatal
4 events occur, bleeding events occur, people come
5 off of drug, and after you've had something happen,
6 your likelihood of something else really bad
7 happening goes up multiple fold at that point.
8 I mean, another reason, frankly, to come off
9 of a double-blind trial like this is you've got
10 very complex medical care going on. And
11 participating in a clinical trial, if you have to
12 go to an assisted living facility or such, is just
13 very hard to do. So all the events are compressed
14 into the time after people stopped study drug.
15 That's our best interpretation of it.
16 DR. FLEMING: So I misspoke. It's threefold
17 numbers of deaths, but the point is the same. So,
18 in essence, what you're confirming, Rob, is that
19 this could be highly informative censoring, because
20 what happens as you discontinue could be very
21 different from when you are on the treatment, which
22 could be partially treatment related and partially

related to the inherent risk of the patient when you discontinue. That's my interpretation.

DR. CALIFF: Yeah. I don't think censoring is uninformative in any clinical trial, and I think --

DR. FLEMING: On the other hand, you might argue it's less informative for some other measures that are occurring with the same frequency when you go off than when you're on, as opposed to here, that are occurring. You've got 20 percent of the person-years of follow-up contributing 200 percent more events.

DR. CALIFF: Yeah.

DR. FLEMING: That's stunning, in my experience, in clinical trials.

DR. CALIFF: You shouldn't be stunned in a trial like this, but I understand it's a new adventure here. So it's informative censoring, but we've done a number of looks at this. And it's mostly informing that bad things are going on with patients. The dropout rates or coming off treatment were the same in the two treatments.

What happens after people stop study drug is they go back to the natural frequency of events. And if you look at the time off treatment, the event rates are the same in the two populations.

DR. FLEMING: We're not saying that, but we'll talk about that in discussion.

DR. CALIFF: So, Dr. Rose -- by the way, my name is Califf, for those of you who were confused. It's not Cahill. I just want to be identified correctly. But Dr. Rose made the point that differences in definitions can move centers around in center TTR. And that is true, but the example he chose was a pretty extreme example. And we just want to make the point that, in reality, when we apply all the different definitions that might occur and look at the quartile analysis, it comes out roughly the same. It's not causing great inconsistency. There is some difference, but it's not causing great inconsistency.

Anyway, just one more slide, I think. And I think this is an important point. When Dr. Rose referred to the RE-LY study in the middle, he

showed that the lowest quartile, the one at the top, and the highest quartile, the one at the bottom, without statistics, and said the hazard ratios are markedly different.

But if we look at the entire RE-LY study with their own analysis, the interaction test is not significant. Things come out in between. There is a trend, but I just think it's important for everyone to recognize, most of what I think we're talking about here is random noise because tests for heterogeneity are not significant in almost any part of this trial that we've looked at.

So we want to be sure that when we discuss RE-LY as it relates to quartiles, that people understand there was not a significant interaction between center TTR and estimate of the treatment effect.

DR. PETERS: If I could just show the clinical pharmacology for the switching study, the table of the INRs, because there is an assertion made that the INR, if we measured it at trough with rivaroxaban, might not be accurate, so that we

1 might not have a good transition strategy.
2 Slide on. So this is a phase 1 study we
3 did, and people were taken to steady state of
4 warfarin and then rivaroxaban added. So we're
5 looking at the pharmacodynamics of the combination
6 of those two drugs. The first day is the day zero
7 trough, the INR on the warfarin without
8 rivaroxaban. You can see it's a little over 2 in
9 both groups. And then rivaroxaban is given after
10 that trough value, and the next days reflect the
11 measurements on a 20-milligram dose of rivaroxaban.
12 You can see there's very little difference
13 in that measured INR. It's very slightly higher in
14 the rivaroxaban group, which reflects the
15 interaction at trough. At peak, there is a large
16 interaction, so that you would have to measure the
17 value at the trough to be able to know when to stop
18 the rivaroxaban therapy safely.
19 DR. CALIFF: Mr. Chairman, I think I
20 misspoke a little bit, or Dr. Fleming didn't
21 understand what I was saying. If I could clarify
22 one point about informative censoring.

If you start with the trial we have, treatment A and treatment B, treatment B has a better result than treatment A, which is what we saw particularly in the U.S. We would argue that's random fluctuation because it's totally within the range of the results, seeing it's a whole trial. But it is the actual result in the U.S.

Then you stop treatment in both groups. There's a lot of literature on this, as you know, that typically, the hazard of the group that got the better treatment is going to be somewhat higher because the high-risk patients are the ones who typically are spared from having an event.

So in the midst of the trial, you stop treatment in both groups. The event rates go up in both groups, but more in the group that had the better treatment effect to begin with.

But then particularly at the end of this trial, in the warfarin group, you continue anticoagulation with warfarin as it was before. And in the rivaroxaban group, you essentially have an under-anticoagulated group, so it does go up.

So our argument is not that the event rates are the same in both groups after discontinuation. It's that there's an explanation for the phenomenology that was seen, that at least seems to us to be a consistent explanation with what was observed.

DR. FLEMING: And I don't think I misunderstood. I think I understood. But the issue about what you're saying, it's only 2.8 percent that people had events. So while what you're saying is true, it wouldn't seem to be profoundly influential in a study like this.

DR. LINCOFF: Are there any more questions for the FDA? Yes, Dr. Sager.

DR. PETERS: Could I just make one comment on the dose selection? Just to make a very brief comment about the last comment about the evening of the dosing being a concept that would be better. Certainly, in our data, when we looked at once versus twice daily in the ATLAS study, there was not much difference at the 20-milligram dose between once and twice daily for bleeding or for

efficacy. And there's other data with newer compounds that actually suggest that BID dosing with higher troughs might actually cause more bleeding. So I think it's really an open question about whether once daily or twice daily would actually be better, and we got the results we did with the regimen we selected.

DR. LINCOFF: Dr. Sager?

DR. SAGER: For the sponsor, do you have the categorical cuts like 60 to 70 or 60 to 75 percent TTR, and particularly what the confidence intervals were, the idea being if one looks at that higher range, does it still meet the non-inferiority bound? That's really the question.

DR. PETERS: The only analyses we did were the quartiles and then the sliding figures. I mean, you can draw on the -- so the primary efficacy slide. I mean, if you draw -- that's a center-based ranking of TTR. But you could draw the cut anywhere you would want on that line to at least divide it by centers.

DR. SAGER: I did that. But it still

doesn't give you what the outer confidence interval would clearly be on that.

DR. CALIFF: This is the picture, and if you follow along, you'll notice that right at about 70 is the least favorable point for rivaroxaban. This is the safety on-treatment population, not the ITT to end of study population.

The confidence interval at that point, if you considered all the sites to the left plus, that would look within the bound, and all the sites to the right plus, that would be above the bound. But let me just say that from the point of view of our executive committee, you design a trial to have power to achieve non-inferiority for the trial as a whole.

When we start slicing off subgroups and say does that particular subgroup meet the criteria, it's pretty exciting when it does, but it really shouldn't, because, otherwise, you should just do a trial a quarter as big or whatever that is.

So our interpretation is that these results are largely consistent with random fluctuation when

you get into that range, and you've heard the FDA's interpretation.

DR. SAGER: Thank you.

DR. LINCOFF: One last question.

DR. KRANTZ: Real quickly, just for Dr. Califf. Mori Krantz from Denver. You had mentioned that in explaining this rebound phenomenon that Dr. Fleming mentioned, a threefold increase in events, that, really, the higher-risk patients are spared the events. But I thought the FDA had implied that, in your higher-risk strata, they actually enjoyed a lesser or diminished efficacy of rivaroxaban.

Am I wrong on that?

DR. PETERS: Primary efficacy endpoint by the CHADS score, the subgroup that shows that, actually, there's no heterogeneity across the CHADS scoring, as we're waiting for that slide to come up.

DR. CALIFF: So that raises a critical issue about it, so go ahead and put the slide up. There's no heterogeneity, and Dr. Rose knows his

data well. And when you look at, for example, CHADS score 5 and 6, it's right on the line of identity.

These are all within the range of random fluctuation. And as you all know, when we talk about a general result, the relative effect in all the trials so far has been constant. The difference has been in the absolute difference. So if you have a higher-risk population with the same relative treatment effect, you have, per 100 treated, a bigger impact.

So we're not claiming that the relative treatment difference is usually greater in the higher CHADS patients. It's just for the same relative difference, the absolute difference is greater because you have more people who have events in the control group.

Does that make sense? That's what we're thinking.

DR. PETERS: The other group that went with this was the prior stroke, TIA, and embolism, where the hazard ratio was .91 with that event, still a

good result to the right side of the line. It was a bit more robust in the not-stroke, but in a borderline interaction test, which we really think those are consistent between the two groups.

Questions to CRDAC and CRDAC Discussion

DR. LINCOFF: We are going to have to move on now. So we'll now begin the panel discussion portion of the meeting. And although this portion is open to public observers, public attendees may not participate except at the specific request of the panel. So we have approximately three hours, a little short of that, including the break, devoted now for questions.

There are nine questions. Many of them are subparts. In order to try to keep the time of this okay, what I'll try to do is group some of the subparts together, and then when you speak, you can address any or all of those subparts that you wish to do, because otherwise, I don't think this is -

I think you all have your questions. The cover part is not in a slide. The advisory committee is asked to opine on the approvability of

rivaroxaban, a factor 10A inhibitor to reduce the risk of stroke and non-central nervous system systemic embolism in patients with nonvalvular atrial fibrillation. The support from the claim comes primarily from ROCKET AF, a double-blind study in which 14,264 subjects with persistent or paroxysmal atrial fibrillation and additional risk factors for stroke were randomized to warfarin or to one regimen of rivaroxaban. The trial was event driven. Mean exposure was 19 months. Important results are as follows, which you have.

The FDA review team identified bleeding as the only significant safety issue. However, despite results in the primary endpoint that appear to show superiority to warfarin in reduction of the risk of stroke, and systemic embolism, and no evident increase in bleeding risk, several issues warrant discussion.

So question 1, the committee is being asked to consider how effective rivaroxaban is and whether that degree of effectiveness is adequate for approval. So please comment on the adequacy of

1 the design of ROCKET AF. And I think we should
2 probably take each one of these separately.
3 Was the planned warfarin management strategy
4 reasonable? Anybody? Dr. Nissen?
5 DR. NISSEN: I think it was not. And there
6 are two reasons I come to that conclusion. One is,
7 compared to all of the other contemporary trials,
8 the percent of the time that INR was in the
9 therapeutic range was substantially lower.
10 Secondly, compared to the other two trials with
11 similar drugs, there wasn't an active protocol for
12 how to manage out-of-range INRs. And I think that
13 handicapped the effectiveness of INR management.
14 And I think that is probably one of the reasons why
15 the effectiveness was so much lower.
16 So I don't think the strategy was wise.
17 Knowing that we needed to see robust evidence of
18 non-inferiority with good use of warfarin, it would
19 have been much more prudent to have taken the
20 precautions necessary to make certain that the use
21 of warfarin met the highest contemporary standards,
22 and those measures were simply not taken.

DR. LINCOFF: Dr. Emerson?

DR. EMERSON: So this is just a question. The reason we do attempt-to-treat analyses is because we don't believe that patient's compliance is necessarily voluntary, that there can be medical conditions that are there.

So my question to you is, as the sponsor presented it, they made the claim that this was a far sicker population than is usually used in these other trials. What is your experience in managing warfarin in sicker patients? Are you able to get the same times or do you back off on warfarin therapy because these sicker patients can't handle it?

DR. NISSEN: Let me just respond and say that, as pointed out by the FDA reviewers, there were entire countries where, same population, same entry criteria, they were able to do it. So, clearly, there is evidence that if you do a very good job of instructing the sites, and you stay on top of them -- and, again, I don't know what the role of the data monitoring committee was here in

terms of giving some feedback during the study, but they really fell off the curve in terms of the use of warfarin. But there were specific areas where it was used very well, which suggested, in this population, it can be used well.

DR. EMERSON: I would agree with that as well. But if you went to Seattle and you went to Northwest Hospital, or you went to Harborview for the exact same indications, you'd get very different patients. And so that entire countries and knowing how many sites there were at the country -- and I'll grant you, I don't know that this information is there. But I do just question that, again, is it harder to keep sicker patients in the target range of a proper INR in warfarin?

DR. LINCOFF: Maybe I can address that. I mean, the FDA's own analysis, by three different ways, trying to adjust for the baseline factors and the sickness factors of patients, did not show significant change in the TTR, the average TTR, so at least by criteria that you could measure. But people's individual experience is probably less

relevant than what --

DR. NISSEN: In what country is Seattle? What country is Seattle?

DR. LINCOFF: Other comments on this question?

DR. KAUL: The sponsor has to be congratulated for doing a well-designed study. But no study is perfect, and there are some imperfections in each trial. And if there is one imperfection in this trial, it was the lack of a standardized algorithm for maintenance of warfarin dose adjustment. But some may view this imperfection as a strength of the trial, because that's what happens in the real world. Every local institution has their own practice pattern of doing this.

But nonetheless, I think a dosing strategy, a centralized dosing strategy, would have overcome the mediocre TTRs that were achieved in this study, and in other warfarin dosing procedures, such as the use of centralized unblinded experts, as has been done in previous studies, for example, in the

SPORTIF V or the AMADEUS study. They would have resulted in, perhaps, better TTRs. So that is one imperfection of this study that I identified.

DR. LINCOFF: I would like to make a comment, then. The question is, was the strategy reasonable? I personally think that it was a reasonable alternative. There are really two approaches. One is to have it somewhat artificial, but create the highest standard of a control. The other is to say that part of the advantage of an alternative to a difficult-to-use drug like warfarin is the fact that, in the real world, it is used imperfectly and that an alternative agent may be more effective in that regard.

If you looked at the meta-analysis, which was presented, the TTR rate was very similar, what was achieved in this trial. And so I think it's defensible that one would -- because the key point was not to have provided some sort of algorithm. And I gather that was an active, not a passive -- not an oversight, but an active decision, and a strategy. I think it's a

defensible strategy because it does reflect how warfarin is, in fact, used.

DR. FOX: Just to add to what you just said, I agree with everything you said, and would add that, outside of providing a strict algorithm for adjusting the dose, I think the sponsor and the executive committee did everything else that was reasonable to ensure that INRs were adjusted appropriately in the trial with respect to getting the data centrally, reviewing, getting feedback to the sites, education, and so forth.

DR. LINCOFF: All right. Then we'll move onto the next question.

Was it reasonable to test a single regimen of rivaroxaban and ROCKET AF? Was the specific choice of regimen reasonable, given the short half-life and non-linear kinetics of rivaroxaban?

Dr. Sager?

