such preformed antibodies might preclude them from receiving a kidney transplant or greatly increase their waiting time. I had to warn them that the repeated blood transfusions that they needed would turn their skin bronze colored from deposition of iron, and that iron from the blood transfusion would deposit in their muscles, their heart, their liver, their pancreas, and I wasn't sure of the consequences but they likely were not good. Serum ferritin in such patients generally ranged from 3000-5000 ng/mL, ten times the average value today. In fact, reduction in iron overload was one of the endpoints of the EPO registration trials. I had to draw an emergency potassium level prior to transfusion to make sure that fatal hyperkalemia would not occur because of the large amount of potassium in the blood being administered. Finally, I had to warn patients about the possibility of transmission of potentially deadly infectious diseases, including hepatitis and HIV.

It should also be noted that in current practice, as if those complications are not enough,

blood transfusions are difficult to provide in the outpatient dialysis center and are generally provided at a hospital or infusion center. The logistical burden on patients as well as these facilities of increasing the number of blood transfusions would be onerous.

[Slide]

The availability of recombinant EPO triggered a paradigm change in the care of patients with chronic kidney disease, allowing partial hemoglobin restoration rather than transfusion rescue, as shown in this figure adapted from the seminal paper by Eschbach and colleagues in the *New England Journal of Medicine* in 1987. A typical patient is depicted here and, as shown on the left-hand side of the slide, frequent blood transfusions were required just to keep the hematocrit at the suboptimal level of 25 percent, a hemoglobin equivalent of a little over 8 g/dL. With the administration of EPO, as shown on the right-hand side of the slide, there was a smooth rise in hematocrit over 12 weeks to 36 percent and

blood transfusions were eliminated.

EPO rapidly became the standard of care for dialysis patients as it filled the unmet medical need for improvement in anemia that could not be safely addressed by transfusions, iron or androgens. After only two years of approval, over 80 percent of dialysis patients were receiving EPO.

[Slide]

The dramatic decline in transfusion need in dialysis patients paralleled the increasing use of EPO, as shown on this slide. By the mid-1990s blood transfusion had become a rare event, generally required only after a sudden, acute, severe bleeding episode.

[Slide]

A similar reduction in transfusions has been seen in the non-dialysis CKD population.

[Slide]

As shown on this slide from a study by Bob Dr. Wolfe and colleagues, the standardized mortality ratio of dialysis patients has decreased by 17 percent since the introduction of EPO in

1989.

These data are only corrected for age, gender, race and primary diagnosis but not the co-morbidities which increased significantly over this time period, making these results even more striking.

While these survival improvements are unlikely related to anemia management alone, some hf claimed that despite average hemoglobin levels rising over the past 18 years there, in fact, has not been an improvement in survival in dialysis patients. Such claims are clearly not accurate.

[Slide]

Finally, it is without question that improvement of anemia with EPO has led to significant improvements in patient well-being as observed by nephrologists and reported by patients. It is useful to understand why my fellow clinicians and dialysis patients concur with me on this issue. I will illustrate this point in a personal way but my own experience is mirrored by the thousands of nephrologists who have managed anemia with ESAs over the past nearly 20 years and

the hundreds of thousands of chronic kidney disease patients in this country alone who have received ESAs.

Now, you are going to see data from randomized, controlled trials on this issue a little bit later in dialysis patients, but it should be pointed out that in both arms of the CHOIR study, as mentioned by Dr. Singh, including the one targeting a hemoglobin of 11.3 g/dL in CKD patients not on dialysis and actually achieving a hemoglobin of 11.3, with a range of 10-12, there was a statistically significant improvement in multiple measures of quality of life in both of these randomized groups.

When I first started treating patients with EPO I was astounded by the reports of dramatic improvements in energy, ability to concentrate and to get through their activities of daily living. This seemed truly remarkable and was significant enough that many patients felt so much better that they requested coming off transplant lists because they no longer, quote, felt terrible, unquote, on

hemodialysis. Of course, this wasn't something we endorsed. I was intrigued enough by these patient-reported improvements that I initiated clinical studies to test whether these were, in fact, attributable to improvement of anemia. Along with other investigators, my group was able to show that, indeed, raising the hemoglobin resulted in improvements in symptoms, physical functioning, functional capacity and cognitive function.

I then took a step further by studying brain electrophysiology and was able to show that objective evoked potential electroencephalograms confirm the findings from the patient-reported and neurocognitive studies. I was convinced, as are my colleagues and patients, that there is a significant clinically meaningful improvement in patient-reported outcomes in functional ability with anemia treatment with ESAs.

[Slide]

In conclusion, ESAs have fundamentally changed the practice of nephrology so that nephrologists commonly refer to chronic kidney

disease patient care as pre-ESAs or post-ESAs. ESAs enhance patient well-being, a point strongly agreed upon by nephrologists and patients.

Returning to a time with significant anemia was present in many patients and blood transfusions were frequently required would be setting back the care of chronic kidney disease patients nearly two decades. These patients deserve better. These realities, of course, should not be overlooked while the legitimate scientific debate continues over the most appropriate target hemoglobin.

[Slide]

In fact, recent revisions to evidence-based clinical practice guidelines take this debate into account and make clear important safety information regarding the use of ESAs in this population. These revisions reinforce current nephrology practice. That is, the benefits of anemia correction, particularly those related to patient well-being, must be weighed against the risks for each individual patient. Management

decisions regarding anemia are made by nephrologists working with patients, with the goal of achieving optima clinical outcomes. The goals for provider and patient are the same, maximizing quality of life and ability to function on a daily basis while minimizing the risks of adverse outcomes.

Recent statements in a variety of venues suggest that non-clinical factors drive these critical aspects of care, including adjustment of EPO dose or selection of the target hemoglobin, are simply not correct. Some have suggested that clinical trials are indicated in chronic kidney disease patients to test whether a target hemoglobin of less than 10 g/dL would be more appropriate. There is not, however, a significant disagreement among clinical experts about the preferred lower target hemoglobin. In fact, independently developed clinical practice guidelines from Australia, the United Kingdom, Canada, Europe and the United States all recommend a minimal target hemoglobin of 10.5 in the case of

the U.K. or greater than or equal to 11 g/dL for the other four guidelines, including the guideline developed by the National Kidney Foundation in this country. Such a trial, therefore, would lack clinical equipoise.

Finally, I implore the committee to seriously consider the 18 years of clinical experience of nephrologists and patients regarding the improvement in patient-reported outcomes and functional status with the partial correction of anemia. When such vast experience is consistent with observational trials and these, in turn, are consistent with the available, though limited, randomized, controlled trials it should not be dismissed. The experience of patients and nephrologists matters.

I would now like to turn the presentation over to Dr. Preston Klassen.

Benefit/Risk

DR. KLASSEN: Chairman and members of the joint committee, in 2001 I was a clinical nephrologist on faculty at Duke University,

focusing on patient care and clinical research, and I was one of a number of individuals involved in the early design of the CHOIR trial.

[Slide]

The results of that trial and others that we are discussing here today targeting high hemoglobin values have demonstrated apparent risk and that risk has raised appropriate concern.

As we focus study on that concern likewise, and as mentioned by Dr. Nissenson, it is appropriate to recognize the fundamental importance of erythropoietins to the medical care of patients with renal disease. In these patients with epoetin deficiency ESAs are essentially the only effective medical therapy since chronic anemia since chronic anemia of renal disease was initially described by Brighton Christenson in the 1800s.

Dr. Nissenson has described some of the clinical perspective and I will simply add to this that transfusions, chronic transfusions are, of course, more than simply an order written in a patient's medical chart. They are disruptive

events that carry real risk and the symptoms and the physical disability associated with chronic anemia are real and every nephrologist has dealt with this.

A target not to exceed 11 g/dL with dose reductions above 11, as we will see, will be expected to increase transfusion rates and increase anemia symptoms compared with a target range of 11-12 with dose reductions above 12. So, as we consider the risk and recognize the importance of these therapies we are really dealing with the key question at hand, what is the appropriate use of ESAs.

[Slide]

We are going to walk through four kinds of information today briefly. The original pivotal trials of epoetin alfa used a target range to guide ESA dosing and demonstrated unequivocal clinical benefit in the form of transfusion avoidance, and improvement in the symptoms and physical function associated with anemia. So, we learned from these trials that ESA dosing can be guided by hemoglobin

targets and that produces clear clinical benefit.

Now, we know that epoetin alfa was rapidly adopted in dialysis and over 15 years of clinical experience with ESAs has actually occurred, as mentioned, in a relatively unique situation in a patient population with comprehensive clinical outcome data collection both through a federally mandated program, the United States Renal Data System, or USRDS, as well as large U.S. dialysis providers who captured detailed patient level information. And, these data show decreasing mortality rates across that population as ESAs were used to target a hemoglobin of 10-12 or, in many cases, 11-12.

These data also show that the occurrence of hemoglobin at lower levels is associated with a higher rate of blood transfusions. These data support that the community use of these agents, focusing on hemoglobin targets based on ESA labeling, have bee appropriate. Of course, these are comprehensive clinical practice data which is good from a surveillance perspective. They are not

randomized, clinical trial data, for example, comparing the target of 10-12 to something lower and we will talk at the end of this presentation about how we have been considering approaching hemoglobin target trials.

