AT

DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

CARDIOVASCULAR AND RENAL DRUGS ADVISORY COMMITTEE (CRDAC) IN JOINT SESSION WITH THE DRUG SAFETY AND RISK MANAGEMENT ADVISORY COMMITTEE (DSARM)

Tuesday, September 11, 2007

8:00 a.m.

Gaithersburg Hilton
Gaithersburg, Maryland

PAPER MILL REPORTING
Email: atoigo1@verizon.net
(301) 495-5831

P A R T I C I P A N T S

Mimi T. Phan, Pharm.D., R.Ph.
Acting Designated Federal Official, CRDAC

Drug Safety and Risk Management Drugs Advisory Committee Members (Voting):

Richard Platt, M.D., M.Sc., Chair
Sean P. Hennessy, Pharm.D., Ph.D.
Judith M. Kramer, M.D., M.S.
Timothy S. Lesar, Pharm.D.

Cardiovascular and Renal Drugs Advisory Committee Members (Voting):

Steven D. Findlay, M.P.H.
Frederick J. Kaskel, M.D., Ph.D.
Michael A. Lincoff, M.D.
John R. Teerlink, M.D.
Michael A. Lincoff, M.D., F.A.C.C

Temporary Voting Members:

Henry R. Black, M.D.
Alfred Cheung, M.D.
Stephanie Y. Crawford, Ph.D.
Lawrence G. Hunsicker, M.D.
Jeffrey Kopp, M.D.
Andrew Narva, M.D.
Ruth S. Day, Ph.D.
Chester B. Good, M.D.
James D. Neaton, Ph.D.
Lewis S. Nelson, M.D.
Malazia Y. Scott, Patient Representative

P A R T I C I P A N T S (Continued)

Guest Speakers (Non-Voting):

Dennis J. Cotter, M.S.E.
Miguel Hernan, M.D., M.P.H., Sc.M., Dr.P.H.
Ajay K. Singh, M.D.
Yi Zhang, D.D.S., M.S.

FDA Participants (Non-Voting):

John Jenkins, M.D.
Richard Pazdur, M.D.
Rafel Dwaine Rieves, M.D.
Ann-Marie Trentacosti, M.D.
Ellis Unger, M.D.

C O N T E N T S

C O N T E N T S (Continued)

C O N T E N T S (Continued)

Open Public Hearing:

P R O C E E D I N G S

Call to Order

DR. PLATT: As our first order of business, because it is September 11, we are going to take a moment to commemorate the events of September 11th and the losses we sustained and the responses to it. Could I ask you all to rise for just a moment, please?

[Moment of Silence]

DR. PLATT: My name is Richard Platt and I am the chairman of the Drug Safety and Risk Management Advisory Committee. We will introduce the full committee in a moment but, before that, Dr. Phan will give us some rules of the OND.

DR. PHAN: 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 at today's meeting will be a fair and open forum for discussion on these issues and that individuals can express their view without interruption. Thus, as a gentle reminder, individuals will be allowed to speak into the

record only if recognized by the chair. 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 any conversation about today's topic takes place in the open forum of the meeting and not during breaks or lunch.

We are also aware that members of the media are anxious to speak with FDA about this proceeding, however, like the advisory committee members, FDA will refrain from discussing the details of this meeting with the media until its conclusion. For the convenience of the media representatives, I would like to identify the FDA contact. Miss Riley and Mr. Kelly, if you are present could you please stand?

Finally, I would like to remind everyone present to, please, silence your cell phones or pagers if you haven't already done so. We look forward to an interesting and productive meeting. Thank you for your participation and cooperation.

Introduction of Committee

DR. PLATT: Now it is time for the committee to introduce itself. Why don't we start to the far right, please?

DR. NARVA: I am Andrew Narva, from the NIH.

MS. SCOTT: I am Malazia Scott. I am the patient rep.

DR. HENNESSY: Good morning. I am Sean Hennessy. I am an epidemiologist at the University of Pennsylvania.

DR. NELSON: Lewis Nelson. I am an emergency physician and medical toxicologist from New York University.

DR. CHEUNG: Alfred Cheung, nephrologist at the University of Utah.

DR. BLACK: I am Henry Black. I am a clinical epidemiologist and nephrologist, New York University.

DR. KRAMER: Judith Kramer, internal medicine from Duke University, with a background in clinical trials and drug safety.

DR. NEATON: Jim Neaton, a bisotatistician from University of Minnesota.

DR. TEERLINK: John Teerlink, cardiologist from San Francisco VA Medical Center and University of California, San Francisco.

DR. LESAR: Timothy Lesar, Director of Pharmacy, Albany Medical Center, Albany, New York.

DR. KASKEL: Rick Kaskel, pediatric nephrologist, Albert Einstein College of Medicine in New York.

DR. PHAN: Mimi Phan, designated federal official.

DR. PLATT: Richard Platt. I am a pharmacoepidemiologist at Harvard Medical School at Harvard Pilgrim Health Care.

DR. LINCOFF: Michael Lincoff. I am an interventional cardiologist at the Cleveland Clinic.

DR. CRAWFORD: Good morning. Stephanie Crawford, University of Illinois at Chicago College of Pharmacy.

DR. HUNSICKER: Larry Hunsicker, a kidney

doctor from the University of Iowa, with a background in clinical trials and large database analysis.

DR. GOOD: Bernie Good. I am an internist from the VA Medical Center in Pittsburgh and the University of Pittsburgh. I have been interested in drug safety.

DR. FINDLAY: I am Steve Findlay, from Consumer's Union. I am the consumer representative of the cardiovascular and renal disease committee.

DR. DAY: Ruth Day, director of the medical cognition laboratory at Duke University.

DR. TRENTACOSTI: Ann-Marie Trentacosti, reviewer, Study Endpoints and Labeling, the FDA.

DR. RIEVES: Hi, there. I am Dwaine Rieves from the Division of Medical Imaging and Hematology Products.

DR. UNGER: Ellis Unger from the Office of Surveillance and Epidemiology.

DR. PAZDUR: Richard Pazdur, Office Director, Oncology Drug Products, FDA.

DR. JENKINS: I am John Jenkins. I am the

director of the Office of New Drugs at FDA.

DR. PLATT: Could we just circle around and fill in the gap, please?

DR. KOPP: My name is Jeffrey Kopp. I am a nephrologist at the NIH Clinical Center.

DR. PLATT: Terrific, thanks so much. Dr. Phan, will you start us with the conflict of interest statements?

Conflict of Interest Statement

DR. PHAN: Thank you. The conflict of interest statement for the joint meeting of the Cardiovascular and Renal Drugs Advisory Committee and the Drug Safety and Risk Management Advisory Committee today, September 11, 2007: The following announcement addresses the issue of the conflict of interest and is made a part of the record to preclude even the appearance of such at this meeting.

Based on the submitted agenda and all financial interests reported by the committee participants, it has been determined that all interests in firms regulated by the Center for Drug

Evaluation and Research present no potential for a conflict of interest, with the following exceptions: In accordance with 18 USC Section 208(b)(3), full waivers have been granted for the following participants, Dr. James Neaton has been granted a waiver for his unrelated data safety monitoring board activity for an affected firm. Dr. Neaton receives less than $10,001 per year.

Dr. John Teerlink has been granted a waiver for being a blinded endpoint reviewer on an unrelated issue for an affected firm and for consulting on unrelated issues for an affected firm. He receives between $10,001 to $50,000 per year for serving as an endpoint reviewer, and less than $10,000 per year for consulting. In addition, Dr. Teerlink has been granted waivers in accordance with 18 USC 208(b)(3) and 21 USC 355(n)(4) for owning stock in a healthcare section fund valued between $50,000 and $100,000.

Dr. Frederick Kaskel has been granted waivers in accordance with 18 USC (208)(b)(3) and

21 USC 355(n)(4) for owning stock in two affected firms, worth between $5,001 to $25,000 per firm.

Waiver documents are available at the FDA's docket website. Specific instructions as to how to access the web page are available outside today's meeting room at the FDA information table. In addition, copies of all waivers can be obtained by submitting a written request to the agency's Freedom of Information Office, Room 12A-30 of the Parklawn Building.

With respect to FDA's invited guest speakers, there are reported interests which we believe should be made public to allow the participants to objectively evaluate their comments. Dr. Ajay Singh would like to acknowledge that he has grants, consulting and speaking relationships with Amgen, Ortho Biotech, Roche, and Johnson & Johnson.

Mr. Dennis Cotter would like to acknowledge that he has a grant with NIH on important outcomes.

Dr. Yi Zhang would like to acknowledge

that she has a grant with NIH on epoetin outcomes.

In the event that the discussion involves any other product or firm not already on the agenda for which an FDA participant has a financial interest, the participants are aware of the need to exclude themselves from the discussion and their exclusion will be noted for the record.

With respect to all other participants, we ask that in the interest of fairness they address any current or previous financial involvement with any firms whose product they may wish to comment upon.