DR. SAGER: Yes. I would say that it was reasonable. I think, with the data that we've seen, while it may have shown some small differences between BID and QD dosing, those really

1 were not of clinical significance. And even though
2 the PK modeling suggests that BID might be better,
3 that was not clearly borne out in the studies that
4 were shown to us.
5 DR. LINCOFF: Dr. Nissen?
6 DR. NISSEN: Yes. I am going to disagree
7 and say that we're dealing with an area where too
8 little anticoagulation can cause serious problems
9 and too much anticoagulation cause serious
10 problems. And given that fact, to have a drug
11 that's approximately half of the half-life -- it's
12 a much shorter half-life than dabigatran, which was
13 developed as a twice-a-day drug.
14 Maybe this is a little bit out of line, but
15 my concern was that the dose was selected more for
16 a marketing advantage rather than for the
17 scientific data that was available, and that was a
18 mistake; that, in fact, there are very wide
19 oscillations. And what we don't know is exactly
20 when during those periods of time, events were
21 occurring. And what you have to worry about is
22 that the more marginal benefits that were seen here

might have been due to this very large oscillation between the peak and the trough.

I think it's very hard to argue, for a drug that's being used for a critical indication, that a drug with a 7- to 9-hour half-life ought to be given once a day. If you think about the history of medicine in general, that tends not to work out so well, and I think it maybe didn't work out so well here. So I think it was not the best strategy, and I think it may have contributed to the less-than-robust results.

DR. LINCOFF: Dr. McGuire?

DR. MCGUIRE: Yes. I'll chime in on both A and B, that I think they were reasonable, arguably not perfect, and certainly with the ATLAS data, it perhaps makes it a little less reasonable. But I think the data support the fact that it was reasonable. It was non-inferior to a very good medical regimen. So I think it was reasonable from the beginning. It may be improved upon, but I think we have pretty good data to support its reasonableness.

DR. LINCOFF: Dr. Fox?

DR. FOX: I'm going to agree that it was reasonable. I think that at the time the dose was chosen, the data that they had at hand indicated a couple of things: number one, that the biomarkers they had at hand to estimate pharmacodynamic effect indicated a very flat dose effect relationship. While I was a bit surprised when I first started reading about it in the briefing materials, that the huge trough ratios around those dosing regimens were not considered as of primary importance, the bleeding data, in fact, don't really bear out much importance to that.

So I think they took a data-driven decision. They didn't have the ATLAS data in hand at the time they made the decision. And, sure, they could have added another arm to the study, but then you have a 21,000-patient study instead of a 14,000-patient study, so it gets a little unwieldy.

So I think maybe some compromises were made, but based on the data they had at hand at the time they designed the trial, I thought it was

reasonable.

DR. KRANTZ: Just real briefly, this is Mori from Denver. I think it's a good decision. It makes a lot of sense from the data. The thing I worry about is the one-size-fits-all approach with this kind of drug. I think, had we not had a look through the prism of that real big spike of events when you came off the drug, I wouldn't be as concerned. But I think that colors my thinking.

I look at some of the BMI data, for example, and the ranges were anorexia nervosa on up to 40 meters per kilograms squared or whatever have you. And there was a lot of divergence in the data. I don't know if you guys have that slide. But I wonder, this might be a problematic therapy in those patients that really fall outside the bell curve. So I think not having an ability to adjust it or having a BID could create problems.

DR. LINCOFF: Dr. Emerson?

DR. EMERSON: So I guess I am certainly in the category of I sure I wish I had data on the BID. And I don't, as much, I guess -- if listening

1 to everyone tell me that, no, this was just not the
2 optimal warfarin therapy, then why couldn't we
3 prove superiority? And we didn't hit that. So
4 this constant focusing back on, yes, we've proved
5 non-inferior to a treatment that there's a
6 suspicion was non-inferior -- or I should say was
7 inferior, then I don't necessarily buy that the
8 fact that we were non-inferior to that was the
9 proof that we had the right dose. And so I was
10 actually swayed by the data that we saw about the
11 pharmacokinetics and the BID versus QD.
12 DR. LINCOFF: Yes. Dr. Papademetriou?
13 DR. PAPADEMETRIOU: I, too, believe that the
14 used-once-a-day regimen of this drug was a
15 disadvantage and didn't let it exert the full
16 benefits of the drug, and probably increased the
17 risk of bleeding, since we have seen some data
18 suggesting that the BID regimen of the same dose
19 were associated with fewer bleeding complications
20 risk. So I would have liked to have seen the BID
21 regimen being used.
22 DR. LINCOFF: I will express an opinion on

this as well. I think that it may well have been better with the BID dose, but the data's presented for this dose -- and we should be looking at that at face value. There may be a better way to give the drug, but I think the data we have is what we have to evaluate.

I'm sorry, Doctor. Go ahead.

DR. MCGUIRE: Just a very brief comment to Dr. Emerson's comment. The question, as I read it, is not was this the right dose, but was it a reasonable dose to study. So I'm not suggesting it was the right dose, either. I would love to see some BID data, but I do think it was reasonable.

DR. LINCOFF: Ms. McCall?

MS. MCCALL: Reading a lot of comments from four different patient forums, I do appreciate that you had a very large elderly, and very ill population in this. But what the growing number is on the forums that I'm seeing are young, healthy people being diagnosed with AF, and it's their first really big disease process. They're not taking medications for anything.

I really appreciated that there was just QD dosing. I think you'll get much better compliance if they're only taking one med a day. It would be nice to see the BID study and to see what's going on. But I think for a younger, under-45 who have never had anything major going on, you're going to get them to take this drug because they've already dug in their heels about warfarin, and they're terrified of taking it. And I think this will be a happier thing for them, with one day, one time a day.

DR. LINCOFF: The next question, the primary analysis included events that occurred within two days of discontinuing study drug. For how many days should endpoint events that occurred after discontinuation of study drug, during the study or at its end, be counted?

Dr. Fleming?

DR. FLEMING: There are different answers that one might give for an ITT for a superiority analysis versus for a non-inferiority analysis. It's very important to preserve the integrity of

randomization. As we can, in fact, be confident, we are, if we have an ITT analysis where everybody is followed to the protocol-specified endpoint, or where censoring occurs in an uninformative way, such as everybody being followed X months from randomization or until some calendar date when the target number of events are achieved.

So I have significant concerns with the sponsor's analyses that were ITT, that were prespecified as the primary -- rather that were superiority, that were specified as primary to have been set up as on-study plus two days. So I believe that, for superiority, it should be ITT.

For the non-inferiority analysis --

DR. LINCOFF: Can I ask you to clarify? The question here, though, is the duration. If you don't think two days is right, what do you think would have been?

DR. FLEMING: So I have answered that. If this is ITT, the duration of follow-up is until you have the target number of events or until you've followed everybody X months from randomization.

It's not specifically following somebody X days after discontinuing drug for a superiority assessment.

For non-inferiority, there's been a lot of discussion about whether it should be ITT or per-protocol plus X days. And those advocating ITT say, even in non-inferiority analysis, you want to preserve the integrity of randomization. For those that are advocating per protocol, the argument is, we don't want to dilute the results toward the alternative hypothesis of no difference by including a lot of follow-up after somebody goes off treatment. In that event, though, when you are doing -- if you're in the camp that says, for non-inferiority, I want to do it per protocol, you surely would want to continue to follow during the time that you're still getting signal about treatment effect.

So in this trial, the concern is not only that there are a lot of events occurring after people go off, two days post-treatment, but that they are in fact not balanced by treatment arm.

1 And you seem to, though, capture a substantial
2 fraction of that when you go out 30 days in this
3 trial. But the number of days you have to go out
4 could depend on the setting. If you have an
5 immunosuppressive agent, and I'm worried about
6 inducing cancer risk, then X could be really large.
7 In this trial, it seems that if they had
8 gone 30 days, they would have captured a
9 substantial fraction of the differential treatment
10 effect. But when you look at what happens in this
11 trial after two days, you've got 20 percent of the
12 primary endpoints, 20 percent of the MIs,
13 30 percent of the strokes, and 65 percent of the
14 deaths are all occurring after that two-day
15 endpoint. And there are 26 excess primary
16 endpoints, 9 excess MIs, 19 excess strokes in the
17 experimental arm, in the rivaroxaban arm.
18 So in this particular setting, it becomes
19 very problematic to advocate for a per-protocol
20 on-study plus two day, as a way of getting rid of
21 the noise that's occurring after you go off
22 treatment, assuming it's just neutral noise. A lot

of the events are occurring there, and they're not neutral, and we've heard some of the reasons why, in terms of the transitions that are occurring thereafter.

So at least if they had gone 30 days here, by my calculation, they would have captured a substantially larger fraction of the continued signal that exists for what the actual treatment effect is in this non-inferiority analysis.

So bottom line is, many of us would argue, for a non-inferiority analysis, you need to be guided by the ITT as well as a proper per-protocol, that it captures sufficient duration of follow-up after treatment, that you're getting signal of what the treatment effects' influence is.

DR. LINCOFF: Dr. Nissen?

DR. NISSEN: In my view, stopping study drug is almost always informative in some way. And the problem is -- and I'm going to give you an analogy. One that really was a wake-up call for me was when I looked at the data for rofecoxib while on treatment, and then when you look at it in a true

intent to treat, where you follow them for the complete observation period, what you saw is that when people stopped the drug over the next few weeks to months, they had a lot of events. And what it was telling us is that people who stop drug, stop drug for a reason, and sometimes that reason is because they're sick. And that's why a true intent-to-treat analysis has to consider the possibility that events that occur after you stop study drug are related in some way to the study drug.

So that's why two days is just too short. It doesn't take into account the fact that something bad may have happened that has made that patient stop the drug, and what happens to them subsequently can be very, very important. And so to take that out and censor that all out of the study I think is not sensible.

Now, I can't tell you how many days to go. I think, appropriately, the agency and the sponsor show us data for various cuts at censoring. And the one thing we can say for sure is, the further

out that you look, the better warfarin looks and the worse rivaroxaban looks. That part, we can be sure of, but how far out you have to go, I don't know.

DR. LINCOFF: Dr. McGuire?

DR. MCGUIRE: I just want, Dr. Fleming, maybe for clarification, the plus 30 days, is that within the proposed, for example, site notification or accumulated endpoints, or is that being on the end of study? I agree completely with the plus 30 days or any point of time prior to the site notification. But to go beyond that after every patient is forced off study drug and onto open-label therapy, or no therapy at all, depending on reasonable preference, those additional 30 days' analyses aren't terribly informative to me for the primary analysis of efficacy and safety.

I think we learn a lot from them in this study about what we need to pay attention to if this moves forward, but I don't know. I would use whichever comes first, plus 30 days, or plus X number of days, or site notification.

DR. FLEMING: I was referring to the sponsor's primary analysis, per protocol, based on, on-treatment plus two days. So I was talking about X from on treatment. So in a per-protocol approach, where the concept is, the argument for per protocol, non-inferiority, is I don't want to include a lot of follow-up after people are off treatment when the effects of the treatment are no longer influencing the outcomes, since those will be neutral, and they will dilute me toward the alternative hypothesis of no difference.

So, in that setting, I want X to be -- those that believe that philosophy, you want X to be large enough that you're capturing the treatment-induced events or influence on events, but small enough to leave out what's noise and what's not.

So I was arguing for the per-protocol analysis, working from the end of treatment, off-study. I would argue X needs to be bigger than two days, in this study, certainly 30 days. This is just evidence based. Looking at the data in this trial, 30 days seem to capture where the

excess events were occurring off study. Once you got past that for most of these, it was more or less neutral. And, in fact, in my experience, in other settings, where people have used per protocol, in my experience, plus 30 days has been a pretty commonly used value of X.

This has nothing to do with what I am saying I would do for ITT. That was a different answer.

DR. LINCOFF: Dr. Kaul?

DR. KAUL: I think the period of follow-up should not be fixed, but should be driven by the pharmacologic properties of the compound of interest and how the trial was designed. So the pharmacodynamic effect lasts for about two to three days, so the two to three days after cessation of therapy is quite useful. And because of the trial design, there was no formal protocol-driven anticoagulation during the transition phase.

I would add another 7 to 14 days because I want to be sure -- I want to rule out the possibility of a rebound hypercoagulability versus a resumption of ischemic events as a result of

1 inadequate anticoagulation. So a reasonable
2 follow-up, given the pharmacologic property of
3 rivaroxaban plus the trial design is about 7 to 14
4 days. And I'm also concerned about the events up
5 to 30 days, but if you carefully examine the data,
6 most of them occur within the first 7 to 14 days.
7 So a 14-day period would be quite reasonable, based
8 on that rationale.
9 DR. LINCOFF: Dr. Sager?
10 DR. SAGER: I'd also like to congratulate
11 the sponsors for doing what's really a very
12 difficult trial to do, a double-blind trial on
13 atrial fibrillation. My concern with the two days
14 is that -- well, one is that many of the events, as
15 Thomas said, have occurred later. And then,
16 secondly, it doesn't really reflect how the drug
17 will really be used in clinical practice. Patients
18 stop drugs for multiple reasons, including that
19 they can't get the prescription filled, and they
20 may not get it filled within two days. So I think
21 a longer time period, both to be just practical and
22 reflect how medicines are used by patients, as well

as the fact that so many events do occur at the later time periods. A longer time period is necessary, whether it's 14 days, or 21 days, or a month. I'm not personally sure, but I think 14 days feels kind of like the minimum.

DR. LINCOFF: Dr. Temple?

DR. TEMPLE: How long to follow is one of the subjects of infinite variability and discussion. But I just wanted to comment a little bit on what Steve said.

If someone comes to us with a safety trial, designed to rule out a risk of X or whatever it is, and proposed to follow people for say -- and they thought the average duration of treatment was going to be six months, and they wanted to follow them out to a year, we would throw them out the door because there's going to be a regression to the mean if you do that, and it's a way to obscure things.

This goes to a little bit about the Vioxx thing. If someone came to us with a proposal to study the toxicity of Vioxx or something and wanted

to look for two months, four months, eight months afterward, we would be very skeptical, indeed. The attractiveness of looking at it in that case is because it came out bad. That's why everybody's so enthusiastic about it. These things have to be done prospectively, with as much rationale as possible.

The other thing, of course -- and I think everybody's reached this point -- if you take the drug away from someone and it was beneficial for a while, after a certain amount of time, they're going to have their events, and the groups are going to be equal.

Well, we know that, but we do intend to treat anyway because we're worried about -- and this hasn't really been discussed much -- informative censoring, that the people who leave the study prematurely may be about to have an event, and you lose them. And we're all perfectly worried about that. But it's worth remembering that the main disaster here in the people off therapy, they didn't leave because they left. They

left because the study was over. And this gave us an opportunity to see that the transition wasn't very well described, and so that's what happened. They had events during the transition.

It's also of interest, for what it's worth -- I'm sorry to obsess about this -- that the thing that happened after they stopped was that the thrombotic events increased in the treated group. It wasn't the hemorrhagic strokes, which is where the advantage was early.

So it's really perfectly plausible about what happened. What to do about it is another question.

DR. LINCOFF: Dr. Emerson?