Now, among the additional clinical trials that have, of course, informed the appropriate use of ESAs are the higher target trials, namely, Normal Hematocrit and CHOIR, providing evidence that targeting a patient population above 13 results in apparent increased risk. Yet, we see in these trials and in clinical practice that patients who actually achieve higher hemoglobin have less risk. That raises the question of what exactly is creating or is associated with increased risk in patients with lower hemoglobin levels and whether, in fact, that risk might be related to dose. But the message here is that targeting hemoglobin values above 13 should not be done in clinical practice and ESA labeling has always been consistent with that message.

I will discuss briefly the issues of ESA

dose and clinical risk, and highlight issues of confounding using the comprehensive clinical practice data, and will spend a brief period of time on the phenomenon of hypo-responsiveness from some of the higher target trial data. Finally, I will wrap up with future research that will inform the appropriate use of ESAs, focusing on the areas that have been pointed out by both the FDA and the sponsors. We think these are important areas to consider and we are seeking further guidance from the panel today on this effort. I will note that we will take a short break. We will try and keep it at the right time at 10:10.

[Slide]

Let's start first with the pivotal trials and the concept of hemoglobin targets to guide ESA dosing. This concept is not unique. Therapeutic targets are used in a variety of settings in chronic disease and hypertensives are titrated on blood pressure, insulin and glucose, and ESA dosing has always been titrated on hemoglobin levels. The rules for that titration come, in part, from the

original pivotal trials, which we will see in a moment, targeting a hemoglobin range with the lower end of 10.7 and a higher end of 12.7. Prior to the recent label changes, as noted, hemoglobin targets have always been a part of ESA labeling, specifically for epoetin alfa a target range of 10-12 and for darbepoetin alfa a target not to exceed 12.

[Slide]

Let's first start with the initial dose-ranging study of epoetin alfa in dialysis patients. The doses tested in this study ranged from 15 U/kg per administration to 1500 U/kg per administration. The hemoglobin value at week 4 is graphed here, on the Y axis, or change in hemoglobin from baseline after four weeks of epoetin treatment. Along the X axis is the cumulative dose so basically the monthly dose.

So, what you see here is that monthly doses up to approximately 400,000 units per month display the clear dose-response relationship, with higher doses producing higher hemoglobin levels at

four weeks. It is important to point out that these data points are from the second to highest group, the 500 U/kg per administration dose. The model is fit using that highest group of 15,000 U/kg. They are not graphed here. That is one million units based on body weight that the patients received. So, the assertion that this study had limited dose ranges is incorrect.

Mr. Cotter's and Dr. Zhang's dose-response model presented on slide 29 here today in their presentation attempts to mimic an actual dose-ranging study using real-world claims data, and their model suggests that doses above 80,000 units per month are on the flat part of that curve. That is a bit puzzling because a real dose-response study, of course, requires fixed dosing or forced titration, which is not the case for population-based data, and this is important because the data from actual pharmacodynamic dose-ranging studies are clear that there very much is a dose response with ESAs, with no evidence of a dose threshold with anything under 500 U/kg per

administration for epoetin alfa which is approximately, based on body weight, 400,000-500,000 units per month.

[Slide]

Now, the transfusion benefit is demonstrated here in one of the double-blind, randomized, placebo-controlled studies in which hemodialysis patients with a baseline hematocrit of 22 or a hemoglobin value of just over 7 were randomized to receive epoetin alfa or placebo. The target in the study in those days was a hematocrit of 35, plus/minus three percent, which is basically a hemoglobin of 10.7 to 12.7. And, you see here clearly the reduction in those randomized to receive epoetin alfa in three months, and at six months, between months three and six, no patients in fact received a transfusion. The placebo data are shown here. Over the first three months, no appreciable change in transfusion. These patients were actually then switched over to receive epoetin alfa and they too derived the same magnitude of benefit as those originally randomized to epoetin

alfa.

Epoetin alfa was approved in dialysis on the basis of two adequate and well-controlled studies in addition to long-term single-arm trials. Now, beyond transfusion avoidance, the other major clinical benefit observed was improvement in the symptoms and physical function associated with anemia. These endpoints were added to pivotal trials in part based on discussions with the FDA, and that makes sense in that we know that hemoglobin values in a population correlate with signs and symptoms and, in fact, double-blind, placebo-controlled trials have demonstrated improvement in signs and symptoms.

[Slide]

I will walk through one such study, the CESG or the Canadian Erythropoietin Study Group, which evaluated anemia symptoms and physical function in a three-arm trial. First, this was a randomized, placebo-controlled trial in dialysis patients with an inclusion criteria of a hemoglobin value less than 9; 118 patients randomized to one

of three arms, placebo, or group A targeting a hemoglobin of 9.5 to 11, and group B targeting a hemoglobin of 11.5 to 13.

Exercise capacity was measured in terms of exercise endpoints as a six-minute walk test or an exercise treadmill stress test, the modified Naughton protocol, standard and objective measures. Patient-reported endpoints were measured using two instruments, the KDQ, Kidney Disease Questionnaire, and a Sickness Impact Profile. Both instruments have been well validated specifically in the CRF patient population.

In terms of the analysis, baseline, two, four and six months. We will just show one data slice at six months for brevity. In terms of the analysis that we will be showing, an intent-to-treat model, we have used a variety of sensitivity measures but what I will show is accounting for repeated measures a mixed model.

[Slide]

Starting with the objective measures of physical function, exercise capacity, we see that,

first, in terms of hemoglobin from baseline to six months the placebo group did not change in terms of hemoglobin. There was just under 3 g/dL increase in group A targeting, again, 9.5 to 11. Group B targeting 11.5 to 13 had just under a 3.5 g/dL increaseB-I am sorry, just over 4.5 g/dL increase. In terms of exercise stress, the modified Naughton protocol minutes walk, just under two minutes over placebo in group A, 3.5 minutes over placebo in group B. In total, looking at both epoetin-treated groups combined, 2.5 minutes over placebo. To put this in context, this amount of increase is not only statistically significant but it is clinically meaningful and on par with improvements shown with other therapeutic agents known to increase exercise capacity. In terms of the six-minute walk we see the same kinds of results, upwards of 30 meters over placebo in group A and upwards of 60 meters over placebo in group B.

So, evidence of not only an effect treating with epoetin alfa but some evidence of a dose effect or a target effect in terms of

targeting hemoglobin values in the 11.5 to 13 range and receiving greater benefit.

Now, this study also measured patient-reported outcomes with those two standard measures, the KDQ and the Sickness Impact Profile. Statistically significant and clinically meaningful improvements in physical function, energy and weakness were seen comparing epoetin alfa to placebo and numeric improvements were seen in most measures comparing group B to group A.

[Slide]

So, in summary in terms of the clinical benefit demonstrated from the original pivotal trials, a clear reduction in the burden and risks of transfusions. This is unquestioned. It is well documented in the literature and in the nephrology community's experience. The double-blind, placebo-controlled data demonstrate an improvement in exercise capacity with objective measurements and improvements in patient-reported outcomes and physical function.

I will point out that in two other

randomized, double-blind, placebo-controlled trials, 8701 and 8904, all scores for energy and weakness favored treatment with epoetin alfa and in 8904 those differences were statistically significant. This has been corroborated by systematic review of the literature. In fact, all published articles in the literature have replicated these types of findings but it must be noted that these studies were either single-arm in design or single-arm in analysis which is a limitation and this is in part because equipoise was lost early on in terms of anemia with therapy with epoetin alfa, and it became unfeasible to conduct placebo-controlled trial in dialysis patients but we do see corroborating results.

[Slide]

So, again, as we look at the clinical experience as these compounds were introduced into the patient population, I think we have a rather unique situation. Again, when a pivotal trial showed such benefit there was rapid adoption. Dr. Nissenson has discussed this. And, the population

is a population with comprehensive clinical practice data collection, again, USRDS and large dialysis providers.

[Slide]

So, what do we see over a period of time when there is almost a single point in time introduction of epoetin alfa into a dialysis patient population that is being actively surveyed? As shown by Dr. Nissenson, what we see is that the population risk is not evident with the introduction of these agents. This graph shows annual mortality rate adjusted for minimal factors from the 2003 USRDS ADR. Prior to 1989, a relatively high and stable rate in about 250/1000 patient-years. That is obviously much higher than the general population. And, in 1989 a gradual decline.

Now, these data clearly don't demonstrate that epoetin alfa was causal in making this change. Other important changes in dialysis care included gradual recognition of the importance of dialysis dose. But in a patient population where almost

everyone is taking the drug, if the drug were toxic one would expect to see an increase in mortality to at least some degree. The 2007 ADR has just come out. The adjustors are slightly different but the most recent report from 2005 shows 200 deaths per 1000 patient-years.

[Slide]

We have also seen in comprehensive clinical practice data as the nephrology community has generally targeted between 10-12 or 11-12, this relationship between observational achieved hemoglobin values and risks, and this is simply to point out what others have already pointed out, the relationship between increased risk with hemoglobin values less than 11. Of course, these data are confounded and we will speak about confounding after the break which is coming up in a moment.

[Slide]

What do we also see when we look at large dialysis data sets and the rate or risk of transfusion? This is looking at the Medicare ESRD set across 160,000 dialysis patients from 2004.