We also would like to announce that Dr. Annette Stemhagen, the industry representative to the Drug Safety and Risk Management Advisory Committee, cancelled her participation very recently and, unfortunately, this did not allow enough time to arrange for a substitute industry representative.

DR. PLATT: Thank you. The only thing that is certain about today's discussion is that it will be a very full one. So, my principal goal will be

to ensure that we have at least the allotted time for each of the segments of the agenda and that we conclude on time. So, we are ready to begin with Dr. Rieves.

Introduction

DR. RIEVES: Good morning.

[Slide]

My name is Dwaine Rieves. I am the Acting Director in the Division of Medical Imaging and Hematology Products. And I would like to thank you for meeting with us today to discuss the use of erythropoiesis-stimulating agents, or ESAs, in the treatment of the anemia associated with chronic renal failure. The next few slides introduce us to the class of ESA products and the topics for today's discussion.

[Slide]

ESAs are a biotechnology-derived form of erythropoietin, the body's naturally occurring protein that is important in the production of red blood cells. Erythropoietin is produced predominantly in the kidney and the loss of kidney

function that occurs in chronic renal failure is commonly associated with decreased erythropoietin production and, consequently, anemia.

Two ESAs are marketed in the United States. Epoetin alfa is marketed under two proprietary names. Epogen is marketed by Amgen for use in anemic patients undergoing dialysis, and Procrit is marketed by Ortho Johnson & Johnson for all other indications. Darbepoetin alfa, the other ESA, is marketed by Amgen under the proprietary name Aranesp. Both ESAs are currently indicated for use in the treatment of anemia associated with chronic renal failure, including patients on dialysis and not on dialysis. The products have other indications, including most notably use among certain patients with cancer who have chemotherapy-induced anemia.

As you probably know, the oncologic indication for ESAs was discussed at an Oncologic Drugs Advisory Committee earlier this year. That committee made several recommendations for revision of the products labels. And, over the past several

weeks Amgen and the Food and Drug Administration have been working to finalize label revisions related to the oncologic indication. Today's discussion which pertains only to the use of ESAs among patients with anemia due to chronic renal failure is a continuation of FDA's and Amgen's ongoing review of ESA safety.

[Slide]

Today's discussion was prompted in large part by two major clinical studies. The first study, which is frequently referred to as the Normal Hematocrit study, was conducted among patients who were undergoing hemodialysis and who also had clinical evidence of heart disease. The second study, the CHOIR study, was conducted among patients who were not undergoing dialysis and included patients without clinical evidence of heart disease.

In both studies anemic patients were randomized to treatment with epoetin alfa, targeted to either a higher or a lower hemoglobin concentration. Remarkably, both studies showed

increased cardiovascular risk for patients in the higher target hemoglobin groups, the groups which also received the higher ESA dosages. These findings illustrate the complexity of deciphering whether the increased risk for these serious reactions was related to the ESA dose itself, the hemoglobin response to that dose, or other factors. Today's discussion will explore these dose and hemoglobin response considerations with the goal of trying to optimize the dosage recommendations within the product labels, as well as design features for any future clinical studies.

[Slide]

ESA dose optimization relates not only to safety but also efficacy. As noted in the first bullet, the benefit supporting approval of these products was the demonstration that ESAs increase the blood hemoglobin to a level sufficient to avoid the need for red blood cell transfusion. This important benefit was robustly demonstrated for both ESAs. Of note, one of the two ESA product labels, the epoetin alfa label, describes other

benefits, a series of patient-reported and physician-assessed outcomes.

The clinical science relating to these types of benefits has importantly changed since the epoetin alfa label was developed in 1989. Last year FDA published a draft guidance to assist sponsor in obtaining the specific types of data necessary to support claims related to patient-reported outcomes. In light of current data expectations for patient-reported and physician-assessed outcomes, this benefit claim in the epoetin alfa label is undergoing reevaluation by the FDA and Amgen. And, for the purposes of today's discussion, we hope to focus upon the increase in hemoglobin and the avoidance of transfusion as the demonstrated benefit of ESAs, a benefit common to both products.

During the first portion fo the FDA presentation later this morning, we will summarize some of the deficiencies detected in FDA's preliminary review of the patient-reported and physician-assessed data available for epoetin alfa.

[Slide]

Today we have two specific topics for discussion. The first relates to identification of the hemoglobin goal when using ESAs. The dosing recommendations within the current product labels do not identify a specific hemoglobin target level. Instead, the labels note that prescribers should use the lowest ESA dose that will gradually increase the hemoglobin concentration to the lowest level sufficient to avoid the need for red blood cell transfusion, and that the achieved hemoglobin should not exceed 12 g/dL. Hence, the determination of a target hemoglobin involves physician discretion and some subjectivity.

The labels lack of identification of a specific hemoglobin target has been regarded as an important deficiency by some clinicians. By the end of the day, we are hopeful that we will have obtained useful advice regarding the identification of a specific hemoglobin target, or target range, or perhaps a maximum target level such that the product labels provide more useful directions. Our

discussions may emphasize the need for additional clinical studies. However, we believe the safety concerns from the Normal Hematocrit in CHOIR studies necessitate that we optimize the dosing information within the product labels now based upon the available data.

[Slide]

The second topic relates to the hemoglobin response to ESAs, and specifically the identification and management of patients who have been referred to as ESA hypo-responders, that is, patients who fail to achieve target hemoglobin levels with administration of usual ESA doses.

Data from the Normal Hematocrit and CHOIR studies suggest that this subset of anemic patients may be especially vulnerable to adverse cardiovascular reactions. Of note, the current ESA labels include a section that lists the various causes for a lack or loss of response to ESAs such as iron deficiency, anemia or underlying inflammatory conditions. However, the labels do not explicitly address how to dose these patients

with ESAs once the underlying conditions are treated.

As our second topic today, we are requesting advice based upon the available data which will, hopefully, improve the product label information pertaining to ESA hypo-responders. Our questions for later today contain proposals that we offer as examples or pivots for a dialogue that will address the sufficiency of these proposals, or result in alternative labeling proposals. To reiterate, I emphasize the importance of the existing clinical data, especially the data from the Normal Hematocrit and the CHOIR studies since these were the only studies that included sufficient numbers of patients to detect important cardiovascular risk.

[Slide]

In our discussions we anticipate hearing about clinical data from two major types of clinical studies, prospective, randomized clinical studies and observational studies. In this context, I use the term observational study to

refer to the studies in which patients are not prospectively randomized to treatment assignments.

In examining drug effect, it is important to remember that the randomization design feature is especially critical to the establishment of drug effects. That is, these types of data provide the most robust evidence that a drug causes a specific outcome. On the other hand, observational data are generally recognized as detecting associations of outcomes with a drug's use, outcomes that may or may not be causally related to the drug. Consequently, when grading the quality of evidence or data to support conclusions regarding an intervention such as a drug effect, the importance of distinguishing observational data from randomized controlled data is exemplified by this quote from the National Kidney Foundation's developers of clinical practice guidelines, who state that in reviewing data the quality of evidence was high if the evidence consisted of randomized, controlled trials but low if it consisted of observational studies.

[Slide]

With these thoughts in mind, let's move to our agenda for the day. Our agenda was developed to prompt an informed discussion of the data, such that this afternoon we can obtain substantive advice regarding the recommendations for ESA dosing within the product label, as well as the design considerations for subsequent clinical studies.

We have arranged our agenda such that our presentations are all completed this morning. First Dr. Ajay Singh, from the Brigham and Women's Hospital in Boston, will provide an overview of the use of ESAs in the treatment of anemic chronic renal failure patients, as well as briefly highlight some of the findings from the CHOIR study. Subsequently, representatives from the Medical Technology and Practice Pattern Institute will discuss their analyses of certain observational databases. Then presentations will be made by Amgen and Ortho Biotech, followed by an FDA summary. This afternoon will be dedicated to the open public hearing, followed by our discussion

of the posed questions.

Again, we thank you for your attendance and look forward to a productive discussion.

Anemia and Chronic Kidney Disease Update

DR. SINGH: I would like to thank the committee for inviting me to speak, and the chair.

[Slide]

My goal this morning will be to review with you some of the studies and some of the practice challenges in managing anemia in patients with both dialysis and non-dialysis kidney disease.

[Slide]

This is the outline of my presentation. I am going to first review the hemoglobin target studies, both the observational studies and the randomized, controlled data. I am going to discuss some of the challenges of treating patients to a target hemoglobin and then, finally, briefly discuss the management of the EPO or ESA hypo-responsive patient.

[Slide]

If one looks at the spectrum of studies

that have addressed hemoglobin targets, it is fair to say that the studies have focused on both dialysis and non-dialysis chronic kidney patients. These studies, as Dr. Rieves has already alluded to, are both observational and randomized, controlled in their nature, and the studies have varied with regards to the use of application of endpoints, ranging from hard endpoints such as death and/or adjudicated cardiovascular endpoints which are softer, if you will, endpoints such as left ventricular hypertrophy or left ventricular mass index, various measures of heart failure and measures of quality of life. Then, more recently there have been studies that have been conducted, some of which are some of the data presented this morning, that represent post hoc analyses of randomized, controlled studies and, in fact, also an analysis will be presented, I believe, of USRDS data as well.