DR. EMERSON: The aspect of using the pharmacokinetic or pharmacodynamic properties to determine the time of follow-up is a very real concern and something that you do like to take into account. It's just that the pharmacodynamic properties, depending upon which one you're talking about, are very different time scales, and most of which we don't know; and so the concept that we

1 have deep compartments. And so I tend to go
2 with -- I'd like at least 30 days follow-up on
3 things, but I am very sensitive to the idea that a
4 treatment might merely be forestalling events, not
5 preventing them, and the second that you remove it,
6 you remove oxygen and I die immediately. It's just
7 the same with heart-lung machines and things like
8 this. The second that you take them off it, you'd
9 be in trouble.
10 But I will say, along with what Dr. McGuire
11 said, that I wouldn't have batted an eye if he said
12 that what we're going to do is follow it, and the
13 final outcome for the primary efficacy endpoint was
14 going to be the day that we declared the study
15 over. I would want the additional safety outcomes
16 for that additional 28 days after the therapy
17 stopped, but you were trying to estimate what would
18 happen undertaking the therapy. So I would treat
19 those who stopped the treatment early very
20 differently from those who stopped the treatment by
21 protocol.
22 DR. LINCOFF: I'd like to comment as well.

I think these are two different issues. And I mean, we recognize it was a problem with the planned discontinuation and the lack of a protocol used for transition, but I don't think that that -- and that's a separate issue that has to be addressed, but I don't think, from my standpoint, that that speaks to the efficacy of the drug itself.

The unplanned or the early discontinuations showed similar event rates in the two treatment arms, and I think that's where the informative censoring might have been important. And even when you follow it out to 30 days, the event rate, 24.8 and 22.1 in the two treatment arms, suggests that when everybody comes off drug for whatever reason and they're sick, this is what happens.

So I agree with Dr. Emerson and Dr. Temple. These are two very different situations. And from my standpoint, I think it was reasonable to take the two days, although one could argue that for the early discontinuations, one might have done differently.

Dr. Kaul, you had it last.

DR. KAUL: Well, I just wanted to follow up on Dr. Temple's comment. In fact, there was increased bleeding with rivaroxaban post-discontinuation during the transition phase, 113 versus 68, with a hazard ratio of 1.65, confidence limits 1.22 to 2.22. So that would, on surface, argue against hypercoagulability. But the problem was that it was probably attributable to poor anticoagulation because there were more in rivaroxaban, where the INR was greater than 3.

The problem with the efficacy versus safety is, if by protocol, you're going to follow the study for safety events for 30 days, then by necessity, you have to follow them for 30 days for efficacy, because the endpoint that is driving the efficacy is a safety event.

So any which way you cut it, I think longer follow-up is warranted, a minimum of 14 days, and perhaps up to 30 days.

DR. LINCOFF: We'll move onto the next question. Are there other aspects of study design

that, importantly, affect interpretation of the study? Dr. Fleming?

DR. FLEMING: We've discussed some of these. There are four or five that I think of as important aspects that impact interpretation. We've talked about time to therapeutic range percentages. We've talked about the handling of the transition and how that's impacted our interpretation.

A related point, then, is the way one is defining the per-protocol analysis, whether it's on treatment plus X, shouldn't X be a larger number than 2?

Missing data, we talked about during the earlier discussion. As has been noted by FDA, and by the Institute of Medicine, and many others, this is a really important issue to interpretation of results. And I'm concerned that you'd have a withdrawal of consent rate on the order of 5 percent.

Then, the choice of the control, and we'll talk more about this later on, so I won't emphasize it now, but it made sense to choose warfarin. I

wish I knew what the results were against dabigatran. But those are some of the issues that I think are influencing the interpretation of the results.

DR. LINCOFF: Dr. Krantz?

DR. KRANTZ: The only thing that was a little bit unusual was the exclusion of cardioversion. And the reason I bring it up is I think Jonathan mentioned earlier the idea of this being a more pragmatic trial, that we're really looking at what real-world anticoagulation is like. And I think it's also relevant what real-world practice is. And a lot of cardiologists will, in first patients with afib, cardiovert them.

So I don't know if you have the data like the RE-LY trial. That might be interesting to look at, but I think that was one thing that stuck out as a design piece.

DR. LINCOFF: Dr. Papademetriou?

DR. PAPADEMETRIOU: I find it a little disturbing that the protocol allowed the investigators to wait up to four weeks to bring the

1 patients back, to check the patients with
2 subtherapeutic INR. That's too long, and it allows
3 for a suboptimal treatment to go forward too long.
4 And this was not the case in the United States,
5 where the patients came back, on average, in a week
6 or 10 days. But in other countries, it was up to
7 four weeks and I found this to be a limitation.
8 DR. LINCOFF: Dr. Kaul?
9 DR. KAUL: One fundamental issue I have with
10 this trial and trials similar to this is this
11 double counting of safety events towards efficacy.
12 And I think I have already made lots of comments on
13 that.
14 If you're going to count hemorrhagic events
15 to borrow an inference of non-inferiority with
16 regards to efficacy, then perhaps it's a good idea
17 not to double count it to infer superiority with
18 respect to safety. I don't think you can say that.
19 You can't have your cake and eat it, too.
20 So if you're going to use hemorrhagic events
21 towards your efficacy assessment, perfectly
22 reasonable to do that because that's the real life,

1 and patients don't really care what kind of stroke
2 they have. In fact, the case fatality rate with
3 hemorrhagic stroke is upwards of 50 percent in this
4 trial, and with ischemic stroke it's only
5 20 percent. So one would argue that hemorrhagic
6 stroke is what's more clinically relevant. I have
7 no problems with that, but I have problems with
8 then double counting it and claiming that it is
9 superior in terms of safety. And I don't think
10 that's appropriate.
11 The other issue I have is a technical one.
12 If you're drawing a non-inferiority margin on the
13 basis of historical placebo-controlled trials, four
14 out of those six trials specify that the stroke was
15 ischemic stroke. The other two don't specify. So
16 it's hard to tell how many of them were hemorrhagic
17 or how many of them were hemorrhagic stroke.
18 So if your non-inferiority margin is based
19 on ischemic stroke, then the relevant dataset for
20 analysis should only be reflected -- the efficacy
21 analysis should only be reflected in ischemic
22 events, because that's why you're borrowing the

historical placebo-controlled data. And I would like to see those results being emphasized in the discussion period in the afternoon, this intention to treat ischemic events at a 14- to 30-day follow-up, for us to be able to estimate what percent of efficacy of warfarin is being preserved. And the efficacy of warfarin is, essentially, prevention of ischemic stroke, not hemorrhagic stroke.

By the way, in this study, according to the data that I read in the briefing material, of the 310 primary ischemic strokes, only 11 went on to have hemorrhagic conversion. So only a minority of them, 3.5 percent, had a hemorrhagic conversion.

If you look at it the other way around, of the 90 total hemorrhagic strokes, only 11, 12 percent were contributed by the hemorrhagic conversion of the primary ischemic stroke. So I think it's important to draw this distinction for both technical reasons and for also philosophical reasons.

DR. LINCOFF: Dr. Sager?

DR. SAGER: Sanjay, I just wanted to follow

1 up on the excellent points you just made. I'm
2 wondering if -- sometimes it's just hard to be sure
3 whether a hemorrhagic stroke is primary hemorrhagic
4 or secondary to conversion. It depends on what the
5 timing was of the MRI, for example.
6 Just clinically, these are often very hard
7 to diagnose with, potentially, a fair amount of
8 clinical inaccuracy unless you get a very early MRI
9 and then serial MRIs, which is practical in some
10 medical centers, but isn't practical in many
11 medical centers, and also the time and
12 presentation.
13 So I mean, even clinically, these are often
14 a little bit hard to figure out which came first.
15 I just wanted to see what your thoughts were on
16 that.
17 DR. LINCOFF: I'd also like to add to that
18 comment, that in these old studies you quote, they
19 weren't doing routine MRIs or CTs, so it's very
20 hard to say that those didn't include hemorrhagic
21 strokes that we clinically considered. I mean,
22 from your own words, from a patient standpoint,

they don't care whether it's hemorrhagic or not hemorrhagic. They care about whether it's a disabling stroke. And I think these are reduced, and it's a reasonable comparison.

DR. KAUL: I don't have a problem with that. The key issue here is that if you're borrowing one endpoint towards inference of one assessment, I think you probably shouldn't be borrowing the same endpoint for assessment of safety.

DR. LINCOFF: Dr. Temple?

DR. TEMPLE: The distinguishing between the two strokes strikes me as really important, because you can easily imagine that a higher dose would affect one favorably and the other unfavorably. And you would really like to know that. And I think in some of the trials we're seeing, you're seeing stuff like that.

But I just want to be sure. You weren't objecting to the idea that the primary endpoint, at least, could be total stroke?

DR. KAUL: No. I'm not objecting to that.

DR. TEMPLE: You just don't want that to be

counted as safety, benefit also.

DR. KAUL: Right.

DR. TEMPLE: Just checking.

DR. LINCOFF: Move onto question 2. The interpretation of non-inferiority study depends upon certain understanding of the effect of the active control. If the active control is used to achieve less than its expected effect, defining of non-inferiority may not be informative regarding the effectiveness of the study drug. Similarly, a finding of superiority to a suboptimally administered active control cannot be used to support superiority of the study drug.

Please comment on the adequacy of the conduct of ROCKET AF. One measure of the quality of warfarin management, time in therapeutic range, TTR, was not as good in ROCKET AF as in many recent randomized controlled studies. So we'll group these together.

Was anticoagulation on warfarin in ROCKET AF good enough so that the warfarin group is an appropriate comparator to show, one, effectiveness

of rivaroxaban, two, superiority of rivaroxaban to warfarin? So please address both or either of those with your answers.

Dr. Nissen?

DR. NISSEN: I was hoping somebody else would step up first. But you look at those studies, all of which have TTRs up in the mid- to high 60s, and you look at ROCKET, and it really stands out. I understand that the sponsor doesn't buy this argument that the TTR is such a powerful predictor, but we've all seen the data on what happens to event rates, both bleeding and ischemic event rates, when you stay within that 2 to 3 range.

We know, and there's very good data in the FDA analysis, of what happens when you go from 2.0 just to 1.9 or 1.8. You're seeing increases in event rates on the order of 20 percent or more. And so all you have to do is downshift patients modestly out of that 2 to 3 range into that 1.8 to 2 range. And a fair number of people are shifting there, and you will have a loss of warfarin

efficacy.

So it becomes a major issue about whether it is an appropriate and adequate comparator. This would be a lot easier if we saw 65 percent mean TTR in ROCKET, rather than 55 percent. And no matter how you cut the data, you have to believe that it wasn't being used as well. It wasn't being used as well as it was being used in the trials that led to the determination of the confidence interval for non-inferiority, and it wasn't being used as well as in other contemporary trials. And so it just kind of falls off the wagon there and it's a problem.

DR. LINCOFF: Dr. Sager?

DR. SAGER: This is a really, really difficult one. It's fascinating that, actually, the reductions in the primary endpoint are predominantly hemorrhagic stroke. You can imagine if the TTR was increased, there'd be even more hemorrhagic strokes in the warfarin group.

Looking at the curve, we didn't have the cut of like 60 to 70 or 75 percent, but if you look

over that 60 to 70 percent range, it actually looks like the confidence intervals in 60 to 70 percent TTR, overall, would still be within the non-inferiority, though we don't have that exact data.

So given the fact that the reductions in the primary endpoint are on something that actually would be worse with a higher TTR, again, is just another factor here that as we deliberate, I think makes this just all the more complicated. I also am concerned that it's at 55 percent and not in the mid-60s, but I think there's a lot of factors here that we have to carefully consider.

DR. LINCOFF: Dr. Emerson?

DR. EMERSON: So I was asking earlier, and that people around the table who actually know seemed to all believe that this is a measure of inadequate treatment more than it is a measure of different sorts of patients. And so I'll defer to that.

But I will point out that the comparisons of the people at 1.9, versus 2.0, versus 1.8 were almost certainly observational data. We didn't

randomize them, and we can't randomize them to a particular INR. We can randomize them to a strategy to try to get that. And so, again, there's just this uncertainty in my mind about what's a patient effect and what's a clinician effect here.

DR. LINCOFF: Dr. Kaul?

DR. KAUL: I read the question differently from Dr. Nissen here. Effectiveness of rivaroxaban, I read it to see whether it is superior to placebo. And even with that poor TTR, I would say, yes, it is superior to placebo. So, therefore, it is effective. The superiority of rivaroxaban to warfarin, now this becomes a comparative analysis, and I would always like to have the best available comparator to make those inferences from. Clearly, it is not superior to warfarin.

DR. NISSEN: Let me --

DR. LINCOFF: Are you responding to this?

DR. NISSEN: I just wanted to make one other comment. And that is that everybody's trying to

1 make the effort here, or at least a lot of people
2 are, the sponsor certainly, to say, well, what
3 might have happened had -- let's look at the
4 tertile or quartile subgroups of people that were
5 at higher TTRs.
6 What that means is that we're being asked to
7 try to figure out what would have happened in this
8 trial if conditions that really didn't exist
9 actually had occurred. And it's almost like
10 saying, well, I'm going to take the subgroup of
11 people in whom warfarin was used well, and based
12 upon the results in that subgroup, I'm going to
13 declare rivaroxaban to be as effective as. And I
14 have a lot of trouble with that.
15 It is what it is. And any effort to figure
16 out what might have happened had everybody been in
17 the 65 to 70 percent range is taking and torturing
18 the data to the point that I don't know that we can
19 go there. I don't know what would happen, had
20 happened, if they'd had a higher TTR. What I do
21 know is what happened here is it didn't have a TTR
22 that fits in with contemporary trials.

DR. LINCOFF: Dr. McGuire?

DR. MCGUIRE: I'll respond to yes and no to the two questions. To Dr. Emerson's point, to underscore, the proposed odds ratios of subtherapeutic INRs that have been presented were, to my recollection, based on 69 cases and 69 controls.

If you look at the bottom of that table that was truncated in the slide, but is in the briefing document, the point estimate of hazard of 1.0 versus 2.0 was 18, which would imply a relative treatment effect of Coumadin just getting to 2.0 of 94 percent, which, even at the 1.4 presented on the slide, that would invoke a 77 percent relative risk reduction, going from 1.4 to 2. So those data are terribly flawed, I would say, and I think they overexaggerate the importance of the therapeutic range.

I'm curious, and I don't know -- maybe our hematologist -- what the variance of the standard deviation of the INR assay is. But I'm comfortable at the 1.8 to 3.2 range. I don't think the data

presented show us that being 1.8 versus 2.0 is so dramatically different.

So getting back directly to this question, do we all wish the TTR was higher here? I would say yes. I've been completely unconvinced that TTR reflects treatment adequacy or very reliably affects the -- as a treatment modifier. There are cases presented where there were numeric differences in the point estimates by TTR, but we've seen a lot of examples where there were no predictable differences either. So I don't have any confidence that the TTR really reflects what we really think it does.