You simply break them up into three categories. Over a six-month period of time did these patients have ever hemoglobin values greater than 11? That is about 40 percent of the population. And, if they did, their rate of transfusion per 100,000 patient-years is about 1. The next group is any hemoglobin value less than 11. That is about 60 percent of the population so these may have had most values above 11 but at least one value less than 11 and that is a 10-fold increase in the rate of transfusions. The final group is a group of patients with hemoglobin values less than 11 consecutively for that six-month period of time, and their rate is another 10-fold increase. This is obviously on a log scale in terms of the increase in rate of transfusions.

I think it is important to note, and both the FDA briefing document and Dr. Singh have touched on this, that there is significant hemoglobin variability in this population. Cross-sectional hemoglobin evaluation does not suggest that individual patients are being targeted

at a given level. In fact, 83 percent of hemoglobin excursions greater than 12.5 g/dL are titrated back into therapeutic range within three months. And, as we see, in terms of values consistently below 11, that is a very small percentage of the population over a six-month period of time. So, dose titration is used to keep hemoglobin within a specific range.

[Slide]

My final slide before the break is that, again, when thinking about the concept of hemoglobin targets, we believe that hemoglobin targets below 10 would certainly not be supported by evidence as clinical trials have demonstrated benefits based on a target range in pivotal trials of 10.7 to 12.7, both in terms of transfusion avoidance and improvement in the symptoms and physical function associated with anemia.

These clinical benefits have been supported by comprehensive practice data, and a minimum target of 10 g/dL was at least necessary in the CRF patients to achieve demonstrated clinical

benefit. In consideration of the upper end of the target range of 12 g/dL, we believe that is well below the target that has been associated with risk in studies demonstrating apparent increased risk with Normal Hematocrit and CHOIR.

I think at this point in time it would be appropriate to take a short break. Then we will come back. I will talk briefly about the association between ESA dose and risk and hypo-responsiveness and then we will end up with our future research program. Thank you.

DR. PLATT: Good. Thanks very much. In 15 minutes we will start up again. Thank you.

DR. PHAN: The committee is reminded again to refrain from discussing the meeting during breaks.

[Brief recess]

DR. KLASSEN: Thank you.

[Slide]

Now I will speak about the higher target trials, but actually I am going to do so very briefly because, of course, they have been covered

in detail so far today.

[Slide]

It is not really a question of why these trials were done. In the mid-1990s that was the question. If partial correction showed this kind of benefit, then normalization or complete correction must be better.

[Slide]

Again, I am not going to go through the details of the Normal Hematocrit but I will simply point out that these results were, of course, rapidly communicated through ESA labeling and have been well-known to the nephrology community for over a decade, as Dr. Nissenson has pointed out.

[Slide]

Likewise, in terms of the CHOIR trial, I will only note that in this study a dose cap was put in place for the protocol of 20,000 units per administration either weekly or every other week, and about 75 percent of patients in the high target arm hit the dose cap at least once during the course of the trial. Because it was an open-label

study physicians knew that their patients were randomized to the higher arm and, as a result, these patients may have had additional maneuvers to enhance hematopoiesis, most probably parenteral iron utilization and we know that iron use was numerically greater in the higher arm and iron stores were significantly higher in the low arm.

[Slide]

These studies, Normal Hematocrit and CHOIR, again raised this apparent paradox of targeted versus achieved hemoglobin where targeting a population to a higher hemoglobin, above 13 g/dL, appears to confer risk. But if you look at the achieved hemoglobin values it is actually patients with the lower hemoglobin values achieved--and these are pooled data from the Normal Hematocrit study so both arms combinedB-are associated with less risk, and this is simply the kind of relationship that is seen in all the observational data sets from the earlier slide that I showed.

So, the question is what is going on with these patients in terms of their underlying health

status, in terms of the relationship between the dose of the ESA that they are on? Is there something that we can determine?

[Slide]

So, now I will actually move into that relationship between dose and risk. As well, I will talk a little bit about hypo-responsiveness.

[Slide]

The question that we are really interested in is whether there is an independent and, in fact, causal effect of ESA dose on clinical outcome? We have spoken in detail, others have, about the confounding related to this. But as a result, the evidence suggests it may not be possible to get a direct answer to that clinical question because of those very close, inseparable links between hemoglobin, ESA dose and the relationship between them and the ESA responsiveness.

One of the very important points to recognize is that the highest dose requirements are given to patients with the worst health status. I will show that in a moment. But because of their

poor health, they also have the greatest risk of mortality. The epidemiologists who are like-minded in the room will, of course, recognize this as confounding by indication, that is to say there are factors like underlying health status that are related to dose and are also related to outcome, and without the appropriate control for these potential sources of confounding the effects of health status on outcome, which are obvious and known, may be incorrectly attributed to dose. In fact, in order to really assess the independent effect of dose on outcome one would be required to control for every factor related to ESA dose and also related to outcomes.

[Slide]

The importance of this relationship and the issues of confounding are fairly easy to spot. These are data from a paper by Zhang and others, published in 2004, observational data again from the USRDS data set. A large number of dialysis patients were examined to determine just their baseline dose of ESA. That is broken into

quartiles in the graph on the Y axis, with the highest doses in the green. Along the X axis is their achieved baseline hematocrit. Along the Z axis or vertically, the column heights, is the mortality rate per 1000 patients. These results are unadjusted for confounding co-morbidities and if you focus on a particular hematocrit groupB-we will say here the 11-12 range, you see that in order to achieve 11-12Bthis is hematocrit 11-12B-sorry, hematocrit 33-36, hemoglobin 11-12. In order to achieve that level of hematocrit or hemoglobin, if you required higher doses you had a greater rate of mortality.

[Slide]

What is going on with those patients? Let's take a look from the Fresenius Dialysis Provider data set, which has a much more rich set of clinical variables collected on a routine basis in all dialysis patients that they serve. We will look at patients grouped by simply the category of hemoglobin 10-12.

That is shown here. So, these patients,

12,000 in total, all had a baseline hemoglobin value or 10-12. I have graphed the dose quartiles again on the Y axis; highest doses in the back. And, we are looking at other baseline characteristics that we know as nephrologists are associated with poor clinical outcome. Have you had a percutaenous catheter? Do you have a percutaneous catheter at baseline? A percutaneous catheter for dialysis access is associated with increased risk of infection, increased risk of hospitalization and increased risk of death. Have you been in the hospital recently, in the last five to six months? If you have been in the hospital recently that proportion goes up as dose goes up. So, the patients who are receiving the highest doses are the patients who have percutaneous catheters. They have been in the hospital recently, and those hospitalizations are often for cardiac and infectious causes and nephrologists will know that these are the components that basically kill dialysis patients--catheters being associated with very poor underlying health status.

They have catheters, they go into a hospital for cardiac and infectious causes.

[Slide]

So, what happens when we try to adjust for that underlying health status in patients who are receiving high dose? This is from the full Fresenius data set so not just patients who are captured by a hemoglobin of 10-12 but the entire spectrum, 24,000 patients total. In this analysis dose is listed on the Y axis and this is in a natural log to account for the skewing of the dose distribution. So, what we are looking at is for each unit increase, that is about a threefold increase in the unit of epoetin alfa administered.

The unadjusted analysis is here on the left. So, not adjusting for anything else, you see, as we have seen in other graphs, a clear relationship between dose and the risk of mortality. Higher dose at baseline is associated with a greater mortality risk.

The second estimate adds in baseline covariates that capture underlying health status,

age, gender, diabetes, hospitalizations, etc. You can see that the association is substantially attenuated with this adjustment.

In the third analysis we actually move to a time-varying model with a two-month lag on the outcome for dose. This analysis also adjusts for time-varying hemoglobin in which hemoglobin is lagged by another month. With this final level of adjustment the association is completely attenuated. This paper is now in press in the *American Journal of Kidney Disease*.

As an aside, Amgen has also used marginal structural models to evaluate the ESA dose and the mortality association, and we have come to similar conclusions as shown here, on this slide-Bno relationship between dose and the level of mortality risk when using these more detailed or sophisticate modeling techniques. This is not to say that it doesn't exist. This is to say that it is very complicated and these models vary greatly on the assumptions and what we really need are randomized, controlled trials. We will talk about

that at the end of the talk.

So, there are a variety of ways to look at the potential relationship between dose and clinical risk, and the bottom line is that that relationship is confounded.

[Slide]

Now, we have talked about greater dose requirements and risk of death and I am just going to very briefly touch on hypo-responsiveness. We have actually looked at hypo-responsiveness specifically in the Normal Hematocrit study. I believe additional analyses from the CHOIR trial will be forthcoming today.

If we look at the high arm of the Normal Hematocrit what we recognize is that it is an opportunity to look at hemoglobin or ESA dose response, so a dose challenge and looking at the hemoglobin response. So, in the high arm of the Normal Hematocrit study all patients, regardless of their hemoglobin level, received immediately a 50 percent increase in their dose so they had a dose challenge. Over the next three weeks they had a

hemoglobin response. If you take the slope of that hemoglobin response and divide that by the dose that gives you an index, a measure of responsiveness, how much did the hemoglobin change for that dose challenge?

If you graph that in terms of quartiles, you see that the best responders, on the far right, compared to the worst responders, on the far left, in terms of the adjusted all-cause one-year mortality risk, greater mortality with poorest response. Why do patients have poor response? They have poor response because they are more ill, underlying health status factors, as well as a variety of unmeasured factors that we don't have full understanding of at this time.