[Slide]

If you review the observational studies, it is indisputable I think that there seems to be

an associative relationship between poor survival and worse outcome--worse outcome and lower hematocrit. So, in this study from USRDS you can see lower hematocrits of less than 27 percent, 27-30 percent, associated with a higher relative risk of all-cause death and cardiac-related death. As one looks at slices of this patient population where the hematocrits are higher, you can see that they appear to be associated with a survival benefit.

[Slide]

As Dr. Kalantar's group at UCLA has pointed out, this relationship, however, is quite severely confounded so here, again, you can see in the unadjusted analyses where he is relating all-cause mortality on the vertical axis with different levels of hemoglobin on the horizontal axis that lower hemoglobins are associated with an unadjusted all-cause mortality as high as nearly 5 in terms of a hazard ratio, and that there is a much lower hazard for mortality when you associate this with higher hemoglobin levels. The

cardiovascular causes of death mirror this relationship.

However, when you start adjusting for case mix or case mix and a number of inflammatory or co-morbid patient factors you can see that this relationship becomes less robust in terms of the association between lower hemoglobin and outcome, suggesting that really this is quite a substantial confounded endpoint, the hemoglobin with relation to outcome.

However, even after adjustment it is important to point out that there is a 1.8 or so higher hazard for worse outcome, whether it is mortality or cardiovascular risk, in patients with low hemoglobin levels. There seems to be a sweet spot. This may represent a U-shaped relationship or J-shaped relationship between outcome and hemoglobin levels.

[Slide]

Given the limitations of observational data, the fact that this is a confounded relationship, a number of studies have been

published that have pursued a randomized, control design to address the issue of what the optimal hemoglobin is in patients both on dialysis and patients who are not on dialysis. As Dr. Rieves pointed out, the largest study published so far in dialysis patients is the Normal Hematocrit study, published in 1998 in the *New England Journal of Medicine*. A second, smaller study, the Canadian-European study, was published a few years ago. The main difference between these two studies is that this study used hard endpoints such as death and cardiovascular risk, where this study did not and used surrogate endpoints such as left ventricular hypertrophy and improvements in heart failure. I am just going to briefly discuss with you the Normal Hematocrit study.

[Slide]

Now, the Normal Hematocrit study randomized approximately 1,230 patients to either a low hematocrit or a normal hematocrit. The hematocrit goal in the low arm was 30 percent and in the normal arm was 42 percent, although the

achieved hematocrit in that arm was 39 percent. You can see that there was nearly a 3-fold higher utilization of epoetin in the higher arm.

When this study was published the results that were presented indicated an increased rate of death and non-fatal MIs. When they looked at the relative risk of death or MI, the relative risk was 1.3 but the confidence intervals span 1, 0.9 to 1.9.

Since the publication of the Normal Hematocrit study there was speculation as to why the study was stopped and what explained this. I took a quote out of the paper: The study was halted when differences in mortality between the groups were recognized as sufficient to make it very unlikely that continuation of the study would reveal a benefit for the normal hematocrit group and the results were nearing the statistical boundary of a higher mortality rate in the normal hematocrit group. Some have thought that this study was stopped because of the higher rate of vascular thrombosis but, in fact, if you read the

paper, it was stopped because of concerns about risk.

[Slide]

Now, this study presented the results from the interim analysis. They did not present results from the completed study in the *New England Journal*. When you look at the results from the completed study--and this is in the FDA briefing document because the FDA has access to the study reportsB-the higher hematocrit arm actually had 634 patients that completed at the end of the study and the lower hematocrit arm had 631 patients. There was a 32.8 percent rate for primary endpoint deaths in the higher arm and 27.4 in the lower arm, and total deaths were 34.9 percent in the higher arm and 29 percent in the lower arm and that, in fact, did reach statistical significance, with a p value of less than 0.001. Also, there were more non-fatal MIs in this arm than in this arm.

So, one could argue that, in fact, this was not necessarily a negative study or one could argue that it was an inconclusive study, but if the

study results are looked at, at the time the study was completed, in fact, there was a statistically significant higher risk for death for patients randomized to the higher hematocrit, an achieved hematocrit of 39 percent, a target hematocrit of 42 percent.

[Slide]

Now, the largest studies that have been done on target hemoglobin in patients who are not on dialysis with chronic kidney disease are two studies, both published in November of 2006 in the *New England Journal*, the CREATE study which is comprised of 603 subjects and the CHOIR study which is comprised of 1,432 subjects.

[Slide]

The CREATE study was a study that was conducted largely in Europe, although it had centers in North Africa and other parts of the world but not in the United States, and it utilized an erythropoietin agent called epoetin beta, which is indistinguishable in its function, as far as we know, from the epoetin alfa. The primary endpoint

was time to first cardiovascular events and the composite was comprised of eight cardiovascular events that ranged from sudden death to acute heart failure, to peripheral vascular disease complications, and so on and so forth. There were 58 events in the higher hemoglobin arm, which aimed for 13 g, and 47 in the lower hemoglobin arm, with a hazard ratio of 0.78 favoring risk in the higher hemoglobin arm, with a confidence interval of 1.53 to 1.14 and that did not reach statistical significance.

However, it is important to point out two important aspects of the study. The first is that they had powered the study such that they would collect 200 events and, in fact, they were only able to collect 105 events. So, this was an underpowered study because they only had nearly half the events that they originally anticipated. Also, the confidence intervals are quite wide, suggesting that there is the possibility, or we can't exclude the possibility of a type-1 error. So, I think a fair reading of this study would be

that this was not necessarily a negative study but an inconclusive study that potentially suggests a trend towards risk in patients randomized to the higher hemoglobin level.

[Slide]

In the last five to ten years there has been a lot of speculation about whether anemia treatment is associated with either stabilization of left ventricular hypertrophy or regression of left ventricular hypertrophy, and the CREATE study provided some insight into that. They had 451 echoes that were performed. Left ventricular mass index decline in both groups but there was no difference between the two groups, neither at year 1 nor at year 2, and it didn't seem to matter whether you started with LDH at baseline. You could not detect a difference between the two groups that was statistically significant.

It also documented quality of life. Quality of life was detected in an open fashion. This was an open-label study and they used the SF-36, a well-established instrument for

measurement of quality of life. It was better for group 1, the higher hemoglobin arm, than for group 2, although the differences are modest and you can see that there were benefits in this group 1, which is in blue, and there seemed to be actually a negative change over baseline for some of the scores for group 2 or the low hemoglobin arm. But the quality of life at year 2 attenuated for most of the scales and was only present for general health and vitality by the time the study reported out two years of data.

[Slide]

The summary of CREATE then is that there was an increased risk, in my opinion at least, for targeting a higher hemoglobin, with a hazard ratio of 0.78. This was an underpowered study. But the 95 percent confidence intervals were sufficiently wide that you couldn't rule out the possibility of a type-1 error. There were a number of dialysis initiations, although this was a secondary endpoint for a study where the primary endpoint did not reach statistical significance. So, I think you

have to interpret this data with a degree of caution. There was an improvement in quality of life and there was no benefit between the two arms for left ventricular hypertrophy.

[Slide]

CHOIR was a study that was conducted by us. The coordinating center was Duke University and my colleagues who participated in this on the steering committee were Linda Szczech, Donald Reddan, and we had a number of investigators, over 100 centers in the United States, using epoetin alfa, and the study was sponsored by J&J.

Median follow up was for 16 months. Patients were randomized to high target hemoglobin, 13.5 g, versus a lower target hm, 11.3 g. There were 715 subjects in the higher target group, 717 in the lower target group. The study had some limitations, one of which was the fact that there was a significant number of dropouts that occurred in both groups. But when you analyze the dropoutsBand there are published papers now that have addressed this issue, which I am happy to

share with you, but most of the dropout rate was accounted for by renal replacement therapy, and this was censored. The withdrawal for other reasons was 21 percent, which is a relatively small number which was similar to other studies. But, certainly, I think that was a limitation of the study.

[Slide]

The study sponsor, as I indicated was Johnson & Johnson. This was a study that was very, I believe, rigorously adjudicated. There was an adjudication committee. All the primary analyses were performed by Duke. Secondary analyses were confirmed by Duke, and we had full access to the information, and this was a study conducted under the intention-to-treat principle.

[Slide]

The primary composite endpoint of the study was death, myocardial infarction, stroke and CHF hospitalization. We did exclude patients who developed the need for renal replacement therapy at the time of admission for CHF hospitalization

because we wanted to make this endpoint more rigorous since we were concerned that if patients developed the need for dialysis, whether the fluid overload that was associated with that could be interpreted as heart failure hospitalization. So, this was adjudicated and excluded out of the analysis. But when we did sensitivity analyses including this it did not modify the results in any meaningful way.