Particularly confounded here, and, again, to a flaw, potentially, of the study design, was the long durations between out-of-range measurement and the subsequent measurement, forcing, probably, unfairly, almost unilateral bias to increase time out of range than the patient actually observed.

So I think this 55 percent is a worst-case scenario. And two other points. I think it fairly closely represents clinical practice. I called our

anticoagulation clinic, and our anticoagulation clinic TTR is 54 percent. We treat an underprivileged county indigent hospital patient population. So I'm not so sure this is so different than clinical practice.

The last point is, this is fairly comparable to the average TTR that was observed in the six trials used to generate the confidence limit for non-inferiority, so I think it's a fair comparison.

DR. LINCOFF: From my point, I would like to add that I agree with those statements, and in particular, the issue that the quartile analyses for each of the three major trials that have been performed, that was shown by both the sponsor, and I think also by the FDA, although without -- all four quartiles showed no relationship between TTR and the relative treatment effect of the new anticoagulants versus warfarin.

I realize that's not the same as saying if everybody was in quartile 4, what would the result be. The question is, looking at the trends, the quartiles, do the quartiles predict treatment

effect? And I think that they show fairly convincingly that they don't.

DR. KRANTZ: I just wanted to respond to that. Mori Krantz. If I recall Dr. Rose's data, when you got the quartile that was a TTR greater than 68 percent, I think they did, indeed, bridge the non-inferiority margin at 1.84.

Is that not correct with the FDA data? Not to be a stickler, but -- so I think one level of the TTR analysis did suggest that there was non-inferiority at that level.

But I think, overall, I agree more with Darren, that it cuts both ways, that if you get too stringent -- I realize it's an efficacy trial -- you really run the risk of not having generalizability. And, to me, I think that's sort of what's important here. And I also wanted to bring up that another advantage of this study that sort of negates that is the fact that it was blinded, whereas I think with dabigatran, when we reviewed that data, there was the issue of ascertainment bias that we don't have here. So I

1 think there's a lot of factors that kind of go on,
2 and it's a very complicated issue.
3 DR. LINCOFF: Dr. Emerson?
4 DR. EMERSON: So I was just going to mention
5 that, if we really did want to study in a very
6 scientifically and statistically rigorous manner,
7 this concept of what would the people in the
8 highest quartile do, we know how to do that.
9 Right? You first identify the people that you can
10 get a TTR in the acceptable range, and then you
11 randomize them.
12 But I would not really recommend that study.
13 I would really rather have the overall
14 effectiveness study of, say, in the patient
15 population. And on the dabigatran advisory
16 committee meeting, I stated that I was perfectly
17 happy with that being an unblinded study because so
18 much of the argument was the difficulty of
19 monitoring the warfarin and a lot of that.
20 So for a superiority study, I think it's an
21 artificial setting not to have people monitoring
22 that as they would monitor it. You don't want to

get complete incompetents doing it, but as long as it's standard clinical practice, you want them to do it the way they would do it. And that would be the way to establish the superiority, but that wouldn't be as acceptable in a non-inferiority study.

DR. LINCOFF: Dr. Fox?

DR. FOX: Thanks. Others have covered what I was going to say.

DR. LINCOFF: Dr. Kindzelski?

DR. KINDZELSKI: Thank you. I would like to continue what Dr. McGuire said about how difficult it is to keep the INR in a very fixed range. And the short answer is, it is difficult, and it also depends on what type of population you select for the clinical trial. It also depends on what kind of resources and which country you are doing the study.

However, we can all talk about how to make the trial better, best. We have the data right now. So we have to base our decisions on the data that has been already received.

Coming back to the question about the effectiveness of the drug and superiority, I do think those are two different questions. And I do agree with Dr. Kaul about, yes, rivaroxaban shows to be more effective than the placebo. And the question about superiority, this is something that is not very easy to answer.

DR. LINCOFF: Dr. Fleming?

DR. FLEMING: For completeness, the low-lying fruit here is regarding superiority by a proper ITT analyses, superiority was not established even prior to the concerns raised by these TTR issues. For non-inferiority, my sense, it's difficult to understand what the influence of TTR is in ROCKET. The FDA slide 23, though, to my view, does support at least the potential importance of a lower TTR in ROCKET regarding the validity of constancy assumption.

So my sense is, while I don't think this invalidates, I think it does meaningfully compromise the interpretation of non-inferiority against warfarin.

DR. LINCOFF: Dr. Sager?

DR. SAGER: Just one quick point, and we've really focused on the TTR, but if one looks at the percent that maintained an INR greater than 2, which includes those above 3, we're actually up to about 71 percent. That might have some bleeding risk, but from the efficacy side, almost three-fourths of the individuals in the study who were actually above 2.

DR. LINCOFF: The next part of this question, disposition of subjects in ROCKET AF is summarized below, and this is a table regarding the patients who completed drug and died, withdrew. Please comment on how the disposition data affect your ability to infer, one, effectiveness of rivaroxaban, or, two, superiority of rivaroxaban, to warfarin. Again, address both or either.

No comments? Dr. Fleming?

DR. FLEMING: My sense about this -- I'm actually answering parts B and C at the same time -- is somewhat similar to what I said in part A. Superiority is not an issue. The ITT

analyses do not establish superiority, even before you get into the irregularities that we're talking about regarding TTR targets, and disposition, and follow-up.

My sense, in general, is, again, I don't know if it invalidates the non-inferiority, but it is at least complicating the interpretation of the non-inferiority with the high levels of missing data that we have due to discontinuation, due to withdrawal of consent, and the very strong indication of informative missingness that we have when we truncate the data for non-inferiority assessments per protocol at day 2 with, as we've already said, not only differential numbers of events occurring, but a very high event rate, particularly for endpoints like death.

So my sense is, in addition to the first part, part A, which talks about the coagulation, parts B and C, I'm looking at in terms of completeness of information and informative missingness. It adds to the difficulty I have in determining the degree of reliability of the

evidence for non-inferiority.

DR. LINCOFF: So for others who may want to combine these answers, C is, was follow-up for endpoints adequate in both treatment groups; and D, are there other aspects of study conduct that importantly affect interpretation of study? In your comments, you can address any of those.

Dr. Papademetriou?

DR. PAPADEMETRIOU: It seems like about 20 percent of the patients completed this study off study medication. And although the numbers are not very different, the small difference is between the two groups. It seems like discontinuing the medication put rivaroxaban in a disadvantage, knowing that in the transition period, there were more events and that these patients were started, presumably, back to Coumadin. And in the transition period, if there was no coverage or bridging with heparin, this may have resulted in more events during that period in the group that received rivaroxaban as compared to warfarin.

So the disposition may have affected at

1 least 20 percent of the patients that completed
2 this study off drug. It may have affected
3 adversely the outcomes for rivaroxaban.
4 DR. LINCOFF: Any other comments on this
5 question?
6 DR. SAGER: Just quickly, I was somewhat
7 comforted by the sensitivity analysis that Jonathan
8 had asked for, showing that if one looked at worst-
9 case scenarios, still, the non-inferiority bound
10 was met.
11 DR. LINCOFF: Dr. Kaul?
12 DR. KAUL: I think Tom probably had
13 something first.
14 DR. FLEMING: I would have felt more
15 comforted if I had seen that on the ITT analysis.
16 I mean, I'm trying to put together all the pieces
17 here that are confounding our interpretation. And
18 it was reassuring for the non-inferiority, but much
19 more fragile is the ITT comparison, and we didn't
20 see any sensitivity analyses for that.
21 DR. PETERS: That was ITT to site
22 notification for the 77 events.

DR. LINCOFF: Could you go to microphone? And if you have a slide, you can show that.

DR. PETERS: What I showed was 77 events in the ITT to site notification, so it wasn't the safety on treatment. It was up to the notification.

DR. FLEMING: Yes. I was thinking in terms of what the agency had done that was presented in detail. I thought the agency's presentation was for the on-treatment plus two days.

DR. LINCOFF: Does somebody from the agency want to address that?

DR. ROSE: I'll give it a try. For the overall analysis, not looking at any quartiles or anything like that, we looked at both overturning non-inferiority and overturning superiority. Superiority is very easy to overturn. I think it took about 13 events.

John, why don't you talk about it? You did the analysis. I'm sorry. I didn't see you.

DR. LINCOFF: Please identify yourself to the mic.

DR. LAWRENCE: I'm John Lawrence. That graph is for the on-treatment with less dose plus two days, as it was for the primary analysis for non-inferiority.

DR. ROSE: There was another statistician who worked on this project, and he's left us, unfortunately, but he did other analyses. And they included ITT analyses of, I think, all the way to -- regardless of treatment. I'm sorry. I haven't looked at those recently.

I think, at a minimum, to overturn non-inferiority, it was something like 70 events, and there were some -- it's in the review. And then for on-treatment, it was upwards of 80 events. But even to overturn non-inferiority, it took at least 70. It might have been 80 also.

DR. LINCOFF: Dr. Fleming?

DR. FLEMING: It took 70 or 80 when you did this analysis in the on-treatment plus two day cohort, correct?

DR. ROSE: Now, are you speaking. There were several analyses done, and maybe I responded

wrongly. There was one analysis in which we looked and made assumptions about patients who had dropped out or were lost to follow-up. And then there was another analysis where we just looked and said, how many extra events would it take to overturn the finding of non-inferiority?

So which one are you talking about?

DR. FOX: The one that I asked for in the morning was for the patients who withdrew consent, so they say, I don't want to take the drug anymore, I don't want to talk to you anymore, I don't want to come for visits anymore. I'm gone.

If all the patients on rivaroxaban who left the study under those circumstances had actually had a primary event, or if you want to take a less extreme approach, had events at the rate that the patients on treatment had events, versus the patients in the warfarin arm, and if you assume they had no events after they left the trial on treatment -- essentially a worst-case scenario biased against the test article, what's that sensitivity analysis? And I believe that's what

the sponsor showed in an ITT population.

DR. LINCOFF: There is a table 29 in the FDA's -- so it's on page 116 -- sensitivity analysis of non-inferiority to superiority. And for intention to treat, regardless of exposure, it requires 88 events to overturn non-inferiority, and intention to treat to follow-up requires 80 events.

So I don't know, Dr. Fleming, if that addresses the issues.

DR. EMERSON: So I don't believe they did exactly the analysis that you were supposing. Instead, they said, of those that we were missing data, how many did they have to have? They didn't go to say everybody we were missing data on would have failed on one arm, and everyone not on the other.

DR. FOX: No. That's right. If you apply the observed event rate in the patients who were on therapy, still in the trial. I think that's a reasonable sensitivity analysis, as opposed to assuming everybody on test article drops dead as soon as they withdraw their consent, versus nobody

in the other arm, and that's kind of ridiculous.

DR. TEMPLE: But they did the more extreme analysis.

DR. EMERSON: Actually, as Tom's pointed out repeatedly, the death rate is actually far higher in the people as they've come off study. And so we ought to probably apply that to the people who dropped off, rather than the event rate that we're seeing in the people who did complete or were still on treatment at the end.

DR. LINCOFF: Question number 3. Please comment on effectiveness. How does rivaroxaban compare with warfarin as used in ROCKET AF, as used in the U.S., when it is well managed? Dr. Emerson?

DR. EMERSON: According to the non-inferiority margin that they agreed on beforehand, it's non-inferior. There was nothing that they did that didn't hit that question. So just that question, in terms of as it's being used, we saw that it didn't hit the margin in any of the analyses.

DR. LINCOFF: Dr. Nissen?

DR. NISSEN: Make sure I understand the context of the question. If that's the question, my answer is the same. But, in fact, when warfarin is well managed, I don't know that we know the answer, and that's the problem. And I want to make sure I articulate this, because, again, it's not so much how we're going to vote here. It's what are the policy implications of what we do here.

I worry about the slippery slope of creeping non-inferiority. So we say, well, okay, they had non-inferior with 55 percent of the patients in the therapeutic range. And then somebody comes along in a couple years and they are 50 percent. And we say, well, it's almost as good as they were in ROCKET.

The problem is that this can get us into some very dangerous territory. And I do think that we have to look at the other two comparators that have been approved. One is warfarin and the other is dabigatran. And we know that warfarin is very effective when used well, when used in warfarin clinics, the kind that we have. We do a little

better than you do at the Cleveland Clinic. I actually checked as well, and it's not uncommon to see rates up around 70 percent in well-managed warfarin clinics.

So are we going to take a step back if we do this? And my worry here, when you ask the question about when it is well managed, is that we approved dabigatran with something in the mid-60s. We approved, based upon rivaroxaban, in the mid-50s. Where do we draw the line? At what point do we say being non-inferior is no longer acceptable for patients when the outcome is a morbid mortal event of grave consequences? And my concern is if we don't draw the line in the sand somewhere, we're going to get into a serious problem with this kind of creeping non-inferiority margin. So that's part of my worry.

DR. LINCOFF: Dr. Krantz?

DR. KRANTZ: Yes. I kind of agree with Steve. I think for the first two, it's a no-brainer. I think it's got the efficacy versus warfarin. But then the question of when it's well

managed is really the lynchpin, isn't it? And I know I shouldn't worry about quartiles of TTR because we don't think it's a great surrogate and it doesn't track, necessarily, with outcomes. But in that analysis, the point estimate was 1.14 with an upper bound going to -- what was it -- 1.84. So, again, technically speaking, you would say it was non-inferior when it didn't meet non-inferior criteria when it was well managed.

So I think it's a lingering question and I think I don't know that we can adequately answer it. But that's the lynchpin, really.

DR. LINCOFF: We've been talking about upper bounds of confidence intervals on subgroups. And I'm surprised that the statisticians in the groups haven't yelled out. I don't know why we're talking about reaching boundaries or not reaching boundaries when we're talking about subgroups. The trials are never powered for that. I don't think that that's, at all, a valid point and we shouldn't be looking for that. That would be my comment.

Are there any other? Dr. Kaul?

DR. KAUL: For 3A, yes. It was non-inferior to warfarin with regards to efficacy. But I seldom interpret non-inferior to efficacy without asking the important question, what does it offer?

Okay. If I'm willing to tolerate some degree of non-inferiority, what is the tradeoff in return? It failed on primary safety endpoint, which was the major bleeding. It would still have met non-inferiority in the U.S., although it's getting closer to the margin; 1.33 is the upper bound, even though the point estimate is in the right direction. And if you just focus on the ischemic endpoints, it exceeds that non-inferiority margin.

The answer to 3C, which is when it is well managed, if you use 65 percent TTR as the cutoff for well-managed warfarin use, which is the mean, non-inferiority for a margin of 1.38 would still be met in the on-safety plus two days dataset, but not on the safety plus 30 days dataset. The hazard ratio of the point estimate there is 1.1.

If you use 68 percent TTR as the cutoff,

which is the median, non-inferiority would not be met for either the plus two or plus the 30-day dataset. The point estimates for the hazard ratios all exceed 1, 1.02 for the plus 2 and 1.30 for the 30-day datasets.