[Slide]

So, in conclusion, the unadjusted association between dose and clinical outcomes are clearly confounded by underlying health status but, as mentioned by others today, there are other unmeasured confounding variables.

In terms of responsiveness, we think it is

a risk factor. We think it should be recognized and evaluated. In fact, all clinical practice guidelines and all ESA labels currently recommend evaluation of poor ESA response. We think that working definitions of hypo-responsiveness can and have been developed, and Dr. Eisenberg will talk about some of that as he discusses the risk minimization plan in terms of the ESA label. But it is clear that precise quantitative definitions should be explored in future research.

[Slide]

So, speaking of future research, let's finally again talk about our research program that we believe should inform the appropriate use of ESAs. I won't talk about the TREAT study because that has been, of course, covered by Dr. Pfeffer. We are pleased with the conduct of this trial and we look forward to the upcoming results.

[Slide]

We are interested in three areas, and these areas have also been highlighted by the FDA in their briefing document: hemoglobin target, ESA

responsiveness and hemoglobin cycling. Now, we haven't talked much about hemoglobin cycling today but, again, the FDA briefing document has covered this topic well and we are basically in agreement.

For targets, I think the key question is what do we know about current targets and risk? But current targets, one could mean epoetin alfa labeling of 10-12. This is the range we think, in fact, is appropriate for labeling. Or, one could be describing the nephrology community standards as represented by clinical practice guidelines, basically 11-12. So, for the sake of description I will represent this as 11, but this does not represent a target with an upper end of 11 g/dL as mentioned in the questions today by the FDA.

[Slide]

Because TREAT is the only large-scale double-blind, placebo-controlled trial of anemia therapy in chronic renal failure of this magnitude, it makes sense to factor in those results in terms of picking comparison groups. We would, of course, be looking both at cardiovascular risk and

important secondary measures such as transfusions, exercise capacity and patient-reported outcomes.

Now, on face value, given the apparent risk in Normal Hematocrit and CHOIR, the first choice might seem to be to compare 11 to something lower like placebo or rescue arm, such as is occurring in the TREAT study. But if TREAT shows that 13 g/dL is, in fact, better than placebo in terms of cardiovascular risk, the primary cardiovascular endpoint, one might consider either foregoing additional target studies or, in fact, choosing to look at 11 versus 13.

If TREAT showed neutral results, no discernible benefit on cardiovascular risk but perhaps improvement in transfusions, exercise capacity and patient-reported outcomes then, depending on the magnitude of those secondary endpoints, again, one might consider examining a variety of targets including 11 versus placebo or, in fact, 11 versus 13, depending on the magnitude of the secondary benefits.

Finally, if TREAT shows negative results,

higher cardiovascular risk in 13 versus placebo/ rescue, then the most pertinent question would seem to be 11 versus placebo.

Again, a range of options, each design would need to be a cardiovascular outcome study, clearly probably on the order of the size of TREAT or larger. And, in the next slide we will talk about responsiveness very briefly because we believe that using a run-in phase to calculate ESA responsiveness at baseline and stratifying on that factor could provide useful information on whether hypo-responders have different outcomes by target compared to the others.

I will just briefly touch on this. In our analyses the best definition or the best way in terms of looking at clinical trials that we have found to measure responsiveness is actually with that dose challenge, particularly in EPO-naive patients. So, we believe that a run-in period with a dose challenge is the most feasible in a non-dialysis patient population, and when you think about hemoglobin target trials, again, that placebo

arm or placebo rescue, as Dr. Nissenson has spoken toB-we have been in contact with many individuals and dialysis provider organizations within the nephrology community and we strongly believe that a placebo or rescue arm or, frankly, anything at 10 or less from a target perspective would be difficult to enroll from a feasibility perspective. So, we believe that a non-dialysis patient population with assessment for ESA responsiveness at baseline in terms of a target trial would be the best way to go. And, we are, of course, open and very interested in receiving information and feedback from the panel.

[Slide]

Now, in terms of responsiveness itself, hypo-responsiveness, the key question, outlined by both the FDA's and the sponsor's briefing documents, is whether the hypo-responsive patients would have a better clinical outcome with a different dosing strategy than what is currently employed. In sum, if you have a patient with a low hemoglobin value who is not responding, should you

increase the dose? We believe that that will increase hemoglobin values and, in fact, with successive increases in hemoglobin or an ESA dose you do see a rise in hemoglobin values. That may rise to a level that has been associated with better clinical outcomes. Or, should you maintain the dose and accept a lower hemoglobin value that has been associated with worse outcomes? Those issues are issues of confounding. Or, should you, in fact, be concerned about dose toxicity and reduce the dose, knowing that hemoglobin will likely fall further?

It is important to point out that there is no clear evidence that reducing the dose in these patients will improve clinical outcome. So, one may actually want to investigate multiple arms in the study. But what I have represented here--and these are simply concept ideas and, again, we would appreciate discussion during the panelB-is to identify patients who are hypo-responsive and simply randomize them to current management on label, or a dose reduction. Again, given the issue

of potential equipoise in terms of dosing, one could consider a more aggressive dosing arm as well.

The primary endpoint, again if we are talking about the risk that has been apparent in some of these trials, we are talking about cardiovascular outcome studies. So, something like time to all-cause death or first non-fatal cardiovascular event, again with secondary endpoints of transfusion, patient-reported outcomes, etc.

[Slide]

Finally, I am just going to note that we have been interested in hemoglobin cycling and we believe that it is possible, using a hemoglobin endpoint-based study, to examine on-label dosing versus some kind of alternative graded dosing approach, and we can certainly speak to this in the open session.

[Slide]

Finally, we have covered a lot back from the break. We have covered pivotal trials which

clearly establish the benefit using hemoglobin targets to guide ESA dosing in a patient population that has an amazing unmet medical need. Over 15 years of comprehensive practice data has seen a decline in mortality over the ESA era. And, we recognize the risk of transfusions that still exists in these patients, particularly when hemoglobin values are maintained at lower levels. The higher hemoglobin target trials have shown an apparent increased risk and have been incorporated into ESA labeling. Finally, we have a variety of interests in terms of future research programs. We intend to discuss that further at the open session.

Now I would like to turn the podium over to Dr. Eisenberg who will speak to the risk management program.

Risk Management Program

DR. EISENBERG: Thanks.

[Slide]

Clearly, we have heard a fair amount today. We have some randomized, controlled data from two studies both of which, in fact, were

stopped at the interim and represent a level of evidence that needs to be considered, particularly because it is evidence of risk and that is critical.

[Slide]

We have to look at the practice to date and we have to think about appropriate risk management for patients treated today. Then we need to ask the important question. We have seen some interesting questions raised about target, responsiveness, cycling. What is the appropriate way to evaluate them? And we are delighted actually to have the opportunity to talk about that this afternoon in the specific questions from the FDA that will allow us to address that.

[Slide]

I first want to talk briefly to the issue of target. FDA has posed this question separately for patients on or not on dialysis. I want to address the dialysis patients first. Clearly, the data there are richer both in terms of the comprehensiveness of the observational data and

clinical practice data, as well as the pivotal trials which were clear in demonstrating transfusion reduction, improvements in physical and cardiovascular performance. And, the comprehensive clinical trial databases strongly suggest that the target ranges running roughly from 10.7 to 12.7, 10.5B-you can take a number but it is roughly in this range, confer the greatest benefit for patients on dialysis.

Now, our label has recommended prior to the recent changes, and we still recommend that the target range be between 10-12 g/dL. We believe that means a target not greater than 12 g/dL. We believe as well that if one considers the apparent riskB-these are the point estimates. Let's keep in mind the confidence intervals. These trials were stopped. We do not believe this is overwhelming evidence of risk and certainly will not trump the data that we will be getting from TREAT, which already has accumulated greater experience than these two trials combined. We believe those data suggest that an upper target of 12 adequately

protects patients today both in dialysis and, with respect to the second question, in the patients with non-dialysis chronic renal failure. Clearly, once one gets below 10 there is transfusion-related risk. There seems to be little dispute around this issue.

[Slide]

So, from our perspective, with regards to the question of hemoglobin target, we believe the target of 10-12 with an upper limit of the target, not a ceiling, is appropriate. Targeting 11 g/dL, you have seen the dose variability.

[Slide]

You have seen the risk. It will increase transfusions based on our data. There are no randomized, controlled trials to test regimen or the dosing algorithms around it. We are strongly committed to physician and patient education. We have communicated and we are happy to discuss further, but you have seen the communications that promptly occurred on all the clinical trial data as soon as they were available. We have only

communicated 10-12 as has been in our label for the last decade. In addition, we are working on strengthening recommendations with FDA related to the black box warning recently introduced in the label to clarify that, and introduce patient education for medication guides and other means. I won't comment further, since Dr. Klassen has commented, that we are committed to additional randomized, controlled trials. We believe that should be informed by TREAT.