[Slide]

Baseline characteristicsB-these were generally well balanced sets of groups. Patients were randomized to the high versus the low hemoglobin groups were balanced with regards to hemoglobin level at baseline, renal function at baseline. Estimated GFR was about 17 in both arms, and about half the patients had diabetes.

Now, it is important to point out that this study had more diabetics than the CREATE study and appeared to have slightly more co-morbid factors than the CREATE study, but otherwise was relatively similar to the patient population that

was studied in the CREATE study.

[Slide]

Cardiovascular historyB-there were two differences at baseline that did reach statistical significance, the history of hypertension but not hypertension itself. Control of blood pressure was very similar between the two arms and, in fact, blood pressure fell in both arms during the course of the study. So, this was baseline history of hypertension as reported by patients. And, the CABG rate, coronary artery bypass graft rate, was significantly higher, with a p value of 0.05 in the higher hemoglobin arm. So, there were two differences at baseline but otherwise the groups were quite balanced when you looked at other cardiovascular risk factors.

[Slide]

The epoetin alfa dosing will be addressed in a little bit in more detail by Linda Szczech will is going to be commenting on post hoc analysis of CHOIR in the public discussion phase of this hearing, but dosing that was used was epoetin alfa

10,000 units once a week for three weeks. Then after this period, after three weeks, the epoetin was adjusted based on the hemoglobin response. Epoetin was dosed weekly initially and then, when there was some stability, it was dosed every other week. But the vast majority of patients had weekly dosing. Hemoglobin monitoring was every week.

[Slide]

This is the data with respect to dosing. You can see that in order to get to the higher hemoglobin level, approximately a mean dose of 11,215 units were used in the higher hemoglobin arm, 6,276 units in the lower hemoglobin arm. For the hither hemoglobin arm we did not reach our target. We reached a mean hemoglobin of 12.6 g/dL. That was the achieved mean hemoglobin. But for the lower arm the mean achieved hemoglobin was 11.3, and the 95 percent confidence intervals of that were from 10-12 g/dL approximately. So, that was the range of most of the hemoglobins in this arm.

[Slide]

We tested the difference of this on the composite endpoint that I discussed with you and this is the Kaplan-Meier plot unadjusted analysis. So, this is what we saw. There were 125 composite endpoints. The primary composite endpoint was death, MI, CHF hospitalization and/or stroke, with a hazard ratio of 1.337 and a p value of 0.03. So, 125 composite events in the higher arm and 97 in the lower arm. This, to us, indicated that there was increased risk for targeting patients to a higher hemoglobin where the mean achieved hemoglobin was 12.6 in the higher hemoglobin arm.

[Slide]

Now, we did a per protocol analysis and I am not going to go through that in much detail, other than to point out that we did do it. That excluded 37 patients and in the analysis of 1,395 patients the hazard ratio was not much different. The p value did not change. There were other sensitivity analyses that were in the paper and, unfortunately, time constraints don't allow me to present all of that in detail.

[Slide]

The components of the primary endpoint is shown here. Now, remember, these components of the primary endpoint--while you see the data here, the study was not powered for any of those individual components of the primary endpoint. But the primary endpoint was explained by two components largely, death where the hazard ratio was 1.4 but did not reach statistical significance but there was a higher rate of deaths in patients randomized to the higher arm, and CHF hospitalization patients randomized to the higher arm had a 40 percent higher risk. For non-fatal stroke and non-fatal MI, you can see, there was no meaningful difference and this did not reach statistical significance.

[Slide]

Now, CHOIR did not show that there was an incremental benefit in quality of life. We analyzed this using three instruments, LASA, which is a linear analog scale; KDQ disease questionnaire that uses as part of the central component of the questionnaire SF-36 but is relevant to kidney

disease; and the SF-36.

Time doesn't permit me to show you even all the unpublished data, but truly there was no difference either on longitudinal analysis for any of the domains of quality of life in any of the scales, except for one, role emotional, which was worse in patients randomized to the higher hemoglobin arm but there was no difference. Now, much has been said about quality of life, and this was an open-label study so much as we may or may not want to interpret the CREATE study quality of life, you should know that this was also an open-label study but, nevertheless, we did not see any difference between the two arms.

[Slide]

There was a significant difference in adverse events. So, serious adverse events, any SAEs were significantly higher in those patients in the higher hemoglobin versus those treated in the lower hemoglobin. The p value of 0.02 was largely explained by a higher rate of heart failure in the higher hemoglobin arm versus the lower hemoglobin

arm. When you look at any SAEs associated with epoetin alfa, there were threefold higher SAEs in this arm than in this arm but the total number of events was low. Most of these were related to a thrombotic phenomenon but it did not reach statistical significance, although I think you have to be cautious in applying statistics to SAE data but, nevertheless, you can see that there were more SAEs in this group than in this group.

[Slide]

So, the summary for CHOIR is that there was increased risk with targeting hemoglobins to 13.5 and achieving a hemoglobin of 12.6, with a 34 percent increased risk and a p value of 0.03. Strong trends for death and CHF hospitalization. What I haven't discussed with you is the trend for time to RRT which was higher but didn't reach statistical significance. There was a higher rate of cardiovascular and all-cause hospitalization that did reach statistical significance. There was no incremental quality of life benefit with higher hemoglobin, and the number of SAEs in the higher

hemoglobin arm were 15 percent higher which reached statistical significance, largely explained by heart failure.

Our conclusion at the end of this study was that given the fact that there was increased risk, that there were increased serious adverse events and no quality of life benefit, we did not think that it was justified to treat patients to a higher hemoglobin when you have non-dialysis chronic kidney disease.

[Slide]

Subsequently there was a meta-analysis published, and this meta-analysis was published in February of 2007 in the *Lancet*, and used Cochrane methodology. Basically, when you put all the largest studies together there was a higher risk, a 17 percent higher risk if you treated patients to a hemoglobin above 12 in this meta-analysis. Now, there is another meta-analysis, recently published from NKFK Drueke, where they separate out non-dialysis and dialysis CKD patients and they don't see a difference as pointed out in this

study. So, there are two different meta-analyses and they are methodologically different, and those methodological differences may explain the differences in the findings.

It is important to note that in the Normal Hematocrit study they used the published analysis and not the completed study analysis with respect to risk. So, if you included the completed study that risk may have been higher.

[Slide]

There are several post hoc studies and I am not discuss all of them. In fact, I can't discuss all of them, but I will very briefly discuss one of them. So, there is the Normal Hematocrit study that is part of the Amgen report which I guess can be discussed; the CREATE post hoc analysis, I am not aware of any. There is the USRDS data, which will be presented by Dr. Cotter, and the CHOIR post hoc analysis, Linda Szczech, from Duke, will present this in the public forum but I am going to just very briefly touch on some of the highlights of that in my presentation.

[Slide]

So, in one of the post hoc analyses we looked at association between the epoetin alfa dose and outcomes in CHOIR. The hypothesis here was to examine the association between epoetin alfa dose, achieved hemoglobin and targeted hemoglobin, and outcomes. This will be presented as a late-breaking trial by Linda at the American Society of Nephrology annual meeting later this year.

[Slide]

The methodology used was Kaplan-Meier plots and Cox proportional hazards regression. Outcomes were the primary component endpoint that you are now familiar with, and the use of death. There were two approaches. One was to use time-varying covariates and the second was to use a landmark analysis where you fix a landmark time endpoint, and in this case it was four, six and nine months but the data that we present is for four months and nine months to try and also tease out whether in a post hoc way you can avoid some of

the confounding and see whether there is a relationship.

[Slide]

This is just a distribution of the average dose by treatment arm for CHOIR. You can see that the patients in the hemoglobin 11.3 group generally had average doses which were lower than those in the patients treated in the higher arm, but these were wide distributions and these were non-normal distributions.

[Slide]

When you look at this from the perspective of the first analysis using time-varying covariates, there are a couple of slides that I just want to highlight and, as I said, Linda will present this in more detail. This is the number of subjects. This is in group A, the higher hemoglobin arm, and group B, the lower hemoglobin arm. You can see that most of the patients achieved their hemoglobin of greater than 13 in the high hemoglobin arm and greater than 11 in the low hemoglobin arm.

[Slide]

But when you look at the proportion of subjects in each group experiencing the primary endpoint, you can see that the primary endpoint was higher in the higher hemoglobin arm if you did not achieve your hemoglobin of 13, and you can see that if you achieved your hemoglobin of 13 this was associated with a lower risk in the higher hemoglobin arm compared to not achieving your target hemoglobin. This was true also for group B or the lower hemoglobin arm as well. When you adjust for this in a Cox model this difference is actually statistically significantly different.

DR. PLATT: Dr. Singh, I am sorry to interrupt, will another minute be sufficient?

DR. SING: Yes.

[Slide]

This just shows you very briefly the landmark analysis, and again just shows you four months and nine months, and you can see that there seems to be a raltegravir between dose in both arms.