So if you were to borrow from Dr. Rose the operational standard for as effective as warfarin, which means that the point estimate favors rivaroxaban with the upper 95 percent confidence interval close to superiority, it could not be supported in cohorts where warfarin was well managed. And if you can do the additional analysis, the fractional preservation analysis, it preserves 58 percent using one analysis synthesis method and 75 percent using the other synthesis method of the warfarin effect.

So using well-managed warfarin, an inference of as effective as warfarin, based on the operational or the quasi-operational standard I heard from Dr. Rose, it does not meet that criteria.

DR. LINCOFF: Dr. Fleming?

DR. FLEMING: So regarding 3A, if we look at the primary endpoint, the hazard ratio by the per-protocol on-treatment plus day 2 is .79, stroke .85, death .85, MI .81. Those are all favorable point estimates. And by the way, they illustrate that you can actually do trials with rigorous non-inferiority margins and not have extraordinary sample sizes, because if you're a little better, you can rule out that you're modestly worse without huge sample sizes.

I think often we say, well, we can't do rigorous non-inferiority margins because the sample sizes will be enormous. Yes, if you're identical. And, in fact, how important is it, then, to approve something that in truth is identical that could have an estimate that says it's worse?

So if you're a little bit better -- and these analyses would suggest you are -- then you can, in fact, readily rule out you're worse, and the sample size could have been, actually, a fair amount smaller.

The issues, though, that come to

complicating this, as we've already extensively discussed, are the issues around the truncation at plus day 2, when things are happening afterwards in large fractions of patients, in a way that is imbalanced by treatment arm.

So if you look at the ITT analysis, those four relative risks become .91, .94, .89, .89. They're all still in the right direction, although much more consistent with no difference, but still, in my view, results that would establish non-inferiority.

So the complication to this is, how much influence does the joint issue of time in therapeutic range and missingness have to compromise the interpretation of those estimates that are near 1 that would, in fact, rule out the non-inferiority margin, if it weren't for those other issues, that in my view -- I don't know whether it means we cannot conclude non-inferiority. We may well, but it does complicate the interpretation.

DR. LINCOFF: Dr. Emerson?

DR. EMERSON: I just would like to ask, as we're looking at, for instance, the U.S. question, and seizing on the 64 percent that's the mean, I'll just note that unless we are only working in Lake Wobegon, where everybody is above average, even as we focus on the U.S. clinics, the fact that the average is 64 percent doesn't mean all of the U.S. clinics would be at 64 percent. And so, really, pulling down below that leaves us lots of room for the robustness of the results in the U.S.

DR. LINCOFF: Dr. Temple?

DR. TEMPLE: Is anybody moved at all by the following observation? The ITT for hemorrhagic stroke was essentially -- the point estimate was essentially identical to the estimate for the treatment plus two days. The change was entirely in the ischemic stroke, which is not surprising because we already sort of knew that. And there is a plausible explanation for why that occurred.

Does that influence anybody's view of the data? Because it doesn't seem surprising, if you didn't anticoagulate them, that they threw a few

strokes. I guess what I'm asking is, don't you have some sense that we might understand what's going on here?

DR. LINCOFF: From my standpoint -- I mean,, I've already said I think they're two different processes, what's happened afterward, and that although that's an unfortunate occurrence and may mandate either another study or some careful evaluation of what to do afterward, that that doesn't speak to the efficacy on therapy or the results of coming off therapy if you had adequately anticoagulated. So I agree.

DR. NISSEN: Bob, we can't know. I mean, that's the thing that makes this very frustrating. We don't know what would have happened had the protocol and the study conduct been done in a way that protected patients during that transition period because that wasn't what was done. And so we're in the position here of having to guess at what might have happened.

I would make the same comment with regard to this issue about, well, what about you take the

1 group of countries where TTR was better? What
2 we're then trying to do is decide whether or not a
3 drug is approved based upon a subgroup in a larger
4 trial, where we really don't -- I mean, subgroup
5 analyses, as we all know, are fraught with hazards.
6 So I think we have to judge this the way it
7 was conducted, not the way it might have been
8 conducted. And it wasn't conducted in a way -- and
9 the way it was conducted, some bad things happened
10 after that 2 to 30-day period. And that's all we
11 can say. And until somebody can prove to us that
12 that isn't going to happen if the drug is used
13 better, we don't know the answer.
14 DR. TEMPLE: No. I think that's true. But
15 what happened was that the rather modest, barely-
16 there advantage on thrombotic stroke clearly went
17 away. And maybe we don't know exactly why. But
18 the other advantage -- it always surprises me; I'm
19 not sure I understand why -- wasn't affected at all
20 by that. I mean, it's true, you've got to look at
21 the overall data, but we just always, nonetheless,
22 sort of look at the data, too, and help us

interpret it, and I just wondered what you thought.

DR. LINCOFF: We do know what happened in the sizeable proportion who stopped the study drug early, who didn't have an excess risk. I mean, they didn't have that differential effect of anticoagulation. And it's one thing to say that it's a subgroup, so we can't interpret it. But the subgroup of good TTR is the group that is concordant with the others.

Dr. Fleming?

DR. FLEMING: Just quickly on those numbers, Bob, for hemorrhagic, the difference between the on-study plus day 2 and ITT was plus 8 against plus 7. Now, it's only 15 events. The ischemic was 7759. And you're right. That may be where you would expect it. Of course, that's where most of the events were, and there we saw it, 7759. The hemorrhagic was 8 against 7.

DR. TEMPLE: But it's also where the drug didn't seem to have much advantage. The hazard ratio for the on-therapy was only .95 in the first place. It was a very, very modest difference,

barely there, you'd have to say.

DR. FLEMING: But you went from plus 12 to minus 6.

DR. TEMPLE: Yes.

DR. LINCOFF: Dr. McGuire?

DR. MCGUIRE: It's a nice segue. I was going to call Dr. Temple out the second time. When we talk about a point estimate of .95 against a placebo, it's awfully disappointing. Against aspirin for primary prevention, it's awfully disappointing. But against warfarin in this condition, I think it's fairly impressive to even come close.

We don't have any drug in all of cardiology that comes close to the treatment efficacy of warfarin in stroke. And so to have a drug that comes close at a .95 point estimate is awfully impressive.

DR. TEMPLE: Okay. But we just approved one that actually was better.

DR. MCGUIRE: Did better. Understood.

DR. TEMPLE: It's not impossible.

DR. MCGUIRE: A couple of comments. I do think it was -- I think answer A is yes. Answer B is probably. I agree, all the limitations of subgroup analyses, but often it's better to be lucky than good, and we're lucky in this case, in that the U.S. population, the point estimate actually achieves the -- it's favorable, and the confidence limit is well within the 1.38, the conservative confidence limit of the FDA. And I agree, we can't answer C.

DR. LINCOFF: Dr. Fox?

DR. FOX: So I think the answers to A and B are fairly straightforward. As used in the trial, yes. As used in the U.S., I'll take a slight different interpretation. Maybe the FDA wants to correct me on how they wrote the question. But as used in the U.S., to me, goes back to the meta-analysis of how warfarin is actually used in the U.S. It doesn't say as used in the U.S. cohort of ROCKET AF. So if you say as used in the U.S., it still looks pretty good.

I want to interject, maybe, some commentary

that goes back to the way Sanjay Kaul answered the question about, when should the per-protocol observation period have properly ended? And in my view, it ought to be 4 to 5, or 4 and a half pharmacodynamic half-lives of the test article, which, whether it's 1.5 days or two days, we can argue about. But it's not 7, and it's certainly not 30.

I think that the excess of events that happened at that 30-day visit is an unfortunate design feature of the study around the transition plan, which could have been a whole lot better in retrospect. But I don't think it's a reflection, necessarily, of the test for effectiveness on a per-protocol on-treatment patient population in the study. So maybe we should try and separate those two issues because that whole transition plan idea is supposed to be another part of our discussion, right?

So then, to answer C, when it's well managed, I'm actually not as persuaded as others that TTR is, in fact, a reliable -- call it a

surrogate for treatment effect or treatment modifier, if you prefer. But when the sponsor did the analysis of the other trials in terms of what the center TTR was, proportionately speaking, in the regions in which the trials were conducted, it seemed comparable.

So I think it is codependent on where it's done and in the sickness of the patient population. I think that's important.

DR. LINCOFF: We are going to take three more comments on this, and then we're going to take the break, so Dr. Sager, then Dr. Emerson, then Dr. Nissen.

DR. SAGER: Yes. One is I just wanted to respond to you and just raise the point that these early temporary discontinuations who then restart may be very different than the later discontinuations. These may have been people having procedures who were put on heparin, were temporarily stopped for some very specific reason that might have been managed in a very different way.

1 But I also wanted to come back to your
2 comment. Yes. No, it looks like, with a different
3 transition anticoagulation regimen, we might well
4 get pretty different results. I guess the only
5 thing that concerns me a little bit is, A, we don't
6 have the data, and, B, just how we do that so we
7 don't, for example, have an increase in bleeding.
8 And without data, it's hard to be sure. But I do
9 think it's hard for me to imagine that a regimen
10 can't be developed that would significantly change
11 the type of data that had to have been put in
12 place.
13 DR. LINCOFF: Dr. Emerson?
14 DR. EMERSON: Others have made it.
15 DR. LINCOFF: Dr. Nissen?
16 DR. NISSEN: I feel compelled to point out,
17 as somebody who takes care of a lot of these
18 patients, that it is very common to interrupt
19 therapy with warfarin. And patients will interrupt
20 it on their own. They'll get a lot of bruising.
21 They'll have bleeding gums. Things will happen.
22 People don't get to the pharmacy, elderly people,

and don't get their drug.

One of the things about warfarin, even though it's a drug we love to hate, is that its pharmacodynamic effects are slow on, slow off, and it's pretty forgiving. And so, as Sanjay Kaul has pointed out to us, in those first few days after you stop rivaroxaban, pretty bad things happen. And my fear here is without having studied a regimen for what to do with those patients, the drop-in, drop-out -- in a clinical trial, people are more likely to actually take the drugs in a reliable fashion. They know they're being studied. They're part of a project. There's certain phenomenon that occurs when you get in a clinical trial.

I'll tell you, your 80-year-old AF patient that can't get a ride to the pharmacy, who stops rivaroxaban, what we seem to have here, is in the first week, the potential for some really high relative risks of events, and we've not studied that. We don't have any information on how to ameliorate that risk, and it makes me concerned.

DR. LINCOFF: Can you do this in 30 seconds?

DR. KAUL: Yes.

DR. LINCOFF: Okay. 30 seconds.

DR. KAUL: I think it's important to understand the question here. It says well managed. While I agree that TTR may not be an effect modifier, it certainly is a marker of skillful care. How else do you expect or explain why rivaroxaban event rates goes down when the TTR goes up?

So it is a marker of quality of care, and that's what this question asks. And you have to use the TTR cutoffs that you pick up to define that quality of care. Is it 60 percent? Is it 64 percent? Is it 68 percent? We can disagree on that, but the key point here is that this is a marker of the quality of care, not an effect modifier.

DR. LINCOFF: We have now completed three of the nine questions.

DR. MCGUIRE: After publicly admitting our TTR rate at our institution is about 54 percent, I

have to defend that only one component of TTR is the care of the physician and/or the process and system. We saw Dr. Harold's (ph) multiverbal model, and about 20 patient-level characteristics also contribute to predicting time in therapeutic range.

So I will agree with Dr. Califf's comment earlier. I take some offense at the term "unskillful use" because it's a little bit accusatory, and it's a complicated therapeutic strategy.

DR. LINCOFF: We will resume this controversy in 10 minutes. We're going to take a 10-minute break. Again, there should be no discussion of the meeting topic. Please come back by 3:50, so we can finish the six questions.

(Whereupon, a recess was taken.)

DR. LINCOFF: Dr. Fleming, did you have a comment on the last question before we moved? I think you wanted to respond to the U.S.

DR. FLEMING: Sure. I want us to go onto question 4 as well, but just to take 10 seconds in

the U.S., just a note. As has been noted by others, we have to be very cautious in subgroups with 2,000 people here. It is noteworthy, though, that this issue of the events, the primary endpoints occurring after two days, almost one and a half times as many events in the U.S., patients occurred after they went off the treatment, from two days onward.

The other was that death and MI, even from the per-protocol analysis, had hazard ratios that were near 1 and were in fact effectively 1 when you did the ITT analysis, so results that, in fact, were consistent with the larger study, including the concern about large numbers of events on rivaroxaban that occurred after the patients went off for two days. But interestingly, the MI and the death were neutral in the U.S.

DR. LINCOFF: Moving to question 4, then. As part of the Clinton administration Reinventing Government initiative, FDA published a policy that said, in part, the agency does not require new human drug products or medical devices to be more

effective than existing therapies, nor does it necessarily require the product to be compared to other products. However, for products intended to treat life-threatening diseases, diseases with irreversible morbidity and contagious diseases that pose serious health risks to others, it is essential for public health protection that a new therapy be as effective as existing approved therapies.

The as-effective policy explicitly does not apply if the new therapy is studied in a new population. And considering how this exclusion might apply to rivaroxaban, here are some points for comparison of the warfarin arms in RE-LY and ROCKET AF.

So this is all warfarin arms, and you can see here different criteria, both on the left and on the right, the right being endpoints according to TTR.

The question, then, is, is the population in ROCKET AF sufficiently distinct from the population in RE-LY, that the as-effective policy does not

apply? And if so, how? Dr. Emerson?

DR. EMERSON: Can I ask for a point of clarification here? Because I would have said that this policy should be about what the indication is and whether the indication is in the new population, never mind the population that the drug was studied in, which, ideally, would be the population that the indication was there. But as I read the indication, it's very similar that they're looking for, and that that's what's important; not what the eligibility criteria were for the studies.

DR. TEMPLE: I think the question is trying to indicate our flexibility and willingness to consider a variety of potentially interesting but important differences, whether it really ended up in a different indication or not; that is, does it study a new population in a way that enthralls you?

I guess we thought that could be relevant, even if the indication turned out to be the same. I mean, you're sort of implying that you indicated only for people with CHADS whatever, and we wouldn't necessarily do that. We don't necessarily

do that unless, for some reason, you thought the data didn't apply to people who are less sick, which you might think, but you wouldn't necessarily think.

So if you thought studying the sicker people gave you information about the less sick people, you might indicate for everybody, but you still might conclude we had interesting additional information here. That's what we were asking for, advice.

DR. LINCOFF: Dr. Fleming?

DR. FLEMING: Actually, I had written down something very similar to what Dr. Emerson said. It seems that most of the discussion of this issue is in question 5. But in question 4, my sense was, if the label, if the indication, is the same as for another therapy, then I would think it applies.

DR. LINCOFF: Dr. Nissen?