[Slide]

With regards to ESA responsiveness, I guess after listening to all the ten years of research and multiple analyses, if I say "confounded" one more time you will probably shut your ears. So, it is challenging. We know that patients who don't respond have an increased risk of outcomes that are negative. We also know that they have poor underlying health status. We have had in the label for a number of years the concept that if a patient is recognized to be hypo-responsive to evaluate them. We believe that

needs to be strengthened. I think our thinking around this has continued to evolve. The CHOIR data, the post hoc analysesB-Dr. Singh cited one and we will hear more about that. Dr. Unger will provide some analyses. All of these suggest we should be thinking about this. We should be increasing the awareness for labeling and patient education. But we think that this needs to be approached in a way that is individualized to the patient.

So, concepts that we have been working on and will continue to refine, and we are interested in any additional input or thought, are that what we believe that what is appropriate is to define hypo-responsiveness by the patient's response. That is not dissimilar to how we define poor response. In a diabetic, someone with hypertension, hypercholesterolemia we look at a patient's response.

In terms of ESAs, it is generally I think reasonable to assess that response over a period of about 12 weeks. We always titrate to a target.

So, if we don't reach that target over that 12-week period we think that is a useful working definition. There are others out thereB-the Drueke guidelines suggest another. We are not wedded to one or the other but we think this is reasonable and implementable from a label perspective, and would include the guidance that already exists on reversible causes.

But, more importantly, we would in fact recommend that in that patient who is identified in this matter using the lowest dose to maintain a stable hemoglobin, in other words, not continuing to push the dose, seems prudent based on the data that we have seen and the analyses we looked at recently. But, of course, this should be periodically reassessed since the patient's health status does change over time.

[Slide]

Finally, the cycling question, and since we haven't discussed it Dr. Unger may touch upon this. It was in the briefing book. It is an important concept. There is this concept that if

you make frequent dose changes or even possibly due to health status changes there may be relatively rapid increases or decreases in hemoglobin. We are not entirely sure how one implements management of that, other than to believe that the way to do this is not to overact to minimal changes, minimal excursions above or below a target range; that one should manage those excursions with dose reductions, not dose cessation or concern since there is no evidence that a transient achievement of a hemoglobin above 12 g/dL confers risk. There is no evidence of that.

So, we will conclude by saying we believe that if we target 10-12 g/dL this is an appropriate approach for risk management. This is consistent with the safety and efficacy that has been demonstrated for ESAs in this range over a decade. It is guided by randomized clinical trial evidence and comprehensive experience in the nephrology community.

I want to thank the committee for your time and attention, and we look forward to the

discussion this afternoon.

DR. PLATT: Thank you very much, Dr. Eisenberg. We are switching now to presentations by the FDA and the first presenter is Dr. Trentacosti.

FDA Presentations

Epoetin alfa: FDA Overview of Patient-Reported Outcome, (PRO) Claims

DR. TRENTACOSTI: Good morning.

[Slide]

My name is Ann Marie Trentacosti and I am a study endpoints and labeling reviewer at the Office of New Drugs at the FDA. I will be discussing the FDA review of the patient report outcomes, also called PRO claims, in the epoetin alfa label.

[Slide]

The term PRO as used in this presentation refers to all claims which describe how epoetin alfa makes patients feel or function better. Although in a few instances the instruments used to justify the claims are physician assessments, the

majority are patient report outcomes and, for simplicity, all the instruments and claims considered in this presentation will be considered PROs. The epoetin alfa PRO claims include performance, symptoms of anemia, al and health-related quality of life.

[Slide]

In March, 2007 the FDA sent Amgen a letter describing a post-marketing commitment. Amgen was requested to reevaluate the data used to support the PRO claims in the epoetin alfa label. This reassessment was to be consistent with the principles found in the FDA draft PRO guidance. The draft PRO guidance provides recommendations for sponsors in the development and validation of patient report outcome measures, and the incorporation of those measures into clinical trials to support labeling claims. Based upon this re-analysis, Amgen was requested to provide additional recommendations for labeling revisions.

[Slide]

After review of the information submitted

by Amgen as of July, 2007, the FDA has determined that based upon the principles found in the draft FDA PRO guidance, and the review of the PRO instruments and the clinical studies submitted, the PRO claims in the epoetin alfa labeling are not adequately substantiated.

The FDA conclusions have been deemed preliminary, however, because Amgen continues to submit additional information that has not been completely reviewed at this time. My brief presentation will not summarize the FDA review of all of the information submitted by Amgen. Instead, I will focus on only a few of the major deficiencies identified to reach this preliminary conclusion.

[Slide]

First I will discuss the draft FDA PRO guidance and explain some important concepts which will aid in the review of this material.

[Slide]

The following are some important definitions from the PRO guidance which may add in

this discussion. The draft PRO guidance was developed in February, 2006. This publishing of data occurs many years after the epoetin alfa clinical trials submitted to support the PRO claims were initiated, which is during the 1980s and early 1990s. The development of the PRO guidance was based upon the principle that in defining clinical benefit of a treatment it is extremely important to define how a patient feels or functions as a result of treatment, essentially the patient's experience.

In addition, consistent with the FDA's regulations for medical product approval, assessments of clinical benefit must be well defined and reliable. In the case of PRO claims the assessments must be well developed and adequately validated to measure what they are intended to measure. The draft guidance describes how the FDA evaluates patient report outcome instruments when used as efficacy endpoints in clinical trials, and provides recommendations for developing and validating these instruments in order to support labeling claims. The final

guidance will provide additional clarity but will not change in terms of the basic principles for instrument development and validation.

[Slide]

The following are some definitions from the guidance. A patient-reported outcome or PRO is any measurement that is reported as a direct response by patients without any interpretation by anyone else. For example, asking patients to rate their pain severity on a scale of 0-10. A PRO assessment is different than a physician outcome assessment in which physicians rate their patient's function on the basis of an aspect of a condition or disease. An example of a physician outcome is the New York Heart Association classification for congestive heart failure.

[Slide]

Quality of life is a general concept that measures the impact of all aspects of life on general well-being. The term implies the evaluation of non-health related aspects of life such as economic status. For this reason, quality

of life instruments are not health specific and cannot be used to support labeling claims.

On the other hand, health-related quality of life implies an evaluation of the patient's overall perception of the impact of an illness and its treatment. Even though health-related quality of life is a very complex concept, it is limited to the physical, psychological and social functioning related to the health of the patient and may be utilized as an efficacy assessment in a clinical trial. It should be noted that quality of life is often used as a generic term to refer to quality of life, other PRO measures or health-related quality of life rather than referring to the specific definition described.

[Slide]

Content validity is evidence that the items and domains of an instrument are appropriate, comprehensive and interpretable relative to its intended measurement concept, population and use. Content validity is established by documentation of patient input that confirms that the concept

measures are important, meaningful and well defined. For example, physical functioning or performance is a complex concept. In order to adequately develop and instrument to measure physical functioning input from the patient population of interest is important to ascertain what clinically meaningful items should be included.

A physical functioning questionnaire for patients with rheumatoid arthritis would require different items compared to a physical functioning questionnaire for patients with congestive heart failure. The rheumatoid arthritis questionnaire would probably contain items referring to the ability to open doors or open jars, while the questionnaire for congestive heart failure patients may contain items referring to ability to walk distances or climb stairs. Although both questionnaires measure physical functioning, based upon patient interviews each questionnaire should be tailored to its specific target population and indication.

[Slide]

The measurement properties of an instrument assess the instrument's ability to measure a concept. These properties include content validity, construct validity, reliability and the ability to detect change. The content validity is the first measurement property which should be evaluated before the others. Essentially, you just understand fully the concept you are trying to measure from the patient before you test the instrument's ability to actually measure it.

[Slide]

In summary, several important concepts from the PRO guidance are important to consider for the remainder of this presentation. PRO instruments should be developed rigorously with clear documentation, similar to any other measurement of effective. In order to adequately capture the patients' experience, direct input from the patients themselves is vital. This input is necessary to fully identify the concepts of

interest and ascertain the validity of the instrument.

Once the instrument, such as a questionnaire or diary, has been created patient input is also necessary to ascertain that the instructions, questions and response options are understandable. Once an instrument has been developed and validated, any major alteration or modification, such as using an instrument in an entirely different population, will require a reevaluation of its measurement properties. In addition, any study that uses a PRO measure must be adequately designed, including blinding.

[Slide]

Next I will discuss the PRO claims in the epoetin alfa label.

[Slide]

The approved epoetin alfa label contains the following information, once the target hematocrit was achieved, statistically significant improvements were demonstrated for most quality of life parameters measured, including energy and

activity level, functional ability, sleep and eating behavior, health status, satisfaction with health, sex life, well-being, psychological effect, life satisfaction, and happiness. Patients also reported improvement in their disease symptoms. They showed a statistically significant increase in exercise capacity, energy and strength with significant reductions in aching, dizziness, anxiety, shortness of breath, muscle weakness, and leg cramps.

Many of the claims in the label are quality of life as opposed to health-related quality of life claims, such as well-being, life satisfaction and happiness. Amgen and FDA both agree that these quality of life claims should be removed from the label.

[Slide]

The PRO claims which Amgen has proposed to retain in the label can be summarized into two categories, improvement in physical function and activity level and improvement in anemia symptoms of decreased energy, muscle weakness and shortness

of breath.

[Slide]

Overall, the FDA issues identified from the information submitted can be categorized into two categories. The first category represents the design inadequacies of the clinical trials. Essentially, the clinical trials which support the claims were not designed to measure health-related quality of life or symptoms of anemia. The second category is the problems with the instruments. The instruments used in these trials were not developed or validated to measure the clinical benefits implied by the labeling claims.