[Slide]

So, basically just very briefly, I think the observational studies show a relationship but the randomized, controlled studies do not. Randomized, controlled studies suggest increased risk.

[Slide]

I am just going to very briefly talk about the other two because of time limitations.

DR. PLATT: We are just about out of time.

DR. SINGH: Yes, just very briefly and then you can look at the presentation that has been presented. Just very briefly on the target hemoglobin issue, the potential reasons why a dialysis population, despite this data, are achieving hemoglobins of greater than 12 g/dL have multiple factors and I wanted to just very briefly point out, and this is in the presentation, that there is a variety of reasons. One is a genuine concern that we want to prevent blood transfusions in a dialysis population, as well as try and improve quality of life. So, there is some goal to

try and increase average hemoglobins in non-dialysis patients. But other factors are also important, including the ownership characteristics of the dialysis facility; the anemia algorithm used, etc.

And, variability is an inevitable part of management of anemia in the dialysis population. In an analysis we have done using Kaufman's data that was published in the *New England Journal*, even in the most well managed dialysis patients about 50 percent of them had some variability. The variability is very modest but there is variability. That is one important point to be made.

The second important point is algorithms in the dialysis do make a difference in modifying this variability. So, that is the second point.

DR. PLATT: Dr. Singh, I think we really do have to move on.

DR. SINGH: I am just going to summarize. Then, in the final part of it I just want to let you know, in the third part, that management of the

hypo-responsive patient really involves not only managing these patients with ESAs but also managing patients for their iron deficiency, and it is important to recognize that the management of a hypo-responder patient is not simply the management of ESA dose in the hypo-responder patient but also a variety of other factors. Thank you very much for your attention.

DR. PLATT: Thanks so much. Mr. Cotter, please.

Medical Technology and Practice Patterns Institute

Epoetin Outcomes Research

MR. COTTER: Thank you.

[Slide]

I want to thank the committee for inviting us to present our work in progress. My name is Dennis Cotter. I am with the Medical Technology Institute. It is a non-profit 501(c), established in 1986.

[Slide]

Since 1991 we have been looking at factors related to EPO use. To make full disclosure,

between '91 and '98 our work was sponsored by Ortho Biotech, and between '04 and '06 it was sponsored by an NIH grant. Since 1991 we have looked at a variety of issues related to EPO utilization. This has not been done in a vacuum though. We have a technical advisory committee that met periodically over the course of the last 15 years to advise us on methodologies and interpretation of findings. Most of our work has been published in peer-review journals and today we will present some work in progress that has not quite made it to the journals.

Our advisory group is made up of a variety of NIH researchers, nephrologists that are well-known nationally, providers. Recently we expanded our group to include a larger number of providers and other groups that have a stake in the outcomes of our work.

[Slide]

Between '91 and '97 we did what I guess is best termed as an actuarial analysis of EPO utilization although, in doing this, we developed a

variety of papers dealing with access to EPO. Subsequent to that, we investigated EPO practice patterns and published those works. Then, finally, with a grant supported research we got into the causal link between EPO and outcomes.

[Slide]

As you can see on this is a USRDS slide, the distribution of hemoglobins over time has shifted from relatively modest hemoglobins up to relatively high hemoglobins. We think that this is partly due to objectives of both the renal networks that state that they would like to see providers move as many patients as possible over a hematocrit of 33. I believe CMS' objective was 70 percent of patients in excess of a hemoglobin of 11, and some of the networks have an objective as high as 85 percent of their patients with hemoglobin of 11. Consequently, we can see the shift over time to the higher hemoglobin ranges.

[Slide]

What dives this? Perhaps it is partly based on clinical guidelines that were published in

2000. There were several claims made for benefit of higher achieved hematocrit and, as you can see, many of these claims have no longer been included in more recent guidelines that have been published in '06. Most of these claims were based on observational studies that were highly confounded.

[Slide]

Our interest was in the survival claim. In a paper that was published in 2004, that motivated us to apply for an NIH grant, we simply desegregated data that was used by other investigators and demonstrated that independent of achieved hematocrit, the larger your exposure to EPO, the higher your mortality rates. So, if we go back to the last group of patients we can see that the mortality rates have increased.

[Slide]

This was a demonstration more than anything else of the confounding in observational studies, which led us again to apply for an NIH grant. Dr. Hernan, with Harvard School of Public Health, will explain the methodologies we used in

our grant research and Dr. Zhang will present the results, both published and unpublished results, resulting from an application of this methodology.

Dr. Hernan, would you like to take the podium?

Introduction to Causal Modeling

[Slide]

DR. HERNAN: So, as the title of slide may say, this is not going to be a biostatistics class, first, because many of you don't need it and, two, because we don't have the time. I would like to give you the conceptual background to put into context the work that we have done with observational data.

[Slide]

So, going to the basics, we are going to estimate the effect of EPO on different outcomes like hemoglobin level or death and, of course, RCT would be ideal. If we don't have that, the next best thing that we have is an observational study that mimics an RCT.

[Slide]

You all know that there is a problem with

observational studies, that patients who are worse off are more likely to get higher doses of EPO, and that is confounded by indication. I don't have to dwell on that. I know there is a lot of expertise in this room about this. This is, of course, not a problem in intention-to-treat analyses of randomized clinical trials. As an aside, of course, the problem with intention-to-treat analyses of randomized clinical trials is that they only give you a conservative estimate so any ITT estimate from a trial in the presence of non-compliance is an underestimate of the true effect.

[Slide]

But confounding by indication is a little more complex. There are actually two problems. One is the problem that we may have not measured all the indications for treatment, all the things that make the patients who receive high dose or low dose different.

[Slide]

The second problem is that even if we have

measured all the confounding by indication, that confounding may not have been appropriately adjusted for, and that is exactly what we did.

So problem number one is the main problem that has been discussed already here and that is why we tend to distrust observational analyses because we don't know if we have measured all the indications and we don't know if we were able to make the treated and the untreated patients comparable.

In our analysis we have tried to fight this problem, number one, by using all the information that was in the database that we were working with. Whether we have achieved that goal or not is for you to figure out but bear in mind that it is usually not enough to say that observational studies are biased. You have to say why or else it not a scientific statement. So, when you think about all the analyses that we did, please think of variables that we should have adjusted for but we didn't.

[Slide]

Let's talk about the second problem because that is why we are here really. It is what happens when you have a lot of confounding by indication and the confounders are not appropriately adjusted for in the analysis. You might think that that is not possible, if you have the confounder sheets and you add them to your models you are adjusting for, but it is well-known and is proven in theory and in practice that that is not always the case. When the confounders are time-varying confounders and they are affected by prior treatment--in this case they are affected by prior EPO dose--then if you add the confounders to the model you may get bias in any observational study or in any post hoc analysis of a randomized trial in which, rather than using the randomization indicator as the treatment, we use time-varying exposure that happens after randomization.

Fortunately, this problem can be solved by using statistical methods that are not the standard statistical methods. The one that I am going to discuss here, and which I will say something about

here, is inverse probability weighting, or IPW.

[Slide]

So, the good thing about IPW is that it can be used to mimic a randomized clinical trial assuming that we have measured all the confounders. Of course, there is no way out of the problem in number one, but under that assumption we have a method that can adjust for those confounders in the "right" way even if there are time-varying confounders affected by prior treatment.

[Slide]

Essentially, what we do is that for each subject we estimate what is the likelihood for that individual having had the treatment history that he or she actually had, and then we weight the contribution of each subject by the inverse of that. So, you can think of this in a way as the general version of propensity scores, the general version of propensity scoring which we are weighting by advancing of propensity scores.

One of the problem of propensity scores is that the theory is not developed to work with

time-varying treatments, and here we are talking about a time-varying treatment. This is EPO dose that changes with time. You can also think of this as the general version of standardization. Some of you may have heard the term marginal structural models because technically what we are doing when we do this is we are estimating the parameters of marginal structural models.

[Slide]

IPW has been extensively used in some disciplines, especially in HIV/AIDS research, to the point that in some of the last RFAs by NIH it is written that the groups applying must have expertise in IPW. And, the reason why they say that is because IPW has been the only method that has been able to replicate the results from randomized studies using observational data. In HIV/AIDS there are a lot of independent confounders. There is a lot of confounding by indication but also there is a lot of information about those indications-BCD4 count, viral load, etc., and the problem doesn't seem to be as much

the placebo confounding[?] but the way that the confounders are adjusted for, and IPW has shown that we can get the same with other trials, which is the first step to go beyond the trials, of course.

[Slide]

In this setting we used IPW to estimate the effect of different EPO doses on the hematocrit and survival. And, we needed to use this method because, of course, we have a time-dependent confounder. It is the hemoglobin level that affects not only the outcome and the future use of EPO, but also is affected by the previous EPO history for that patient.