DR. NISSEN: First of all, I do want to commend the executive committee and company for studying a particularly sick population. You know, that's very useful. However, it's pretty clear

1 that there is substantial overlap between the
2 patients that were in RE-LY and the patients that
3 are in ROCKET. And we don't have enough
4 information to suggest that these patients are so
5 different that they constitute a unique population.
6 Having said that, I'm glad that we have data on the
7 sicker population, because, frankly, these are the
8 people that probably the most need to be treated.
9 DR. LINCOFF: Dr. Sager?
10 DR. SAGER: I would say this study really
11 focused on the sicker patient population and
12 doesn't really say a great deal about the less-sick
13 patient population, but there is a lot of overlap
14 in the overall patient groups. So I don't think it
15 really reaches this bar, this policy, in terms of
16 it really being a totally different patient group
17 that we would look at totally separately; though I
18 think, later on, as we get into some of the
19 nuances, it may be difficult to extrapolate this
20 data to a less sick population.
21 DR. LINCOFF: Dr. Fox?
22 DR. FOX: I actually think the question is

inappropriate to the policy. So the policy says something about, it doesn't require the product to be compared to other products. And then if you infer that to be "as effective as" implies that it's to be compared to an existing approved therapy.

Well, when the trial was designed and run, dabigatran wasn't existing or approved, so it couldn't be compared. So dabigatran is irrelevant to the question, as it applies, as it's driven by this policy.

But having said that, I kind of agree with the people who are saying that, yes, it was a little bit sicker, but it was still pretty much the same population and generalizable.

DR. LINCOFF: Dr. Papademetriou?

DR. PAPADEMETRIOU: Yes. I agree that the warfarin arm of the two studies doesn't consist of a new population one way or the other, but there's a lot of other overlap between the populations, except there are differences. In the ROCKET study, there were a lot more sicker patients with CHADS4

to 6 score; a more percentage were in the ROCKET study. They had some differences in age and whether they were taking warfarin before, but there is a lot of overlap between the two groups.

So I don't think this should be exempt from the rule of as effective as.

DR. LINCOFF: Dr. Krantz?

DR. KRANTZ: No further comment on that, just that I think there's an err potentially on the age greater than 75. ROCKET looks like it's lower when, in fact, my sense is it should be higher. If you look at slide CC-37 from the sponsor, it seems like it was 44 percent were over 75, which would make sense for a sicker population. I don't know if that's accurate or not, but I just wanted to make the comment.

DR. LINCOFF: Dr. Kaul?

DR. KAUL: I am somewhat confused. Earlier today, we were told that the non-inferiority guidance document that came out in 2010 sort of reflected or was influenced by this policy statement. And so this new population and new

therapy is not a concern when we are doing a non-inferiority assessment because we are comparing it to a treatment population that was never studied in the old historical trials. And, yet, there is a concern when you apply it to RE-LY.

So I am somewhat confused, and there seems to be a disconnect here. So perhaps the FDA can help clarify this.

DR. TEMPLE: Well, I don't blame you for being confused. This was not my favorite question.

[Laughter.]

DR. TEMPLE: I mean, there is an interesting and provocative issue here that I'm not going to try to address our policy on. But suppose, in the course of a study, something new and really hot comes along? Do we ever say, I'm sorry you compared it to a dog, we have something better now? We don't usually do that, but I wouldn't rule out the possibility. And this is a little bit about that possibility, but I don't think we really meant to get too much into that discussion.

I mean, it's the big dog; in the range,

there's this other drug that does something. But I don't think -- as many people have said repeatedly, I don't a cross-study comparison here is plausible. It's a different population, and I don't know whether this study tells you anything about how the drug compares to dabigatran. I'm not sure it does at all.

So I think this reflects the fact that it's in the air. We know it's there. It's hard to ignore it, but I'm not sure this is the most important question. The comparison that we focused on when we discussed the trial with the company and that we're focusing on now, really, is with warfarin and not with dabi.

As I said at the opening, what dabi does makes me, at least, more interested in being sure that the comparison with warfarin is really solid; whereas, before there was any alternative, just having an alternative to the difficult-to-use Coumadin would have seemed like a very good deal, and any evidence of effectiveness might have been okay. But I don't think that's true anymore,

because you have alternatives that are really solid. But I don't think it makes dabi the comparator.

DR. LINCOFF: Dr. McGuire?

DR. MCGUIRE: Yes, just a couple points of clarification. From the policy, I'll just point out that the words "when skillfully used" are not in the policy. They're commonly linked in the FDA briefing document, but it just has to be proved against existing therapies. Now, it may be implied that it should be skillfully used, but it's not in there.

I think there's a lot of confusion here. Dr. Fox just alluded to approved therapy, but in fact, the word is plural in the policy. And so the final word -- existing approved therapies, does that imply it has to be proven against every approved therapy, at least one approved therapy, the very best approved therapy? The guidance is, as is in other places, non-operational.

So my interpretation is at least one approved therapy is sufficient. Warfarin was that

1 therapy at the time, and I don't think we should
2 consider the dabigatran comparison.
3 DR. LINCOFF: Mr. Coukell?
4 MR. COUKELL: I think the Federal Register
5 is certainly getting at a public health impact.
6 And so what the study population is, to me, is
7 irrelevant. The question is, is there a population
8 out there? And I asked the sponsor about that this
9 morning, and the answer was, essentially, no.
10 DR. LINCOFF: The next question actually
11 overlaps quite a bit, so we'll go to it, we'll read
12 it, and then those who feel like they've answered
13 with the previous one don't need to answer it here
14 again, but those who have additional comments or
15 want to amplify, by all means.
16 So number 5. If you concluded that the
17 policy does apply and that rivaroxaban needs to be
18 as effective as something, A, what does as
19 effective mean operationally; B, is it sufficient
20 to be as effective as warfarin? If so, is it? C,
21 is it necessary to be as effective as something
22 else? And if so, do you need a direct comparison

to the something else? And is it? I assume that means is it effective as something else.

So all of those questions, I think, are quite a lot of overlap with the last one, but are there any comments?

I think, Jonathan Fox, you had a comment before?

DR. FOX: Yes, just to Dr. Temple's comment before, and then I'll segue into my answer to these questions. I guess one extreme case would be that somebody comes along with a new therapy for a disease, for which there is a long-standing approved existing therapy and is so overwhelmingly effective that the trial is stopped early, and it makes big headlines. And, meanwhile, that new therapy still has to go through an approval process before it can be approved, and launched, and made available to people.

What would the trialists working on a competing therapy that was about as effective as supposed to do? Are they supposed to stop their trial because it's unethical to continue because

the newfangled thing is so much better, even if it's not available yet? I mean, there's a whole host of things that come into -- so dabi doesn't quite get to that extreme, not anywhere near it, in my view, but it's something you want to keep in mind when you start down that road.

As effective, I think, Dr. Temple, you set this up at the very beginning. Operationally, the agency, and this division in particular, declared, in your words and actions over the last number of years, that preserving about 50 percent of the treatment effect of the established therapy was a pragmatic way of getting innovative new therapies out there for patients, and you've been consistent in that way.

So is it sufficient to be as effective as warfarin? Well, warfarin was the only approved existing therapy at the time the trial was done, so the answer to that is yes.

Necessary to be as effective or as something else? Well, I guess that's the interpretation of the policy. And I think industry's been okay with

that, and most of the science world has been okay with that. And the direct comparison, well, I'm not sure how else you'd do it but to compare them directly head to head. And is it? I think we've had a long discussion about that already, about meeting a non-inferiority standard. So that's it.

DR. LINCOFF: I believe the second part of this question was directed at dabigatran, but I may be wrong.

Dr. Fleming?

DR. FLEMING: So to answer all three parts here, the first part, what does as effective mean? Well, technically, to be at least as effective as -- and I think Dr. Temple already pointed this out -- technically would mean you're ruling out a quality, which is what superiority would allow you to conclude.

Having said that, my interpretation of this, my sense of what was meant I think is -- I think similar to also what I heard Dr. Temple say a little bit earlier today, not quite so stringent that it would be sufficient to be non-inferior with

respect to an NI margin that preserves an adequately large fraction of the -- to comparator's effects. It might be more than 50 percent, but I don't think the intention was that we would have to show superiority.

So for part number B, is it sufficient to be as effective as warfarin? My sense is, if no other agents in this indication have been established to be superior in benefit to risk to warfarin, then it should be sufficient with a properly designed and conducted trial to be non-inferior to warfarin.

What I struggle with, though, is if there is another agent that's clearly superior to warfarin, why is it sufficient to then be non-inferior to warfarin? It was stated earlier that this policy says that we have to be as effective as other therapies.

I think it's implicit to mean not as effective as the worst of those. I think it's implicit to those other therapies, i.e., to whichever would be the best of those, or at least generally would be argued to be among the best of

1 those.

2 Does that mean that, if you're underway with

3 a trial and something else evolves, then you might

4 have to start again? That's not unprecedented.

5 I go back to 1987, when in the colon adjuvant

6 setting when 5FU leucovorin was being compared

7 against placebo. And two years later, in 1989, FDA

8 approved 5FU levamisole, having been shown to be

9 clearly superior to placebo. Hence, the 5FU

10 leucovorin trial stopped in midstream, saying

11 that's not the question anymore. The question now

12 is, 5FU leucovorin being non-inferior to 5FU

13 levamisole. That's what happened.

14 My sense here is, do we have something

15 that's clearly superior to warfarin? Obviously,

16 what comes to mind is the RE-LY trial with

17 dabigatran. There are some negatives that the

18 sponsor and others have correctly pointed out about

19 that trial. It's an open-label trial, and the

20 sponsor, to their credit, did a placebo-controlled

21 trial.

22 The dabigatran had unfavorable trends in MI,

and here we had positive trends in MI. But the primary endpoint, the primary stroke-based endpoint, has a hazard ratio of .65 by ITT analysis with four standard errors, ruling out a quality on an ITT analysis, pretty robustly superior. I think the agency approved it with the label of being superior.

Is that correct? No? Just non-inferior? Well, to go on -- they can come back and ask that.

DR. TEMPLE: Well, it shows -- I mean, I don't know if superior is in the indications, but if you read the description of the trials, it shows a P less than whatever and shows an upper bound that's well below 1.

DR. FLEMING: So, at least then to make the comment that it was four standard errors in a superiority assessment, it was stated by the agency that when the TTR subgroup greater than 67 percent was assessed, that it was robustly favorable on non-inferiority in those centers. It didn't have the interaction by prior stroke. There were 6,500 U.S. patients to give you much more robustness in

the U.S. There were two active arms.

So, collectively, at least I ask the question, is there sufficient data to say dabigatran is superior? And, therefore, if we would say it is, then I'm uncomfortable by stating that the policy is addressed simply by showing that you are non-inferior to something that is inferior to an existing therapy.

The last part of the question is, well, is this agent, then -- if one would say that we would have to not only be non-inferior to warfarin, but non-inferior to dabigatran, is it? I don't know. It might be, but we would need more data to answer that question.

DR. LINCOFF: Dr. Sager?

DR. SAGER: Tom, I just wanted to ask you, there are also differences in the patient populations between RE-LY and ROCKET. Does that concern you at all, in terms of thinking about the superiority issue?

DR. FLEMING: A lot of things concern me, and that's a very relevant concern. I come back to

what Dr. Emerson said earlier, though. In essence, we're labeling this for a given indication. And it seems to me, the policy should be, if you're going to use it in the same indication as another therapy that is superior to warfarin, then why is it not necessary to know that you are not giving up efficacy, meaningful efficacy, against that other agent with the same indication?

DR. LINCOFF: But there is room for the true scientific debate that a drug proven in a less sick population to have a degree of superiority may not have been superior in a sicker group of patients, even ostensibly with the same indication.

I mean, I guess if we had unequivocal evidence of superiority, that would be one thing, but we have one trial with, as you point out, some flaws, and in a somewhat different population. So I think, in this specific case, not as an overall policy --

DR. FLEMING: If one were to say, even -- to correct what I'm saying. If one were to say, well, even if dabigatran has the broad label, quite

frankly, it's really not clear what its true efficacy is against warfarin in a sicker population, and this agent is non-inferior in a sicker population, then I think that is a relevant consideration.

DR. SAGER: I think I would agree with you, Tom. I think that makes a lot of sense.

DR. LINCOFF: Dr. Nissen?

DR. NISSEN: This is more of a question for everybody, but as I recall the data, warfarin's efficacy is greater amongst sicker patients. That is, the sicker you are, the greater the relative reduction in stroke risk. And somebody correct me if I'm wrong about that.

Is that right or wrong?

DR. LINCOFF: I think the EAFT was similar, which was the highest-risk trial to all the others, or relative risk.

DR. NISSEN: So either way, I guess the question, then, would be, is it likely to be the case here that the degree of risk is a modifier of the efficacy of the drug? And if it's flat, that's

fine. If sicker patients get more effect from warfarin, that's fine. The issue is, do sicker patients tend to get less benefit from warfarin? And I don't think we have any evidence to suggest that they do. So, to me, it becomes a non-issue.

DR. LINCOFF: Dr. McGuire?

DR. MCGUIRE: Yes. Just a few points. I agree completely with Dr. Fleming's comments on an academic level, but from a practical level, I completely agree with Dr. Fox's. It gets to the point where it would be impossible to continue to execute these large-scale randomized clinical trials without having to stop and start, and stop, and start, as new agents come along. We see, in this space, there are four agents within a 12-month period of time that we're going to be dealing with. And so I think that just raises a bar that's impractical to achieve.

I think it benefits patients to have competitive therapies. We have lots of examples where some drugs within a class for a given indication are slightly better or slightly worse,

1 and there are a lot of reasons why some may be
2 preferred for given patients, side effect profiles,
3 tolerability, third-party payer decisions, other
4 issues.
5 So for those reasons, I don't necessarily
6 think we have to continue to chase the tail of the
7 efficacy as long as it's defensible at the
8 execution of the trial, that it's a relevant
9 comparator to begin with.
10 Then there's also the issue, the long-term
11 track record of dabigatran isn't proven yet, we've
12 seen. It wouldn't be the first drug that comes
13 into problems that were unforeseen, maybe down the
14 road. There's no suggestion that's going to
15 happen. But, again, to have extra in our arsenal
16 is not adverse.
17 I would just make a final comment. We have
18 aspirin, and aspirin plus clopidogrel, or
19 clopidogrel available for this indication, for
20 patients who choose not, are unwilling, or are
21 unable to take warfarin. So this certainly would
22 be better than those options.

DR. LINCOFF: Dr. Kaul?

DR. KAUL: What does as effective mean operationally? I really don't know, but this is excerpted from the non-inferiority guidance document. "In most cases, a successful non-inferiority study supports effectiveness of the test drug, but it only rarely will support a conclusion that the drug is equivalent or similar to the active control," which is what I imply as effective as to mean, a concept that has not been well defined for these. "Such similarity might be concluded, however, if the point estimate of the test drug favored it over the control, and the upper bound of the 95 percent confidence interval was close to showing superiority." This is what the statement says in the guidance document.

Now, if you use that as the operational standard for some analyses, it meets that operational standard. But for other analyses, which in my opinion are more robust than the on-treatment plus 2 days or plus 30 days, it does not.

So is it sufficient to be as effective as warfarin? Yes, if that's the standard of proof. If so, is it? Using that operational standard, it is not.