[Slide]

In discussing the FDA's review of the PRO claims, I will first provide a brief description of the clinical trials which support the claims and a few of the major deficiencies.

[Slide]

Four clinical study reports were submitted to support the PRO claims. Study 8601 was an open-label, single-arm study. The data from this

open-label study provide the support for the PRO claims which were originally placed in the epoetin alfa label. This study included an acute treatment phase and maintenance phase. PRO questionnaires were obtained at baseline, once the target hematocrit was reached and three to four months after the target hematocrit was reached.

Studies 86701 and 8904 were randomized, double-blind, placebo-controlled, partial crossover studies designed to evaluate the safety and efficacy of epoetin alfa in ameliorating anemia and reducing or limiting transfusions in patients undergoing dialysis. Both studies were similar, however, study 8701 enrolled hemodialysis patients where 8104 enrolled peritoneal dialysis patients.

The studies included a 12-week treatment period in which patients were treated with either placebo or epoetin alfa. After the treatment period all patients entered a 12-week open-label period. PRO questionnaires were obtained at baseline, 12 weeks and 24 weeks. The 24-week assessment, therefore, occurred during the

open-label period.

Study EP 86-004 was a randomized, double-blind placebo-controlled study. Patients were randomized to receive either placebo or epoetin alfa doses to maintain hemoglobin levels at one of two predetermined ranges, a medium or a high range. Study drug was administered for 26 weeks. Efficacy was evaluated by the ability of study drug to raise hemoglobin to a predetermined level and by correlation of hemoglobin to quality of life. PRO assessments were obtained at baseline, two, four and six months.

[Slide]

The studies submitted were not adequately designed to evaluate PRO claims. The primary study which supported the claims was open-label in design. The other two studies had crossover design and were, therefore, partially open-label. The studies didn't include a prospective statistical analysis plan and a proposal for handling missing data and multiplicity. In three of the four studies, in order to be considered an evaluable

patient questionnaires from all three time points had to be completed. Therefore, patients who withdrew from the study or did not complete all three questionnaires were excluded from statistical analyses. Most of the analyses which provide justification for the claims were obtained post hoc rather than a priori. All of the studies enrolled patients based upon their hemoglobin or hematocrit and not based upon symptoms. There was no demonstration of a correlation of a hemoglobin or hematocrit with anemia symptoms in any of the studies.

[Slide]

Next, I will discuss the issues concerning instruments used in the clinical studies.

[Slide]

In discussing the instruments I will focus on only two examples, the Karnofsky Performance Scale and the National Kidney Dialysis and Kidney Transplantation Study Symptom List. These instruments were two of the main instruments upon which the PRO claims rest, and the problems

illustrated for these instruments are representative problems across instruments used.

[Slide]

The Karnofsky Performance Scale was developed in 1949 to evaluate performance status in cancer patients who were undergoing chemotherapy. The KPS was designed as a physician assessment and not a patient-reported outcome. As part of the KPS, physicians rate their patient's performance on a scale of zero, which is death, to 100, which is normal.

In open-label study 8601 the KPS was administered as a physician assessment. However, this scale was modified and used as a patient-reported outcome in two of the other studies. Information has not been submitted to suggest that the KPS can be administered adequately as a PRO measurement. At face value, it is uncertain how a patient can understand or appraise their own performance status based on response options provided.

The terms to describe each score are

non-specific and ambiguous. For example, a score of 50, "requires frequent medical care," could be interpreted to mean frequent dialysis treatments, frequent physician visits or frequent hospital visits. There are no specific criteria associated with each score to help physicians rate performance which can lead to arbitrary scoring.

Two subjects with similar physical ability may be rated with different KPS scores based upon the interpretation of normal activity or work. For example a heavy machine operator who informs his physician that he is unable to perform his job might receive a score of 70, which is unable to carry on normal activity or work. While a desk worker with similar physical ability who states that he can perform his job might receive a score of 90, "able to carry on normal activity."

We conclude that the instrument is inappropriate and uninterpretable as utilized in the clinical studies provided. In other words, it lacks content validity for the target population.

[Slide]

The National Kidney Dialysis and Kidney Transplantation Study was developed in 1987 to compare the costs the of treatment and outcomes in patients with end-stage renal disease. The patients in the study were not necessarily anemic or in renal failure since the study included renal transplantation patients. The NKDKTS symptom list used in the study was derived from a 1978 survey of disability and work. The sponsor added some items from the original list to create a revised list which they subsequently used in several of their epoetin alfa clinical trials.

In the epoetin alfa clinical trials subjects were asked if in the past two weeks they had any of the listed symptoms or health problems. Their response options were often, sometimes, rarely and never. Although the entire list was administered during the clinical trials, the sponsors selected specific items, which are highlighted in clue, that they believe represent anemia symptoms and which they propose to retain in the epoetin alfa label. These items were

identified post hoc. As specified in the FDA PRO guidance, adding items to a questionnaire and plucking out individual items for a longer questionnaire is not acceptable without formal reevaluation of the modified instrument.

The generic NKDKTS symptom list is not designed to measure any specific concept, to be a complete list of important anemia symptoms or to be used to derive a total score. The list was not designed or developed as an instrument to measure anemia symptoms and is, therefore, not acceptable as an endpoint to support labeling claims.

[Slide]

The instrument deficiencies which are noted in the examples provided, as well as the other instruments in the other clinical trials submitted can be summarized as follows: None of the instruments have been shown to have content validity, that is, documentation based upon input from patients that the items and concepts in the instrument are appropriate, comprehensive and interpretable, specific for the target population

and indication.

Essentially, none of the instruments were developed for the target population or indication, which is the measurement of performance, health-related quality of life or symptoms associated with the treatment of anemia in chronic renal failure patients. Many of the instruments were generic. They were developed to evaluate health problems in any population or, as in the case of the Karnofsky, developed for cancer patients, a different population. In order to justify anemia symptom claims, items and subscales were chosen and analyzed post hoc.

[Slide]

In summary, the clinical studies provided were not adequately designed to support the PRO claims in the current labeling or to measure health-related quality of life or anemia symptoms. The instruments used in the clinical trials submitted are not adequate measures of anemia or health-related quality of life for the target population or indication.

[Slide]

In conclusion, based upon the data submitted as of July, 2007 and the principles delineated in the draft PRO guidance, the FDA preliminary findings suggest that clinical benefit or epoetin alfa in the improvement of patient performance, anemia symptoms or health-related quality of life has not been adequately established.

DR. PLATT: Thanks very much. Dr. Unger?

FDA Perspectives on Erythropoiesis-Stimulating Agents Anemia of Chronic Renal Failure: Hemoglobin Target and Dose Optimization

DR. UNGER: Good morning, everyone.

[Slide]

I am Dr. Ellis Unger. I am the Acting Deputy Director for Science for the Office of Surveillance and Epidemiology in CDER, and I am very pleased to be here this morning to present FDA's perspectives on erythropoiesis-stimulating agents for anemia of chronic renal failure, with emphasis on hemoglobin target and dose

optimization.

[Slide]

So, I am going to spend about half my time talking about hemoglobin target for anemia of chronic renal failure. I will talk about the randomized, controlled clinical trials, albeit very briefly because you have heard about them three times already this morning. I will talk about the observational data. Then I am going to spend time talking about ESA responsiveness, dose optimization challenges and then discuss a potential path forward.

[Slide]

So, the ESAs as you know, at this point are epoetin alfa and darbepoetin alfa. These are the only two agents approved in the United States for this indication. You need to understand that we are talking about these agents as a class, although the data I will discuss in the next 20 minutes are basically generated on epoetin alfa only.

[Slide]

In terms of the correct hemoglobin target, the randomized clinical trials.

[Slide]

These are the discrete hemoglobin target trials that you have heard about, the Normal Hematocrit, CHOIR and CREATE.

[Slide]

The Normal Hematocrit study goal was to assess the risks and benefits of achieving a normal hematocrit in hemodialysis patients with clinically evident congestive or ischemic heart disease. It was conducted a decade ago. You have heard about this.

[Slide]

Again, I will go quickly. Patients were randomized 1:1 to the low target or the high target, 30, plus/minus 3 or 42, plus/minus 3 respectively. The patients had clinically evident ischemic heart disease or heart failure. They were on hemodialysis, that is a key point. And they were clinically stable at the time they were randomized. The primary endpoint was time to death or

non-fatal myocardial infarction.

[Slide]

Randomized 1:1, 634 to 631. As you have already heard, the trial was terminated early. I won't read this because you have heard it and read it.

[Slide]

Indeed, patients randomized to the normal hematocrit group did achieve on average the higher target versus the lower target for the low hematocrit group.

[Slide]

This is the important slide that shows the Kaplan-Meier for probability of death or non-fatal MI. You have seen this. The results were unfavorable for the higher target.

[Slide]

This breaks down the primary composite endpoint by death and non-fatal MI. It is important to see that it is basically driven by death and not non-fatal MI.

[Slide]

This is a very complicated slide and I will spend some time on it. There is a very similar slide in the *New England Journal* paper by Besarab. What this shows is a negative association between the mean hemoglobin throughout the study and mortality. So, a lower hemoglobin is associated with higher mortality. So, in this analysis, and this was an FDA analysis, the mean hemoglobin was calculated for each subject in the study throughout time and the subjects were divided into quintiles irrespective of whether they were in the lower target or the higher target. The Ns are shown here. So, these are the quintiles from low to high, lower target, higher target.