So, what you are going to see is essentially our attempt to mimic a randomized trial in which a random sample of the patients have been assigned to different doses. So, you are not going to see what happens in those who get a dose of EPO of 10,000; you are going to see what happens if a random sample of the patients have been assigned to 10,000 units per week or 20,000 units per week, or

whatever. Now I give the podium to Dr. Zhang.

Research Findings

DR. ZHANG: Thank you, Dr. Hernan.

[Slide]

I am Yi Zhang, and I am now going to present our research findings.

[Slide]

In this study we evaluated the effect of EPO dose on population hematocrit response. This research is supported by R01 and 90DK grant, and the paper has been accepted by *Kidney International*.

[Slide]

Between 1991 and 2004, the average administered EPO dose increased more than 300 percent in the U.S., as shown by the right line in this figure. The blue line, here, is for the average hematocrit levels. However, the population dose and hematocrit response has not been evaluated.

[Slide]

The only existing clinical study that

actually examined dose and hematocrit relationship is Phase 2 trials. Phase 2 study did not fully capture the population dose and hematocrit relationship because of strict patient selection criteria, small sample size and the limited dose used. Subsequent studies have primarily used an observational database or administrative databases to evaluate the relationship, and they have shown an inverse relationship between dose and hematocrit, which is clearly the case of treatment by indication.

[Slide]

So, our aim was to estimate EPO dose and population hematocrit response by using advanced statistical techniques to appropriately control for treatment by indication bias. In doing so, we tried to mimic an RCT in which subjects are randomly assigned to different levels of EPO doses. Then we compared achieved hematocrit in each arm.

[Slide]

We used the 2003 and 2004 United States Renal Data System. This is an administrative

database which contains EPO information and hematocrit values based on monthly dialysis claims.

[Slide]

We restricted our study population to patients who were 65 years and older and who had first dialysis claims with 90 days of their first ESRD service data. These two criteria were used to avoid less censoring. Patients who used pre-dialysis EPO were excluded from study because we wanted to make sure that we have some completed patient EPO therapy treatment history. Then, patients who were diagnosed with HIV or cancer were also excluded because those patients may respond to treatment differently.

[Slide]

These are the censoring events that we used. We censored patients at these events, whichever occurred first. They we estimated average EPO dose in the first three months. We estimated the effect of average EPO dosing in the first three months and the hematocrit at month four. So, by choosing EPO-naive patients and by

restricting the study period to the first three months, we selected the period in which increase in hematocrit will be greatest. Outcome is chosen at month four because there is usually a two- to three-week lag period for EPO to affect hematocrit values.

[Slide]

We used inverse probability rating technique to control for treatment by indication bias, and we constructed a dose-response curve. Each point in the curve shows the estimated average hematocrit if subjects had been randomly assigned to that EPO dose. Since Dr. Hernan has given us an overview of how this works, I will focus on the results.

[Slide]

I will first show the observed data and then show the model of the data. This is showing the initial EPO dose distribution. Thirty percent of patients received initial EPO dose within the FDA-recommended range, received 50-100 U/kg. More than 60 percent of patients received initial EPO

dose higher than that level.

[Slide]

At the end of month three, more than 30 percent of patients achieved a hematocrit greater than 39 percent, and this group of patients also received highest average EPO doses in the first three months.

[Slide]

This is the estimated dose-response curve. The X axis is average EPO dose, and the Y axis is hematocrit values at month four. The dotted line indicates confidence intervals. Again this is the estimated curve so each point in this curve indicates the population hematocrit average if the subject had been assigned to that level EPO dose. So, this is model data, not observed individual patient data. So, the average EPO dose of approximately 13,500 units/week would result in a population average hematocrit of 36 percent. The greatest increase in hematocrit takes place with EPO dose between 9000 units/week and approximately 20,000 units/week or 22,000 units/week. For larger

EPO doses, higher than that, the population average hematocrit tends to plateau at about 38.5 percent. Those red dots indicate the range of study EPO doses recommended by FDA, which is located by the linear portion of the curve.

[Slide]

The curve, based on the standard linear regression adjustment, is much flatter and the plateaus are lower hematocrit so that it is it is less biologically plausible compared to our estimated curve.

[Slide]

This study has several limitations. First of all, there might be residual confounding introduced by unmeasured clinical factors. For instance, we have hematocrit taken at end of month and total EPO dose from an administrative database, but we may not have the hematocrit that physicians actually based upon when they are making dosing decisions. And, several clinical factors such as iron level, blood pressure and nutritional status are not included in the United States

administrative databases.

So, although in practice EPO dosing decisions are largely based on hematocrit values, meaning that those factors could be mediated through hematocrit values, several additional caveats include that the study research question may not reflect current anemia management strategies. We did not consider dynamic EPO dosing regimens, and the conclusion of the study may not be generalizable to different study populations or may not be generalizable to different study periods.

[Slide]

Conclusions: The dose-response curve is S-shaped. Hematocrit plateaus are at about 38.5 percent and normal hematocrit target might not be achievable for a dialysis population and starting doses recommended by FDA are appropriate.

[Slide]

Now let's move on to the relationship between EPO dose and patient survival. This research is also supported by an R01 grant, and we

have recently presented our findings at the 2007 joint statistical meetings in Salt Lake City.

[Slide]

So, our research goals--similar to our dose and hematocrit response study, we tried to mimic an RCT, but in this case we are modeling patients survival so we want to compare survival at different levels post dose. You are all very familiar with previous research.

[Slide]

Especially, Dr. Singh has already given us an overview about existing studies so we can directly go the last point, which is that to date the EPO dose and survival relationship has not been empirically determined.

[Slide]

The study design is similar to our dose and hematocrit response study, and we used exactly the same data source. The main difference is that for this study we selected patients who are from free-standing facilities and we allowed patients to use pre-dialysis EPO.

The exposure for this study is cumulative average EPO dose, and we extended our study follow-up period to month 12. We treat before three months as our baseline.

[Slide]

Again, we did not change our analytical approach. We used the inverse probability weighting technique to control for treatment by indication bias and time-dependent confounding.

[Slide]

This table shows the association between cumulative average EPO dose and patient mortality risk based on IPW, inverse probability weighting, which is the method that we used and standard Cox regression model. Results based on inverse probability weighting show that the mid-range doses, especially those in the second quartile, are associated with better patient survival compared to the lowest and highest EPO dose groups. In contrast, standard adjustment shows that larger EPO doses are consistently associated with higher patient mortality.

[Slide]

Instead of using dose quartiles, this curve shows the association between continuous EPO dose and patient mortality risk. In general the S-curve is U-shaped and doses larger than 15,000 units/week are progressively associated with increased risk of mortality.

[Slide]

This is the survival function over time for three selected EPO doses based upon the different parts of the curve that you just saw. Consistent with what we saw in the previous curve, the mid-range doses, the line on the top, are associated with better patient survival compared to low and high doses.

[Slide]

Study limitations: We had similar study limitations compared to our dose and hematocrit response study so I will not repeat this slide.

[Slide]

Conclusions: Lowest mortality found for average EPO dose of approximately 8,500 to 15,000

units/week, treating all patients with higher EPO doses, greater than 15,000 units/week might decrease average survival.

That is the end of the slides so that is it for my presentation, and next Dennis will summarize our research findings in light of FDA concerns on the issue. Thank you.

MR. COTTER: Just as a time check, we are talking about five minutes max.

DR. PLATT: How about one minute?

MR. COTTER: Good.

[Slide]

As we can see based on Dr. Singh's report, for FDA discussions today perhaps some of these findings might be of interest. Sixty-one percent of the incident patients obviously have doses higher than the FDA-approved label. So, this is a measure of clinicians' attitude and knowledge of an appropriate dose. This has nothing to do with the patients but this is really a physician issue.

Second, dose response, as we can see, once the dose on a population average of our

sampleB-epidemiologists have to continue to remind people of that, that doses in excess of 12,000 units/week will exceed the FDA recommended upper target of 36.

In terms of risk, doses in excess of 15,000 units/week will result in a progressively higher mortality risk.

What this has led us to believe is that hypo-responsive patients are at greatest risk, and this is the subject of our continuation grant where we are looking at hypo-responsiveness as well as anemia management strategies. I think both of these things will have some bearing on future guidelines in terms of appropriate treatment.

Thank you very much for your time, and we will be available for questions later.

DR. PLATT: Thanks so much. We are going to move now to the part of the presentations organized by the sponsor, and the first presenter will be Dr. Eisenberg.

Amgen Introduction

DR. EISENBERG: Good morning, Dr. Platt,

committee members. My name is Paul Eisenberg. I am responsible for Amgen's global regulatory affairs and safety organization. I would like to parenthetically note that, in addition to my recent interest in drug safety and risk management, I have over 20 years of experience in cardiovascular clinical investigation so it is a particular pleasure to be in front of these first joint committee exercises. I think this is extremely important and I am delighted to provide some introductory comments today on behalf of Amgen and J&J.