C, is it necessary to be as effective as something else? Not necessary, but it's a shifting target. Maybe in the future, it might be necessary to be as effective as another drug.

DR. LINCOFF: We are going to move onto the next question, 6.

DR. TEMPLE: Can I just say something about equivalence? Equivalence is a poorly defined term. We use it for bioequivalence, by which we mean between 80 and 125 percent. Most people would say that's okay.

When you try to translate that into a clinical endpoint, it gets very, very difficult, and that's what that document says. It says if you almost make it, maybe we'll buy the idea that you're equivalent, if you're almost superior. But it's very hard.

We wanted to remind everybody that just

because you are non-inferior doesn't mean you can make the case that you're equivalent. We don't know that, usually, because the confidence interval is larger than what we have historically called equivalence, which is 80 to 125. We're facing this, in a major way, on the whole biosimilar thing, I can tell you.

DR. LINCOFF: Question number 6. Are there adequate instructions for use with regard to A, what regimen to use in most patients? If not, does this matter? B, what dose adjustments are needed in patients at extremes of exposure or risk? If not, does this matter? C, transitioning between rivaroxaban and other anticoagulant therapy? If not, does this matter? And D, actions to take in the event of serious bleeding? If not, does this matter?

In this regard, I've been told that the sponsor would like to clarify one point regarding the discontinuations. And I think considering there was a bit of discussion around it and relates to item C, if they can do that quickly, we could do

that now.

DR. CALIFF: We got a bit concerned when the treatment interruptions were being discussed, that some people may have been confusing what happened at the end of the study from what happened in the midst of the study. We had over 7,000 treatment interruptions in the midst of the study, with no difference in the event rates in either group, in that, during those interruption periods.

So I think -- the concern that was raised about, well, people will forget a dose or stop a drug, in fact, in the first two days after a treatment interruption, the event rates favored rivaroxaban over warfarin, which is a little counterintuitive for the reasons that were given.

So I would just go back to the main slide that we showed, if we look at rows 3 and 4, just to repeat this, early discontinuation and temporary interruptions, you can see that there were a lot of interruptions, over 3,500 in each group, and you had a very small number of events with no significant difference between the two.

I just wanted to clarify that. There's no argument about the end of the study being a transition problem.

DR. LINCOFF: Thank you.

So in your answers, again, please address as many of these items as you would like.

Takers? Dr. Nissen?

DR. NISSEN: I don't know how to transition the drug, based upon what we've got from the ROCKET trial. I just don't think we have any information. And so the question is, there will be people, for a variety of reasons, as everyone's pointed out, who may not tolerate a drug. Who knows what they have? They have a skin rash or they have GI intolerance. And we need to know what to do. And, unfortunately, we're not informed about that.

I do think that's an issue. In the absence of data, physicians will do what their gut tells them to do, and that may be right or that may be wrong. And so they may overlap therapies. Who knows what they'll use. Will they use enoxaparin? Will they use heparin? Will they start warfarin

and have some overlap period? None of that has been answered here, and so we just don't have the answer.

DR. LINCOFF: I'd like to make a comment in this regard as well. When we sat here, maybe a year ago, around the RE-LY panel, there was a lot of criticism of the open-label design, but one of the purported advantages was that they could adjust the discontinuations specifically for the two drugs because that was open label, and that did inform us.

Now, we have a trial that was done as many -- in the more pure way, wanted to see these kind of trials done, as a blinded trial, but as a necessity. Then one couldn't adjust what they did afterward, at least in a transparent fashion, based upon what their treatment was.

So I agree that we don't know what to do, and I think we probably need more data. And I think that in the FDA's booklet, there were some suggestions on smaller trials that might be quite reasonable to provide some information. But I

1 think probably one could have predicted ahead of
2 time that we wouldn't be able to tell, from a trial
3 such as this, the optimal way. I don't think we
4 could have predicted that the outcomes would have
5 been as bad as they were, and I think that's a
6 point of concern, but I don't think that speaks to
7 something we could have gained from this particular
8 trial.
9 With regard to the other points, I
10 think -- although, again, if the dose wasn't
11 optimal, BID may have been optimal, we know what we
12 know from this trial, which was the QD dosing, and
13 I think that's appropriate.
14 In terms of actions with serious bleeding, I
15 think that's a problem with most of the
16 anticoagulants that we have. In virtually none of
17 them, the anti-platelets or anticoagulants have
18 specific reversal agents. And we end up treating
19 them all about the same way, with some
20 consideration of exactly where the hemostatic
21 defect is. But I think that most people who use
22 these drugs are fairly accustomed to dealing with

non-specific means of dealing with bleeding.

Dr. McGuire?

DR. MCGUIRE: I'll agree both with Drs. Lincoff and Nissen. The reality is, though, I don't know that we have instructions, adequate or otherwise, for transitioning of any anticoagulant to any other presently approved. We have no idea how to transition heparin, low molecular weight heparin, warfarin. We all make our best guesses, and we think we're probably doing it the right way, but I don't think this absence of data distinguishes this drug from all others previously approved.

I do think it would be certainly important to pursue the proposed strategy, put in a PK/PD trial, ideally in a smaller trial. And certainly to point D, it's imperative that we, for all of these agents, explore how to reverse their bleeding propensity for emergency, trauma, and et cetera.

DR. LINCOFF: Dr. Nissen?

DR. NISSEN: Just one follow-up comment. I agree with you. The thing that makes this more

poignant is the fact that such bad things happened in the first week or two after the drug was discontinued. So it makes the imperative different. And, yes, I don't know how to go from enoxaparin to heparin and back again. But we don't have, at least any evident data that -- well, we have some data there, but not a lot that suggests that it's really going to produce serious problems.

Here, we've got what's in front of us. And what's in front of us says that we can get into a lot of trouble when we go off drug here, and we don't know what to do, and that makes it a problem.

DR. LINCOFF: Dr. Fleming, did you have a comment? Oh, okay.

Yes?

DR. KINDZELSKI: Just a comment. Just wanted to support what Dr. Nissen said. We don't really know how to bridge for practically any anticoagulant drug. And currently, there are a couple of clinical trials that are looking specifically to address these issues.

Again, one of them is for warfarin, as we are

1 considering the gold standard for the
2 anticoagulation.
3 So I don't think that the purpose of the
4 ROCKET was to answer that question. That was found
5 during that clinical trial. That raises a huge
6 issue, large issue. And I believe at this point in
7 time, we can address that to FDA. And if that drug
8 will be approved, there should be some instructions
9 about not termination of the drug, no rapid
10 termination of the drug, and some actions regarding
11 finding the way how to do the bridging or moving
12 out of the clinical dose. Thank you.
13 DR. LINCOFF: Dr. Papademetriou?
14 DR. PAPADEMETRIOU: I, too, wanted to
15 register my concern about the lack of data on the
16 fact that we don't know how to bridge from one type
17 of anticoagulation drug to the other. And we all
18 have examples in cases of patients that we see day
19 in and day out, that go into trouble when we try to
20 transition them from antiplatelet therapy, the
21 other one, catherization, and then we want to get
22 them back on Coumadin.

In an effort to bridge them with low molecular weight heparin or heparin itself, many times we go into issues with bleeding complications, and other times, we have events, actually, during the transitional period. And this is a problem with this particular anticoagulant, that we have seen more events at the end of the study until the effect of Coumadin took over. And it's important to acquire some organized data in the best way to transition up from this anticoagulation to long-term Coumadin.

DR. LINCOFF: Dr. Sager?

DR. SAGER: Yes. Like everyone else, I appreciate the increased event rate after discontinuations. I do believe that with a properly performed, relatively small PD-type of study, that one could develop and test a schema for this, which wouldn't be tested in a large number of patients, wouldn't be able to assess some things, but probably would be sufficient to address these medical needs, which, as we've already said, we actually don't have for many of the other drugs

that we use in this setting.

DR. LINCOFF: Any other comments? I'll warn you that we're coming up on the voting questions, so if you have other comments, this is the time.

All right. Now, we will be using the electronic voting system for this meeting. Each voting member has three voting buttons on your microphone, yes, no, and abstain. What you'll do is you'll vote by pushing the button located immediately below the corresponding letter. Again, firmly push the same button three times. After everyone has completed their vote, the vote will be complete, and the vote will then be displayed on the screen. I will read the vote from the screen into the record.

Next, we will go around the room, and each individual who voted will state their name and their vote into the record, as well as the reason for why they voted as you did. So this is your opportunity to make your comments about why you made your discussion.

So we will now begin the voting process.

Please press the button three times on your microphone that corresponds to your vote. You'll have approximately 20 seconds to vote. Don't start the flash yet, please, because I haven't even read the question. Please press the flashing button firmly. And after you've made your selection, the light will continue to flash. If you're unsure of your vote, please press the corresponding button again.

So the question is, should rivaroxaban be approved for the reduction of stroke and non-CNS systemic embolism in patients with nonvalvular atrial fibrillation? So again, firmly, three times.

[Vote taken.]

DR. LINCOFF: So I'll now read the result into the record. The voting result is yes, 9; no, 2; abstain, 1; and no voting zero.

Now, we'll go around the table, starting please with Dr. Fleming. If you could read your vote into the record and then any comments you'd like to make regarding why you voted as you did.

DR. FLEMING: I abstained. I think there are many strengths in this blinded and very informative, important trial. By the way, I think Dr. Califf's thoughtful presentation and the FDA's reviews were also incredibly thorough and informative, and I'm greatly appreciative to both in helping us to be as informed as possible in addressing these difficult issues.

Rivaroxaban is clearly superior to placebo, and I believe these data do not allow us to conclude that it's superior to warfarin. On the additional key issue, I think it might be possible to justify a conclusion of non-inferiority to warfarin. And, in fact, I think I'm leaning in that direction, but concerns still remain about the influence of TTR levels, the per-protocol on-study versus ITT, particularly in the context of the increasing events that are occurring at transition, the overall total excess events in the wrong direction, and the large fraction of events, particularly deaths, that are occurring after 2 days post, being on study drug and then the

missing data.

Also, I'm perplexed about the proper way to address the dabigatran results, which I believe is an agent now that we have considerable evidence to be superior to warfarin in a clinically related setting. And if that's the case, it leaves me more uncertain about how to interpret the adequacy of showing non-inferiority to warfarin.

My vote a little bit reflects the fact that I don't think an advisory committee -- being advisory, the voting concerns me. The essence of what matters is the great discussion that occurred. I hope that is what the agency will largely be influenced by. But I go beyond that to say, it's difficult to answer yes/no in my case because I see such issues that are still uncertain in my mind on both sides, and I hope the agency will be enlightened by all the discussion, and can make the proper decision.

DR. LINCOFF: Dr. Sager?

DR. SAGER: Having treated thousands of patients with atrial fibrillation and dealt with

the tremendous problems of using warfarin, and patients having issues on warfarin, I think there's a tremendous unmet medical need for new therapies.

I think the executive committee and the sponsor have done a really nice job with this. Being able to do a double-blind study here, though, as we've just discussed, there are some significant issues, such as the TTR, the transition period.

Despite all that, I believe that the robustness of the non-inferiority being met through multiple different types of analyses leaves me with the belief that it is actually non-inferior to warfarin, and, thus, would be a real advantage for patients with atrial fibrillation. There are obviously a number of other things, though, that will have to be discussed and worked out.

In particular, the study was done in high-risk patients, which I think, again, is a group that hasn't been studied in other venues, and I think is a group that actually is at most risk for having major events.

DR. LINCOFF: For the formality, can you

just state your name and your vote into the record?

DR. SAGER: Philip Sager.

DR. LINCOFF: And your vote into the record?

DR. SAGER: Yes.

DR. LINCOFF: Dr. Nissen?

DR. NISSEN: Dr. Nissen. No. Here is a summary of my thinking. First of all, I am worried that we may have the wrong dose administration schedule for the drug, that a drug with a relatively short half-life was developed as a once-a-day drug and that there may be a better way to give this drug. And we're not likely to find that out at this point, if we approve the drug with the current data.

But more importantly, if you look at the evidence for non-inferiority, it is strongest in the per-protocol population, but when you start to look at differing censoring rules, particularly if you go out -- we don't have the intermediate values of, say, going on a couple of weeks as Dr. Kaul suggested. Basically, you end up with point estimates that start to get pretty close to 1.

They start to approach 1.

If you then factor into that the fact that the TTR was well below other studies that we looked at -- I mean, again, they were all pretty much in the mid- to high 60s -- now you introduce a level of uncertainty as to the efficacy of the drug that is substantial.

A question you can ask is, could the TTR have been better? And my conclusion is that it could have been, that there was a fatal flaw in the study design, in that there were not instructions given. And so, if you were cynical, you'd say that warfarin was designed to not be given particularly well in the trial. And I think it was an oversight, not necessarily an intent, but the other trials in this field used algorithms for treatment that allowed them to get the TTR up into the mid-60s.

That wasn't done here, and it left me with a sense of doubt that using the ITT analysis and then further deflating evidence of non-inferiority because of uncertainty about the TTR, now, you get

to the point where I worry.

Again, as I said earlier, I worry about creeping non-inferiority. This has happened in other fields, where a trial comes along and the comparator is not used as well as it was used previously. Now, I worry someone's going to come along and have a TTR of 50 percent, not 55, and they're going to argue, well, we're pretty close to what they did in ROCKET, so maybe we're okay, too.

When you deal with a morbid, mortal complication like this, we owe patients the best therapy that we can get. Now, if we didn't have another agent, if we didn't have another oral agent like dabigatran, maybe I might view it differently. But it does change my thinking, and it makes me concerned that this may be a step back, not a step forward, and that we will come to regret approving a drug without having the right pharmacokinetic, pharmacodynamic properties studied.

Then, finally, the issue of what to do when therapy is stopped or interrupted is a huge approvability issue, from my perspective. I'm

concerned that it hasn't been studied. It could have been studied in parallel. It didn't have to be studied in ROCKET. It could have been studied in another small trial, as has been suggested. And I still think that the agency should not approve this drug without having the data on what to do when you stop the drug.

We have to stop anticoagulation frequently in patients for a variety of reasons, and when we do, we need to have at least some information on what we should do in those patients, particularly when we have evidence that when we stop the drug, there is a rather significant excess of morbid, mortal events. So those are my reasons for voting no.

DR. LINCOFF: Dr. McGuire?

DR. MCGUIRE: I voted yes. The reasons are that I think the study was well designed and well executed. Certainly, by no means was it perfect, and no trials are. I think it's a large, well-represented, across-the-globe, very diverse patient population at a very high risk that was studied

bravely. And I think the results, fairly convincingly, demonstrate the primary objective of the trial.

The TTR, although it deserves the conversation we've had today, I don't think is a fatal flaw of the trial. I think going forward, it will be a point of interest, and perhaps even better modeling of time in therapeutic range or a metric of INR/Coumadin performance could be established.

Like Dr. Sager, I'm quite impressed by the stability of the estimates across a wide array of sensitivity analyses, including different data scopes, different stratifications, and with fairly monotonous point estimates of effect.