You can see in the lower target group that the Ns are here and patients are over-represented in the lower hemoglobin quintiles because they were assigned to a lower target. Conversely, patients in the higher target are over-represented in the higher quintiles. Then, these subjects are very under-represented. There are only three subjects here. So, what you see both for the lower target

and the higher target is this negative association between the mean hemoglobin achieved and mortality.

Now, I need to make sure you grasp a couple of things about this slide. Look at the mortality rate here. These aren't cancer patients. You know, these aren't class IV heart failure patients. These are patients with chronic renal failure on hemodialysis and the mortality rate in this groupB-it is an isolated group, these These are patients who were assigned to a higher target and didn't achieve much of any response. Their hemoglobin is in the lowest quintile. Their mortality is 77 percent. That is startlingly high.

Although I am not a nephrologist and don't treat these patients, I am a cardiologist and I feel qualified to make this statement, and that is that you don't know who you hurt. Unlike other situations where you are looking at risk and benefit of a drug, you know what the benefit is. The risk is often in an entirely different organ system. It may be cardiovascular risk for an NSAID. It may be hepatic toxicity for some other

kind of drug. Here, these patients die because of their cardiovascular disease and the doctor, you know, looks at the patient and says, well, you know, he died; he had a stroke. And, we know he had bad coronary disease and bad cerebrovascular and it is not a surprise. We don't know who we hurt in this patient population.

[Slide]

So, the summary for this study, these were hemodialysis patients with clinically evident congestive heart failure or ischemic heart disease. Targeting the higher hematocrit of 42 was associated with increased mortality and, somewhat paradoxically, the higher mean hemoglobin concentrations were associated with survival in both treatment arms.

[Slide]

CHOIR, you have heard about. In retrospect, this looks like a brilliant idea. These were patients who were not on dialysis so, basically, these two studies peg the ends of the spectrum for patients with chronic renal failure.

The Normal Hematocrit study included, patients on hemodialysis with clinically evident heart disease; these patients in CHOIR were not on dialysis yet.

Here the patients were randomized 1:1 to a target of 11.3 or 13.5. The primary endpoint was a composite of mortality, congestive heart failure hospitalization, non-fatal stroke, or non-fatal MI.

[Slide]

Randomization was 1:1, and you have heard that this was also stopped at an interim look because of an unfavorable outcome. Indeed, patients assigned to the higher hemoglobin group tended to achieve the higher hemoglobin, although not so much here possibly because, as you heard from Dr. Klassen, there was a cap at 20,000 units for any given dose.

[Slide]

Be that as it may, when you look at the probability of a composite event, it trended unfavorably to the high hemoglobin group, as you have seen.

[Slide]

If you break down components of the primary composite endpoint, mortality was much lower in this patient population. Again, these are pre-dialysis patients so mortality was 5.5 percent versus 3.6 percent. Congestive heart failure hospitalization, 8.3 versus 5.9. And, those were really the two drivers here. Non-fatal MI and non-fatal stroke were similar in the two treatment groups.

[Slide]

Again, if you look at mean hemoglobin for each subject throughout the study, divided into quintiles, and then look at mortality as a function of that hemoglobin quintile you see this negative association between mean hemoglobin and mortality. Importantly, if we looked at the data from these two studies as we look at observational data, if we ignored the randomization and said what is the right hemoglobin target, well, I would say most people would say that higher is better. That is how we would interpret these studies if we didn't know anything about the randomization. So, it is

very important to understand that there is a difference between the achieved hemoglobin and the targeted hemoglobin; it is better to be higher, but versus the targeting a higher hemoglobin carries cardiovascular risk.

[Slide]

So, for CHOIR, again, pre-dialysis patients administered epoetin alfa to a target of 13.5 versus 11.3 was associated with increased mortality and CHF hospitalization. Paradoxically, higher mean hemoglobin concentrations were associated with survival in both treatment arms.

[Slide]

CREATE, I will talk about briefly. It was a much smaller study. These patients had mild anemia. They were also not on dialysis. They were EPO-naive. They were randomized to targets of 13-15 or 11-12.5. I should point out this was epoetin beta which is not approved in the United States but it is one of the ESA class.

[Slide]

The primary composite endpoint, there were

a number of components. You can read them here. I am not going to go through them. About 300 subjects were randomized to each of the treatment arms.

[Slide]

By and large, the patients achieved the targets.

[Slide]

Not that many primary endpoint events, 58/301 in the normal hemoglobin group versus 47/302 in the sub-normal, for a hazard ratio of 0.78. Surprisingly few endpoint events, I would say, given how broad the composite endpoint was. So, the results were not statistically significant but they did directionally support the lower hemoglobin target, like the Normal Hematocrit and CHOIR studies.

[Slide]

You have seen these observational data before from 58,000 U.S. dialysis patients. This is the database from DaVita, a large provider of hemodialysis. It shows a J-shaped curve, here,

with a nadir somewhere between 12 and 12.9. Again, these are associations. The hazard ratio has a ratio across here of 1. You know, for mortality a hazard ratio less than 1 in these two groups but, again, these are associations.

[Slide]

The National Kidney Foundation K/DOQI guidelines say cohort-based observational trials and cross-sectional analyses of large medical databases consistently show that higher achieved hemoglobin values, including greater than or equal to 12 g/dL, are associated with improved patient outcomes. The failure of observational associations to be confirmed by interventional trials renders use of observational evidence unsuitable to support the development of an intervention guideline statement.

[Slide]

If this looks like the Forty-Niners and the Cardinals, you were up too late last night. This is our perspective on hemoglobin targets. So, in the Normal Hematocrit study, you know, the

targets were 10 plus/minus 1, 14 plus/minus 1; CHOIR, pre-dialysis patients, 11.3, 13.5.

[Slide]

So, what we know is that these targets are too high for these respective patient populations. for pre-dialysis 13.5 is too high; 14 is too high for hemodialysis patients with clinically evident heart disease. The observational data show a sweet spot in here. But those observational data are by association only. Exploratory analysis of the Normal Hematocrit and the CHOIR study show the same thing, associations between higher mean hemoglobin concentration achieved and survival but, again, we would be misled if we only looked at hemoglobin versus survival because, you all know, association does not prove causality. Achieved hemoglobin is different from a hemoglobin target. And, that J-shape relation in the observational data suggests that there is, in fact, some hemoglobin concentration that is excessive for this population. DR. PLATT: Dr. Unger, just I am just giving you a time check, five more minutes.

DR. UNGER: That is good.

[Slide]

So, perhaps patients who achieve a higher hemoglobin concentration have less advanced renal disease or lower cardiovascular disease burden. I think, importantly, we are not aware of a randomized, controlled trial that demonstrates in a convincing way that a higher hemoglobin target is associated with less cardiovascular morbidity and mortality than a lower target.

[Slide]

Now I am going to talk about ESA responsiveness.

[Slide]

This has been touched on before. Can we prospectively identify the hypo-responders? Then, if we could, how would we treat them?

[Slide]

I showed you this slide a couple of minutes ago. These are the hypo-responders. Again, they were randomized to the higher target. They achieved a hemoglobin of less than 10.2.

Their mortality was 77 percent.

[Slide]

There is another way to look at responsiveness. For a therapy such as ESAs, the doses are titrated, which confounds anything. So, patients who were least responsive get the highest doses because you keep pushing the dose because the patient doesn't respond, and the dose and responsiveness are inversely related.

This is a startling slide actually. It shows fraction surviving by epoetin alfa dose divided in quartiles. This is for the Normal Hematocrit study, and there is a very clear dose response. Higher dose, lower survival. But, again, this is entirely confounded. It would be like if I showed you results of a loop diuretic in heart failure--higher doses of a loop diuretic, higher mortality.

[Slide]

So, FDA did an exploratory analysis, much the same as what you saw in Dr. Klassen's slide number 56. The Normal Hematocrit study was unique

because all patients were stable and they were randomized to a higher and lower target. Those who were randomized to a normal target had a standard protocol-mandated ESA challenge and their dose was increased by a factor of 1.5 on study entry.

[Slide]

We can calculate epoetin alfa responsiveness in those patients if, in fact, they had received a stable dose or, I should say, a constant dose for two to six weeks following study entry, and if their dose was increased by a factor of 1.5 at entry. And, responsiveness here was defined as the slope of the hemoglobin time relation throughout the two- to six-week period. In the paper presented by Dr. Klassen, I believe that the hemoglobin was also taken into account. Here it is simply the slope of the hemoglobin time relation.

[Slide]

So, there were 618 subjects randomized to the normal target, and we could calculate responsiveness for 414. Interestingly, 117 patients actually experienced a decrease in their

hemoglobin despite their dose increase, and 297 patients experienced either no change or an increase in hemoglobin.

[Slide]

We could look at survival by responsiveness and we could look at the overall responsiveness by initial response.

[Slide]

This shows survival. Here I have narrowed it down to just three groups that I want to show. The dotted line represents the 117 subjects who actually decreased their hemoglobin despite an increase in dose. They fared pretty well. This is mortality. Q1 is the quintile that was the least responsive, in red, and this is the most responsive, in black.