[Slide]

I first need to note there has been a change from the agenda with respect to our presentation. I will immediately state, Dr. Platt, we will keep to the break as scheduled and the change will only involve Dr. Preston Klassen breaking up his presentation at a different time and Dr. Marc Pfeffer will be presenting, after I make some additional brief comments, on the details of the TREAT study, which is the largest

cardiovascular outcome study in patients with chronic renal failure that is sponsored by Amgen and is currently in progress.

I want to also note that Amgen is the license holder for epoetin alfa and darbepoetin, and that Amgen and J&J are marketing partners for the ESAs in the U.S. and that we have worked together on all aspects of this presentation and that the views that we will present today represent those of both sponsors.

We also want to thank the committee and all the experts who have joined us today to advise us and the FDA in the appropriate use of erythropoietin-stimulating agents in patients with chronic renal failure. I am certain that we all agree that the development of ESAs to replace deficiency of erythropoietin in chronic renal failure patients dramatically changed the lives of these patients, particularly those on dialysis, that was associated with the chronic anemia associated with this disease.

[Slide]

Amgen and J&J have been fortunate to be advised by many distinguished scientists and clinicians on the appropriate use of ESAs and the design of clinical studies. Some of them have joined us today and are listed here. I won't go through this in detail, only to note that we have invited Dr. Allen Nissenson to discuss how the benefits and risks of ESAs have been managed in clinical practice based on evidence-based guidelines.

[Slide]

Of course, the reason we are here today is to review and consider the implications of data from two clinical trials that indicate that outcome was worse when hemoglobin targets greater than 13 g/dL were targeted. The results of these trials are appropriately highlighted by FDA in framing the questions that we will consider. We believe these data need to be considered as well in the context of the TREAT study that Amgen is sponsoring. And, I would like to now invite Dr. Pfeffer to comment on the TREAT study.

The TREAT Study

DR. PFEFFER: Ladies and gentlemen, I am here representing the TREAT study, to tell you about perspective while this is an ongoing, vigorous trial and what the objectives are.

[Slide]

As any clinical trial, it is a team event and I am representing the executive committee. I would point out that for me this was a real culture change to do a multi-disciplinary study where leading nephrologists, cardiovascular investigators, endocrinologists all came together to address this issue. As such, there were obvious differences in the approach. We have an outstanding data safety monitoring board. Many of you will recognize the names; the Who's Who in how to monitor trials are monitoring this particular study.

[Slide]

When we came together there was no such thing as results from CREATE, CHOIR or TREAT. It was really left to the discretion of the physician

and there was really very little in the way of hard outcome studies, and credit to these three investigator teams to attempt to provide us with how do patients do besides raising their hematocrit, reducing transfusions and subjective feelings. Can we alter the outcomes of the patients? So, there were three trials. Dr. Singh has told you about CREATE and CHOIR and I will tell you about TREAT.

Highlighting a few things here, randomize is the key word for all of these. They all have that key word and I think that is a very important factor and a lot of the data you are going to be discussing are observational, the best we currently have but this is the gold standard.

There is something else here, the double blind-Bdouble-blind, meaning the treating physician isn't aware nor is the patient. That is a higher bar to have. And placebo. Now, this was a very interesting observation to me. I just assumed that we would do placebo until I met my renal colleagues

who said you can't; you can't do that. We already know that this is a proven therapy, not for outcomes. So, that was the beginning of the culture shock for me and it was so strong that in some countries we are not permitted to do this. We would be able to do it if we didn't have placebo. So, the concept was so strong that these patients must be treated.

I will also point out that for the target arm both groups are on therapy so we are comparing therapy to therapy as opposed to therapy to no therapy, the placebo. Then I will point out the target for what is called the high arm. From CHOIR it was 13.5; here it is 13-15 and we are 13, the lowest of that.

In CHOIR the target is 13.5 but the achieved is 12.6. So, this isn't like the thermostat in your room. This is biology. You just can't dial a number, give a dose and get to that number. I think the better analogy is that this is the thermostat in your room with the windows open for some patients and not others. So,

physicians no longer even get an EPO level. The level doesn't correlate. So, there is a lot to be learned here but our target was 13.

The outcomes are what we don't want our patients to have but what they will have because of the disease. Can't we modify that?

Then, I think it is a very important number, not just the sample size but how many events are we talking about? How many patients had one of these major cardiovascular events? We believe that in order to test our hypothesis, to test the hypothesis that treatment can alter outcomes, we need to study enough patients to have, unfortunately, 1,200 experience a major cardiovascular event to answer the questionB-to answer the question so the question, we believe, is still on the table and let me tell you where we are.

Another important factor of the design of the trials was that in CHOIR once a patient ended up on renal replacement therapy their data was no longer contributed. Now, in TREAT that is an

assumption we didn't make. I don't think the patient cares if they had a stroke after they get on dialysis or before dialysis. So, we are counting all these events.

[Slide]

So, these are the features. Another one that I will point out is that 100 percent of our patients are diabetic so these are diabetic patients with chronic kidney disease, not on dialysis, who all have anemia. Shall we leave them alone or should we correct their anemia? Now, "leave them alone" is not the proper thing to say because one of the nice things about being in clinical trials is the attention these patients get. If I were to show you our baseline data, which I can't because we are still accumulating it, you would be very pleased to see the blood pressure is controlled; the LDL is down. We are addressing the other issues and now we will ask the major question in a placebo-controlled fashion.

[Slide]

You heard about CHOIR and I think that is

the data we have on the table today. That is the data on the table today and I react to data. So, yes, there is this imbalance and, yes, there is an imbalance in death and predominantly death and hospitalization for heart failure. That is important and let's register that, but actually it is not consistent with the Normal Hematocrit trial. In the Normal Hematocrit trial it was death and myocardial infarction, which is not a big player in this study. So, we have that inconsistency right here.

[Slide]

What did we do? Well, we started our trial years before CHOIR. Fortunately, we were well aware of CHOIR. Some of the members of our executive committee were also in CREATE. Some of the members of our executive committee were in CHOIR. It is small community; we are trying to get the answers. And, when we heard about this we proactively notified the data safety monitoring committee. We proactively notified our sites and, more importantly, we proactively notified our

patients. All patients were re-consented and the consent form was changed. We already had the Normal Hematocrit study in it. We already had the information from some of the oncology trials. We then promptly updated it so that by the time the publication came out in the *New England Journal* our sites had already been notified, as had our patients. But then we sent the new information, once it was published we sent that information to our sites.

More recently the FDA has made a change in the label. The data from this was known to our sites but because of that label change we once again notified all our sites, notified the data safety monitoring committee, and once again changed the consent form so all patients, whether they had been in the trial and had consented were re-consented and all new patients were told about this label change.

Because of the CHOIR study and because this is what is on the table today, we went one step even further and asked our data safety

monitoring committee, which has more experience than all of the members of the executive committee combined, to do something else. We had their wisdom. We had their wisdom on the table. We had them looking at all the information from outside of TREAT plus inside TREAT but we wanted to go one additional step. We asked them to take a very conservative stopping rule. Now, I had my hand slapped when I mentioned that to them and they said we go by guidelines. And, I said with the climate the way it is, we would like you to also have a rule. You can go beyond the rule but we want to know at least we haven't hit that level. And, that level was to say if there is any harm for our primary or mortality for two different things to look at, at an 0.05 level, nominal 0.05, at any time, knowing that multiple looks at 0.5 would be a much conservativeB-that we might even be declaring harm that might not be there between that is the position we took.

That position is in place. They met on July 18th and found, this is a quote, no cogent

reasons to recommend alteration or termination of TREAT. Now, from the knowledgeB-I have no knowledge of the treatment assignments but from the knowledge of our overall event rate, I can stand here and tell you that if there was harm in TREATB-and we are doing a trial to look for benefitB-if there was harm it could not exceed this point estimate. The point estimate, I will remind you, in CHOIR was 1.33 or 1.34. And we continue this monitoring.

[Slide]

Now, the other aspect that made me pause personally was as a cardiologist, treating anemia and heart failure is a very hot topic. We have observational data and you heard a lot about that. We have small trials. You haven't heard a lot about that but there are small trials saying my patients feel better; they do better, but we don't know. We don't know.

So, there is enough of a question in the cardiovascular area that treating anemia in patients with symptomatic heart failure is a

top-line item. And, I was aware of some of the safety data that Amgen had in placebo-controlled trial experience in their pilot studies, so a total of about 500 patients, 400-something patients, placebo-controlled. If the canary in the coal mine is the precipitating heart failure, then patients with heart failure should do a lot worse when their hematocrit is raised with an ESA and the results, at least in this pilot experience, were the opposite.

[Slide]

It is so encouraging that the sponsor, Amgen, has embarked on and started, well under way, a study of 3,400 patients who have symptomatic heart failure, depressed ejection fraction and anemia to treat their anemia. So, in the cardiovascular arena this is an area of uncertainty where data is being accumulated, and I will point out this is also an event-driven trial looking for over 1,000 events in a double-blind, placebo-controlled fashion.