Finally, again, because it's a large, pragmatic trial, it is broadly generalizable, not just to patients in the U.S., but around the globe, where, for whatever reason, Coumadin therapy is not well done. This will provide an additional option.

DR. KRANTZ: Mori Krantz, University of Colorado in Denver. I was a cautious yes. I think

what swayed me to vote yes was really the consistency of the results I think across the different populations and the treatment periods, and I think that was really important to me. I think, certainly, it did fit a new need in terms of the higher risk population, so I think that's really important.

I do have a few caveats that I wonder about, potentially pre-approval or post-approval issues, that I think we talk about in number 9, so I'll stop at that point.

DR. LINCOFF: Michael Lincoff. I voted yes. Dr. McGuire as articulated many of the reasons and the points, and I agree completely. Just a few points. I think that it's clearly non-inferior. I think that the opportunity to prove superiority was lost when the decision was made to use the type of warfarin dosing and to raise those questions. And that may be irrelevant over the long term, but I think that it was a decision, it was the appropriate decision -- it was an understandable, defendable decision, but I think it takes away the

opportunity to have the sort of data that dabigatran has regarding superiority.

Nevertheless, I think the comparator was appropriate and that dabigatran doesn't meet a standard, particularly given that this trial tested a much higher population of patients that would have required it to have been the comparator.

I also agree that additional data probably should be obtained before this drug is approved regarding how to transition; not how to discontinue, because I think the treatment discontinuations did not suggest excess risk, but that if you believe that a patient still must be anticoagulated, I think we need to know better how to do that.

DR. KAUL: Sanjay Kaul. I voted yes. Like Dr. Fleming, I don't pay much attention to the little tally count. I think it's the discussion and the deliberation around that is key. And the only reason why I voted yes is because I was heartened to see an option for a claim, claim number D, claim from patients failing other

anticoagulant therapy. And that's one of the reasons why I voted, because I want it to be available for those types of scenarios.

The key question that I still am struggling, and I was struggling all throughout is are the results of ROCKET AF so persuasive that approval will represent no lowering of its standard of proof? I still don't have my arms wrapped around that standard of proof, so I was willing to give it a benefit of doubt, because I'm loathe to making arbitrary scientific assessments in the absence of any objective-basing points.

But I am concerned about one major issue, and that is when you stop the treatment, what happens? We really haven't ruled out whether there is rebound hypercoagulability or whether this was all due to inadequate anticoagulation.

So the data are not definitive enough to allay my concern regarding hypercoagulability, nor are they persuasive enough to reassure me that this was a quirk of the design, that patients were not properly anticoagulated during the transition

phase. So I think that ought to be addressed. Whether that ought to be addressed pre-approval or post-approval is entirely up to the division, but, nonetheless, I think that certainly needs to be addressed.

MR. COUKELL: Allan Coukell. I voted yes. I'm convinced that rivaroxaban is superior to placebo. I think, clearly, the underlying question in the Reinventing Government policy is would patients be harmed if they were started on a new therapy that was approved through this process?

Had there been a third arm to the ROCKET trial that was a dabigatran arm that clearly demonstrated superiority, I think the answer would be quite clear. We don't have that. and I think we don't have the evidence to make a comparative conclusion.

At the same time, the questions that we've discussed about TTR and the transition issues remain real, and I've heard nothing that convinces me that rivaroxaban should be firstline therapy for any patient or any subgroup of patients. And so

while I voted yes, my lingering concern is that once this drug is on the market, it'll be aggressively promoted on the basis of once-daily dosing or some other purported advantage and will displace a drug that has a superiority indication. So I know we're coming to that part of the discussion.

MS. MCCALL: Debra McCall, and I voted yes. All afib patients are very familiar with the three kinds of drugs they have to take. We need a rate med, we need a rhythm med, and we need a blood thinner. We have options for rate, we have options for rhythm, and we have very few options for a blood thinner. Warfarin has been on the market for 70 years. We have the new kid on the block, dabigatran. There are some physicians that will also use aspirin with Plavix. In my experience, it's not offered very often.

I really liked how this study was done because it is an older population. It was a sicker population. These are mostly the people I see on forums having problems, that they're managing

multiple drugs and multiple issues. So it was nice to see a study that was really, to me, targeted towards this population.

I'd like them to have another choice. Not everyone can take warfarin, the interactions, the contraindications. Dabigatran is already having problems, anecdotally, with patients because of the GI upset. So I'm happy to see that there is another option.

Like has already been mentioned, I'm a little concerned with bridging. Patients go off whatever blood thinner they are for the colonoscopy, for dental surgery. How are they protected? How are they covered?

We don't know. That's not done for anything, not even warfarin. In the long run, I hope that's done. That's something that I think clinicians and patients would really like to know. But, overall, I'm still happy with this drug, and I do hope it goes to market.

DR. EMERSON: Scott Emerson, and I voted no. It was not a firmly held opinion. And it was

1 probably somewhere between Tom's and Steve's. But,
2 primarily, this was more of a documentation sort of
3 issue. I felt that there were issues about the
4 bridging that just raised questions. I think,
5 perhaps, some of the data that they have might have
6 allayed my fears, but they didn't have analyses
7 that I would have wanted to see.
8 Then the other issue is just worrying about
9 the creep of indications, and diagnoses, and what
10 you're going to do about the QD dosing when there's
11 some evidence that BID might be better, and the
12 concept of supplanting that that just said -- I
13 would have liked to have seen a little bit more
14 data on that, but I certainly don't think this
15 would be the worst thing that happens in the world,
16 for it to be approved.
17 DR. PAPADEMETRIOU: Vasilios Papademetriou.
18 I voted yes. And I think this area is an area that
19 we need more drugs because the standard of care,
20 Coumadin, is a very difficult drug to use for many
21 patients. The patients have difficulty in staying
22 on the drug and maintaining therapeutic ranges, and

there are a lot of interactions with food and other drugs.

Frequently, almost daily, we see patients that come in and they're subtherapeutic. We need to adjust their medicines frequently -- I mean, their Coumadin frequently; so seeing drugs that have less interactions and less problems. And the data show that -- I believe that's good in preventing embolic and thrombotic strokes. It's a good thing, and it's just going to be good for our patients. And the data I have seen today convinced me that this new drug, rivaroxaban, can prevent embolic strokes and is not inferior to Coumadin.

Although there are questions that I'm concerned about, the bridging, the lack of guidance, and how to go about bridging with heparin, when we are stopping rivaroxaban -- bridging to Coumadin, I think it's very useful to have it on the market. It's going to help a lot of our patients.

DR. KINDZELSKI: Andrei Kindzelski. I voted yes. And regardless that Dr. Rose mentioned that

it doesn't count really, I had a lot of hesitancy to say yes. And this is due to -- I really want to see therapeutic alternatives for the oral anticoagulation in the market. At the same time, there are a few serious concerns that I think should be addressed. One of them is indication, who will be the target, the general public or only patients who really have some serious reasons to go on these drugs?

Second is the dosage. Will it be once a day, twice a day, and what will be the dose?

Third is, of course, transition or bridging. From the trial that was not designed to answer this question, there are hints that it might have more complications than warfarin when it's stopped, or for whatever reason. But, in general, I really -- and I think the American public needs alternative anticoagulants -- oral anticoagulants. Thank you.

DR. LINCOFF: We have 10 minutes left to answer two questions, to discuss two questions. I think these questions are fairly straightforward.

The first, if you voted to approve rivaroxaban, and even those who did not if you'd like to contribute, to prevent strokes in patients with atrial fibrillation, does it merit, A, a superiority claim to warfarin, B, a claim as an effective alternative to warfarin, C, a claim as effective, or D, a claim for patients failing other anticoagulant therapy? And if so, what constitutes failure?

Dr. Kaul?

DR. KAUL: No, for superiority claim. No, for a claim as an effective alternative to warfarin. I don't know what's C -- what's the difference between B and C? That may be technical. But for certainly D, a claim for patients failing other anticoagulant therapy.

What constitutes failure? If patients, for some reason, are not well-anticoagulated on warfarin. They don't want to be on warfarin because of the diet or because of the medications that they are on, or if patients simply refuse to take warfarin because they don't want to be bothered with monitoring anticoagulation, those

1 would be the types of patients where I think
2 rivaroxaban would be an effective alternative to
3 warfarin and other anticoagulants that are
4 approved. So thirdline option..
5 DR. LINCOFF: Dr. Nissen?
6 DR. NISSEN: I would actually be okay with
7 8D, but under very specific circumstances. I can
8 imagine a patient that's had warfarin skin necrosis
9 and gets put on dabigatran, and has severe GI
10 intolerance, which, by the way, is a problem in
11 some patients. Can't take dabigatran, can't take
12 warfarin. This would be a better alternative than
13 to giving that patient aspirin or aspirin and
14 clopidogrel.
15 So I think that would be a reasonable
16 population, pending further evidence of how this
17 drug compares to warfarin when used skillfully.
18 DR. LINCOFF: Dr. Papademetriou?
19 DR. PAPADEMETRIOU: I would vote no for A.
20 I don't think that the data supports superiority
21 compared to warfarin, but I would vote yes for B,
22 C, and D.

DR. LINCOFF: Dr. Sager?

DR. SAGER: No for A, no for B, no for C. For D, I think that's fine. I think the other issue, though, is should the population be those who are at high risk, CHADS score; for example, a CHADS score more than 2? Because that is the focus of the study.

DR. LINCOFF: Dr. McGuire?

DR. MCGUIRE: I would also agree no for A, but it's not clear to me how we can say no for B and C. We've proven in active control trial that it's non-inferior to an excellent comparator. And so it's effective at least against placebo, so it does have efficacy. Whether it's as effective as warfarin isn't the question. The question is, it's an effective alternative? And so I think all three, B, C, and D, should be yes.

DR. LINCOFF: I echo Dr. McGuire exactly.

Dr. Emerson?

DR. EMERSON: I can't go with A or B, but to the extent that C is really just a claim comparing it to placebo, I believe it is more effective than

placebo, and I think D is as acceptable.

DR. LINCOFF: Dr. Sager?

DR. SAGER: Yes. Now, if C is meant to be against placebo, that's fine with me.

DR. LINCOFF: Dr. Fleming?

DR. FLEMING: No for A. I assume C was asking -- can we give it a claim as being as effective as warfarin. I thought what's what you meant. Is that what you did mean?

DR. STOCKBRIDGE: I think B was mostly intended to imply you thought it was as good as warfarin.

DR. FLEMING: So let me answer the question, then. Whether it's B or C, can we say it's as effective as warfarin? I view that that's a proper statement, when you've essentially shown something very close to superiority. So if the agency believes that the upper limit of confidence intervals here are pretty darn close to 1, then I could accept that.

If, however, because of issues around TTR, and the per-protocol censoring, and missing data,

and other irregularities, you're just confident you can rule out 1.38, then I'm comfortable with saying it's as effective as warfarin in that setting. I would be more comfortable with saying it's established to preserve an adequate fraction of the effect of warfarin, if you're only saying you're ruling out 1.38.

Under part D, while it makes more sense to me, as Sanjay was saying, that its use would be in second- or thirdline, one of the concerns that I have, being an evidence-based guy, is that we didn't study it as what do you do when you can't take warfarin, or dabigatran, or you failed to know that it does something there? So I'm torn because I'd rather see it used in that setting, but I don't have direct data to show its effective in that setting.

DR. LINCOFF: Dr. Kaul?

DR. KAUL: I just wanted to stand corrected for option C, which I voted no, but if it is an efficacy claim against placebo, then the answer is yes.

DR. LINCOFF: Let's move onto the last. If rivaroxaban were to be approved for stroke prevention in patients with atrial fibrillation, A, are there any constraints you would place on the population in whom it would be indicated; and, B, are there any issues you would want to resolve post-marketing? Dr. Krantz?

DR. KRANTZ: I'll take a stab at that. Mori Krantz from Denver.

I think the end-stage liver disease, end-stage renal disease I think would fit pretty clearly. I still have this concern about extremes of body weight with a single, fixed dose, and I really think that might be something to consider.

I guess the other one is those that are already on well-controlled warfarin, because I don't think, beyond a shadow of a doubt, we've proven that warfarin, under skillful or excellent control, is not potentially still a better alternative. So I think I might include that.

To me, I felt the idea of having a pre-approval study, looking at the PK/PD relationship

in terms of number of letter B, would be important. And I would make it bi-directional. I wouldn't just look at going from the rivaroxaban to warfarin, but also going from warfarin to rivaroxaban, so you could get an understanding of bleeding as well as thrombotic risk.

Sorry. I put you guys asleep.

DR. LINCOFF: Is that a hint?

Dr. Nissen?

DR. NISSEN: Yes. I would really like to see this drug tested in this atrial fibrillation population BID, because it makes a lot more sense from a pharmacokinetic, pharmacodynamic point of view. And maybe this drug will have more -- at point estimates and confidence intervals and look more like dabigatran if it's studied in a more optimal dosage regimen.

Actually, we're all sitting here saying, yes, maybe it's better than warfarin, maybe it's not; it's certainly better than placebo. But that kind of robustness that Tom has been looking for here, you might just find if you give the drug in a

more optimal way. And if you do such a study, please, get a TTR up in the mid-60s.

DR. LINCOFF: Dr. Papademetriou?

DR. PAPADEMETRIOU: I would like to see the different bridging regimens to be tested and some guidance from industry on what is the best way to bridge from rivaroxaban to Coumadin.

DR. LINCOFF: Pre- or post-marketing.

DR. PAPADEMETRIOU: Post-marketing.

DR. LINCOFF: Dr. Sager?

DR. SAGER: I don't think that this increase in events that happened after discontinuation represents a hypercoagulable state, but I think still doing some type of study just to exclude that would be beneficial. Again, I think it's very low yield, but I think it is kind of an important checkbox.

DR. LINCOFF: Dr. Fox?

DR. FOX: I just throw my vote in with Vasilios and Philip about those two issues should be further investigated post-marketing.

DR. LINCOFF: So this is not a question, but

just a point of clarification for me. Is there anybody here who thinks that the transitioning of anticoagulation to warfarin needs to be studied before marketing for this, or is everyone willing to accept the sponsor's plan, and then validate that at some point post-marketing? Is there anyone who would say that needed to be done pre-market?

Dr. Krantz, Dr. Emerson?

[No response.]

DR. LINCOFF: Okay, all right, just for my curiosity.

Are there any other further comments, then? Dr. Temple or Dr. Stockbridge, do you have any?

DR. TEMPLE: No comments. We don't care about votes either, but we thought the discussion was extremely good, sophisticated, and well informed, and we are very grateful for the discussion. We really do care about votes a little. I was just kidding.

[Laughter.]

Adjournment

DR. LINCOFF: I want to thank the panel for

a very stimulating discussion, and the sponsor, and the FDA for their excellent presentations. Thank you.

(Whereupon, at 4:57 p.m., the meeting was adjourned.)