[Slide]

You can see there is no difference there, and this fills in the other three quintiles. So, there is no good relation here.

[Slide]

If we look at overall responsiveness

through the rest of the study, this is the mean hemoglobin throughout the study on average. This is what happened with that 50 percent increase in EPO dose. A unique opportunity and, yet, we don't see any kind of a direct relation here between the initial response and the final response.

[Slide]

So, in the Normal Hematocrit study where patients had a protocol-mandated 50 percent increase in EPO dose, the initial hemoglobin response did not predict subsequent mortality; did not predict the overall hemoglobin response. I should also point out that in the analysis presented by Dr. Klassen there were, I believe, 14 variables that were considered in the analysis and this was not corrected. So, ESA responsiveness may need to be assessed on an ongoing basis.

[Slide]

Here are data from an individual patient in the Normal Hematocrit study. Hemoglobin is shown in red and the dose is shown in black. Hemoglobin starts out about 8 and at that point

flattens out around 10, and doesn't move even though the target is up here between these two dotted lines. The dose is taken up to 120,000 units/week. This patient is definitely not responsive.

[Slide]

So, prospective identification of hypo-responders may be difficult, although it seems to be feasible in practice.

[Slide]

For hypo-responsive patients the labeling suggests a search for causative factors, but does not explicitly state a maximum ESA dose, or what constitutes an adequate attempt to raise hemoglobin. And, the key unanswered question is whether less responsive patients or those with specific risk factors would experience fewer events if they were not made to raise their hemoglobin to some ideal target.

[Slide]

Last, I am going to talk about dose optimization and I will try to go fast.

[Slide]

The label warns against excessive rate of rise, greater than 1 g/dL per 2 weeks. Is the risk related to hemoglobin response? Well, it makes sense that it is. If it is better to have a higher hemoglobin but it is dangerous to be targeted to a higher hemoglobin, then maybe there is something about the behavior of hemoglobin or maybe the behavior of the dose that is administered that contributes to risk.

[Slide]

Here is data from another patient from the Normal Hematocrit study. Again, hemoglobin is shown in red and the dose, you see, has jumped all over the place, in black. This is pretty classic cycling where the hemoglobin is going in and out of target. You know, it is overshooting; it is undershooting. The target is these dotted lines, and the dose is going up and down.

Now, if risk is related to the hemoglobin level, then it doesn't make sense that a patient's risk is constant throughout time. It makes more

sense that the risk is a function of what the hemoglobin is and maybe what the rate of rise of hemoglobin is.

[Slide]

So, if you focus on the circled area there, we can look closely at that and we can say, look, this is week 49 and the hemoglobin is 9.3. The next week hemoglobin is 9.9.

[Slide]

There is a certain rate of change associated with this week. Then we can marry these data with the adverse event data and we can associate adverse events to specific weeks. So, this is an unorthodox analysis where each week for each patient represents a time at risk and all the patients are combined.

[Slide]

If you do this analysis and look at serious cardiovascular events per patient-year versus hemoglobin rate of change versus hemoglobin you can see that the higher hemoglobin, which is in the front in red here, tends to be associated with

fewer adverse events. Okay? But the rates of change seem to be strongly associated with events. So, here are positive rates of change. These are hemoglobin increases. If you have a low hematocrit or a low hemoglobin and you have a rapid rate of increase you have a high event rate. Conversely, if you have a rapid rate of hemoglobin decrease you have a high event rate. So, it seems best to be in the middle here in terms of hemoglobin and to not cycle, although, again, these are just associations.

[Slide]

So, the ESA labeling warns against excessive rate of rise, and the oscillations appear to be associated with serious adverse events. The question is which came first, the chicken or the egg? It is a classic issue. Is it due to some underlying patient characteristic or is it worth trying to prevent?

[Slide]

You can actually design an algorithm or algorithms to try to limit cycling. It is doable.

So, you can have an algorithm where you put the most recent hemoglobin here and you factor in what the rate of change is, and you go to your Palm Pilot, the Internet, or whatever, and it tells you what to do with dose. This is fairly complex maybe, but it is doable. Using an algorithm like that, you can prevent excessive rates of rise; prevent overshoot, and perhaps provide appropriate means to identify and treat hypo-responders.

[Slide]

In summary, the best randomized, controlled data available suggest that the ideal hemoglobin target is 10, plus/minus 1 for hemodialysis patients and 11.3 for pre-dialysis patients. The data to support a hemoglobin target as high as 12 are observational in nature and of limited utility because association is not the same as causality, and achieved hemoglobin is not the same thing as target hemoglobin. It is unknown if ESA-hypo-responsive and/or high risk patients should be treated differently. And, we have little data to show that current labeling addresses how

best to reduce hemoglobin overshoot and cycling.

I thank all of you for your attention.

DR. PLATT: Thanks so much.

DR. UNGER: Excuse me, I forgot one. I will do it quickly.

[Slide]

So, the potential path forward for the target would be to conduct prospective, randomized, controlled cardiovascular outcome studies to determine optimum hemoglobin targets; consider a priori that disparate targets might be based on risk factors. I am very happy to hear that Amgen appears to be interested in conducting these. Amgen also mentioned the development of new dosing paradigms. They seem to be open to that. Special dosing strategies might be considered for hypo-responsive patients and those at higher risk of cardiovascular events, and the strategy could consider futility. Both of these would need to be tested in randomized, controlled trials. Thank you.

Questions to Presenters

DR. PLATT: Thanks. We have 25 minutes to begin the committee discussion. There are 19 committee members sitting around the table. The lieutenant commander is pointing out to me that these are questions; we will have discussion later. In my experience, they tend to blend some. I assume we will have the opportunity to ask questions after. But let me just get a sense of how many questions there are now. Show of hands from those who will ask questions?

[Show of hands]

Ten. So, that is two minutes apiece. I assume we have some latitude in the way we use this time. It is obvious we are not going to be able to have ten people ask questions and get answers. So, our choices are we can start the questions and then have them continue after the public comment session, or we could try to get the questions on the table with the idea that we will pick up all the discussion later. Show of hands, how many people would like to sort of state the questions and get them out, with the understanding that we

will pick they up later, and how many people would like to work through as many questions as we can now? For get the questions on the table? For churn through as many as we can? Churn through gets it. Dr. Phan, did you get a good look at who has questions? Hands up who have questions. Keep your hands up so Dr. Phan can do that. We will go clockwise, so while she is taking a list, Dr. Hennessy, will you start us off?

DR. HENNESSY: Thank you, Dr. Platt. My question is for Dr. Zhang. From slide 39, it looks like the lowest mortality appears to be somewhere between 10,000 and 12,500 units/week. So, my first question is have you looked to see whether expressing the doses as units/kilogram of body weight or units/meter^2 is a better predictor of mortality?

My second question is, is there effect modification by the dose required to get to a specific hemoglobin? I would point out that in slide 39 it looks like 10,000 to 12,000 units/week produces an average hematocrit of about 35 or 36 or

an average hemoglobin of about 12. Given this, it seems like a hemoglobin target of 10-12, as is currently labeled, is consistent with the lowest mortality, and I wanted to know if you agree with that. I will stop there.

DR. PLATT: Dr. Zhang, please?

DR. ZHANG: Right. So, your first question is regarding how we can measure an EPO dose and how that relates with mortality so that dose/kg, you say, if that is a better predictor compared to dose/month. I think in actual practice physicians actually prescribe EPO not based on patient weight. The majority only look at hematocrit values. So, I think it doesn't matter how you are going to measure that because patient weight now, in current practice, is not actually a factor.

DR. HENNESSY: Even if it is not used as dosing, couldn't that be a better predictor of mortality than just number of units?

DR. ZHANG: In terms of number of units, when we calculate based on our data what we are finding is that if we measure dose/month and

compare it to dose/kg per week, it is basically the same thing. If you take an average population weight as 75 or 70 you get very similar results. So, I don't think weight really matters.

You second question is?

DR. HENNESSY: Whether I was reading your graphs correctly. Actually, the second question was did you look for effect modification by the dose it took to get to a particular hemoglobin and, if not, are you able to do that?

DR. ZHANG: I am sorry, I didn't quite get the question. Are you talking about our survival findings or dose-response findings? We have different findings.

DR. HENNESSY: Sure. I mean the survival findings.

DR. ZHANG: Right. So, we are showing that approximately 8,500 to 15,000 units is associated with lower mortality risk, which is moderate EPO doses if we look at population dose distribution. But that value is slightly different from the dose required to reach the population average hematocrit

of 36 percent. At that point, I think we need to clarify the difference between two studies.

DR. PLATT: Thanks. Dr. Kopp?

DR. KOPP: I have a question for Dr. Eisenberg. On slide CC17, the last bullet says target hemoglobin range of 10-12 is recommended. Then there is a sub-point, managing risk through achieved hemoglobin ceiling of 12 is not consistent with the results of RCTs. Could you clarify what you are proposing, if there is a target range of 10-12, to do with patients who exceed 12? Should the dose be reduced or should they be allowed to stay at 12.5?

DR. EISENBERG: No, I think our recommendations are very clear. We believe that the target is 10-12 and if someone exceeds that 12 range the dose should be reduced. That is actually what has been in the label and we should get those patients back under 12. We think that is the right way to manage patients.

DR. KOPP: Since that was so quick can I have another question? I didn't hear any talk