So, where are we? In the non-dialysis

patients, the large universe of people with anemia and chronic kidney, we have these two trials you have heard about. At the time, TREAT, when our data safety monitoring board met knowing those results, knowing other results, and taking the totality including looking at TREATB-I am jealous of that, looking at TREAT, they said no cogent evidence to alter. They were looking at 514 events at that time. We are on target with the event rate, the patient population. We will finish enrollment, we are at 4,000 patients, in the next two months.

[Slide]

And, what we have here is that I believe we are addressing the question, can we alter a patient's outcome. When we started the placebo group was the one that was drawing attention; now it is our treated group and I think that defines uncertainty, and I think the only way to answer uncertainty is by what we are doing.

I did tell you we are very pleased about this. It is in 26 countries but more than half the

patients are from the United States. Over 2,000 patients are from the United States. Statin-use blood pressure control all around the world. Enrollment is near complete. We are doing very well on this aspect. You can always do better but if I had to stand here today with you having our data, I would not be ashamed at all with where we are now.

I can also tell you that the event rate is right on track of what we said it would be in our protocol, which means the disease burden these people have was as predicted and we are trying to see if we can alter that disease burden. So, this uncertainty can only be addressed by robust randomized, controlled data, which we are attempting to get.

[Slide]

I will just conclude by saying that TREAT and RED-HF together will be 7,400 patients. If the emperor has no clothes, we will know that. If, on the other hand, we are altering disease progression and improving outcomes, we will know that. The

risk/benefit, we will know that. And, that is what we should use for the most rational judgment for patients in the long run. I think our sponsor, by doing these trials, is trying to improve the practice of medicine, to improve our understanding and, in the meantime, while information is coming in, I think you are going to be hearing reasonable, useful guidelines for risk management of patients from Dr. Eisenberg. Thank you for your attention.

Introductory Comments

DR. EISENBERG: Thank you. I will provide some brief introductory comments before asking Dr. Allen Nissenson to comment from a clinical perspective.

[Slide]

What is shown here are the key points that I think there is actually general agreement on. The erythropoietin-stimulating agents provide clear benefits in chronic renal failure patients. This was based on randomized, controlled clinical trials that were the basis for the original marketing application, so the pivotal registration trials

that targeted hemoglobins from 10.7 to 12.7, demonstrating unequivocally transfusion avoidance. There was improvement in anemia symptoms in those patients and, I think more importantly, the focus, and we will present data on this, on cardiovascular function, cardiovascular performance, and exercise capacity was demonstrated.

Now, FDA has commented on the level of evidence provided by patient-reported outcomes so that would reflect the concept of anemia symptoms. While we acknowledge that the manner in which such evidence is developed continues to be refined and, in fact is the focus of a well-thought out FDA draft guidance, the benefits in terms of signs and symptoms of anemia are well recognized by nephrologists. They are supported by data and we would be happy to discuss that further.

Now, the distinction between target and achieved hemoglobin is actually going to be a major point of discussion. You have already heard some different perspectives today. I will actually go to the last bullet for a moment, we are going to

spend some time recognizing the fact that is it confounded. But there are certain facts. We believe the data do demonstrate that there is risk when you are targeting hemoglobin greater than 13. I think whatever you believe the level of evidence is from the randomized trials, this is a concern. We believe that to be the case. However, there is no question as well, despite the complication by factors such as underling health status, and you have seen that actually in the presentation from Mr. Cotter's group and from Dr. Singh. Achieving hemoglobins greater than 11 g/dL is associated with better outcome.

Dr. Unger, from FDA, has highlighted the issue of rapid rises or relatively rapid rises or decline in hemoglobin, and of cycling. We will come back and discuss this. But, overall, I think what we want to emphasize is this fact, that the target hemoglobin range of 10-12 g/dL which, in fact, was the range that was recommended for the ESAs prior to the recent label changeB-this has been in the label for a decadeB-is a prudent

approach to risk management. In fact, what we are being asked to consider today is effectively a ceiling of 12 by targeting a hemoglobin of 11. This is not consistent with the results of randomized clinical trials.

So, this essentiallyBand I will come back to this at the end, is what we believe the data support in the robust clinical evidence. I will point out that I agree completely, as I think we all do, with the contention that randomized clinical trials are what we should base our evidence on.

On the other hand, in risk management and drug safety we often do observational studies. We look at all of the available data. We are extremely fortunate in the dialysis population to have a different kind of database. This is not a sampling. This is virtually 100 percent of the patients who are on dialysis that have data in the databases we will describe that allows for complete assessment of the available data in the at-risk population.

The hypotheses that have been generated in analyzing these data are hypotheses and we acknowledge and are committed to randomized, controlled clinical trials to test those hypotheses.

I would like to now turn to Dr. Allen Nissenson who will comment on the experience in the dialysis population predominantly but in chronic renal failure in general.

Clinical Perspective

DR. NISSENSON: Good morning, and thank you very much, Dr. Eisenberg.

[Slide]

I would like to point out for the committee that in addition to my academic responsibilities, I am currently the president of the Southern California Regional Disease Council, one of the 18 organizations contracted by CMS to oversee the quality of care and dialysis facilities. Of particular relevance to this meeting, I am also a past president of the Renal Physicians Association, the National Association of

Nephrologists. I have also trained well over 150 nephrologists during my career. I can, therefore, speak with confidence on behalf of my clinical nephrology colleagues.

I have had the privilege of caring for hundreds of patients with kidney disease over my 30-year clinical career, nearly half of which has been spent in the pre-ESA period. As seriously ill as patients with chronic kidney disease were in 1976 when I completed my fellowship, I cannot overstate how ill they are today. With over half with diabetes and typically four-plus co-morbid conditions, chronic kidney disease patients are among the sickest of the sick and the most vulnerable to changes in their treatment.

I am here today to tell you about how nephrologists view the risk and benefits of ESA therapy and to tell you what the care of kidney disease patients was like in the pre-ESA era, but also what it could be like again if the wrong recommendations are made by the committee. That is, the risk of higher mortality, more

hospitalizations, poor quality of life, more blood transfusions and higher costs to the healthcare system.

[Slide]

Depicted on the left-hand side of this slide is the well-known system by which the body maintains red blood cell mass and health where any decrease in oxygen delivery leads to increased production of erythropoietin by the kidneys which, in turn, stimulates red blood cell production. This exquisitely functioning system is disrupted in the presence of chronic kidney disease, as we heard earlier, with the diseased kidneys no longer able to produce sufficient EPO to keep up with demand. The result in a patient with advanced chronic kidney disease or end-stage renal disease is severe anemia.

Some of the consequences of anemia are shown on the right-hand slide of this slide. These include impaired cognitive function, dizziness, shortness of breath, reduced exercise tolerance, fatigue and weakness.

Now, you might ask how do we know these symptoms can actually be attributed to anemia? We had the opportunity to observe an experiment of nature when EPO first became available. That is, we administered EPO; anemia improved but not other aspects of the care of the dialysis patient were altered. We saw remarkable improvement in the signs and symptoms shown on this slide and reported by patients.

[Slide]

Anemia treatment was a rescue therapy in the pre-ESA era and I can't stress strongly enough how badly patients in the pre-ESA era needed rescue. They were terribly debilitated and had great difficulty functioning from day to day. Why did nephrologists approach anemia therapy as a rescue treatment, holding out any treatment until severe symptoms and signs appeared? Because the limited number and serious adverse consequences of available anemia treatment at the time. The treatment options, as shown on this slide, included, first, blood transfusions, which I will

return to in a minute. Second, parenteral iron therapy when iron deficiency was present. Unfortunately, however, the available iron treatment at the time, intravenous iron dextran, was associated with a low but not rare occurrence of anaphylaxis which at times was fatal. Finally, androgens which caused significant liver impairment as well as masculinization, the latter essentially precluding their use in women.

The complications of blood transfusions are worth dwelling on to point out the concern in the nephrology community about the recommendation that one should, quote, use the lowest dose that will gradually increase the hemoglobin concentration to the lowest level sufficient to avoid the need for red blood cell transfusion, unquote.

Such an approach alone would undoubtedly return us to the days of frequent blood transfusions. Unlike cancer patients receiving chemotherapy, patients with chronic kidney disease have anemia that persists for years, not weeks or

months. In the pre-ESA era the majority of dialysis patients received multiple blood transfusions annually based on severity of symptoms. At that time patients were younger and much less likely to have the multiple co-morbid conditions, including cardiovascular disease and diabetes, compared to today's patients. Symptomatic anemia, as we saw in the pre-ESA era, will likely be even more prevalent in the complex patients with chronic kidney disease we currently treat, and frequent blood transfusions will again become common practice.

[Slide]

The complications that we saw in the pre-ESA era and which will be seen again if the frequency of transfusions increases are listed on this slide. I remember having to carefully examine each patient prior to giving a transfusion to make sure they would not develop severe volume overload with one or two units of blood. I would have to explain to each patient the possibility of antibody formation to the infused blood components and that