DEPARTMENT OF HEALTH AND HUMAN SERVICES
UNITED STATES FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH

Arthritis Drugs Advisory Committee Meeting

Monday, November 24, 2008

8:30 a.m.

Hilton Silver Spring
8727 Colesville Road
Silver Spring, MD

PAPER MILL REPORTING
(301) 495-5831

C O N T E N T S

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

Kathleen O'Neil, M.D., Acting Chair
Nicole Vesely, Pharm.D., Designated Federal Official, AAC

ARTHRITIS ADVISORY COMMITTEE MEMBERS (Voting)

Diane Aronson (Consumer Representative)
Nancy Olsen, M.D.
Kathleen O'Neil, M.D. (Acting Chair)
Robert Stine, Ph.D.

INDUSTRY REPRESENTATIVE (Non-Voting)
Mark Fletcher, M.D.

TEMPORARY VOTING MEMBERS

Daniel Clegg, M.D.
John Cush, M.D.
Curt Furberg, M.D., Ph.D.
Allan Gibofsky, M.D., J.D.
Stephen Glasser, M.D.
Robert Harrington, M.D.
Sean Hennessy, Pharm.D., Ph.D.
Suzanne Lindley (Patient Representative)
Tuhina Neogi, M.D.

TEMPORARY MEMBER (Non-Voting)

Milton Packer, M.D.

FDA (Non-Voting)

Curtis Rosebraugh, M.D.
Bob Rappaport, M.D.
Jeffrey Siegel, M.D.
Jane Gilbert, M.D., Ph.D.

P R O C E E D I N G S

Call to Order

DR. O'NEIL: Good morning, everyone. We would like to welcome you to this meeting of the FDA Arthritis Advisory Committee. My name is Kathleen O'*Neil and I am the acting chairman of this meeting.

For topics such as those being discussed at today's meeting there are often a variety of opinions, some of which are quite strongly held. Our goal in today's meeting will be a fair and open forum for discussion of these issues and that individuals can express their views without interruption. Thus, as a gentle reminder, individuals will be allowed to speak into the record only if recognized by the chair. We look forward to a productive meeting.

In the spirit of the Federal Advisory Committee Act and the Sunshine Act, we ask that the advisory committee members take care that their conversations about the topic at hand take place in the open forum in the meeting. We are aware that members of the media are anxious to speak with the FDA about these proceedings. However, FDA will refrain from discussing the details of this meeting with the media until the meeting's conclusion. Also, the committee is

reminded to please refrain from discussing the meeting topic during the lunch break. Thank you.

Introduction of Committee

DR. O'NEIL: We would now like to introduce the members of the committee. I will ask each member to introduce himself with his name and institutional name. I will begin with Dr. Fletcher.

DR. FLETCHER: Good morning. I am Mark Fletcher. I am the industry sponsor for the Arthritis Advisory Committee, and I am from Pfizer.

DR. PACKER: I am Milton Packer, from the University of Texas Southwestern and I am a non-voting member for today's meeting.

DR. GLASSER: Steve Glasser, University of Alabama at Birmingham, radiologist.

DR. GIBOFSKY: Allan Gibofsky, Professor of medicine and public health, Weill Medical College of Cornell University, and attending rheumatologist at Hospital for Special Surgery in New York.

DR. CUSH: I am Jack Cush. I am the director of clinical rheumatology for Baylor Research Institute in Dallas.

DR. NEOGI: I am Tuhina Neogi, from Boston

University, a rheumatologist and epidemiologist.

DR. Good morning. My name is Sean Hennessy. I do pharmacoepidemiology research at the University of Pennsylvania.

DR. *OLSEN: I am Nancy Olsen and I am a professor of medicine and a rheumatologist at the University of Texas Southwestern Medical School in Dallas.

DR. O'NEIL: I am Kathleen O'NeiL and I am associate professor of pediatrics at the University of Oklahoma, in Oklahoma City, where I am a pediatric rheumatologist.

DR. VESELY: Nicole Vesely, designated federal official, Arthritis Advisory Committee.

DR. STINE: Robert Stine, professor of statistics at the University of Pennsylvania.

MS. ARONSON: Diane Aronson, consumer representative.

MS. LINDLEY: Suzanne Lindley, patient representative.

DR. CLEGG: Daniel Clegg, professor of medicine and rheumatology, University of Utah.

DR. HARRINGTON: Bob Harrington. I am a cardiologist at Duke University.

DR. FURBERG: Curt Furberg, Wake Forest University.

DR. GILBERT: Jane Gilbert, medical officer, FDA.

DR. SIEGEL: Jeff Siegel, clinical team leader, Division of Anesthesia, Analgesia and Rheumatology Products.

DR. RAPPAPORT: I am Bob Rappaport. I am the director of that division.

DR. ROSEBRAUGH: Curt Rosebraugh, Director, Office of Drug Evaluation II.

DR. O'NEIL: I will now turn the microphone over to Lt. Nicole Vesely who will speak to us.

Conflict of Interest Statement

DR. VESELY: Good morning. I would first like to remind everyone to, please, silence your cell phones if you have not already done so. I would also like to identify the FDA press contact, Miss Karen Riley. If you are here, present yourself. Please stand. Thank you.

The Food and Drug Administration is convening today's meeting of the Arthritis Advisory Committee under the authority of the Federal Advisory Committee Act of 1972. With the exception of the industry representative, all members and temporary voting members of the committee are special government employees or regular federal employees from other agencies and are subject to federal conflict of

interest laws and regulations.

The following information on the status of this committee's compliance with federal ethics and conflict of interest laws covered by, but not limited to, those found at 18 U.S.C. Section 208 and Section 712 of the Federal Food, Drug and Cosmetic Act are being provided to participants in today's meeting and to the public.

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

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

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

Today's agenda involves new drug application 21856, Uloric (febuxostat), sponsored by Takeda Pharmaceuticals North America, Inc., formerly TAP Pharmaceuticals, a wholly-owned subsidiary of Takeda Pharmaceutical Company, Ltd. through a licensing agreement with Teijin Pharma, Ltd., a part of Teijin, Ltd., a member of the Teijin Group, for the proposed treatment of hyperuricemia in patients with gout. This is a particular matters meeting during which specific matters related to Takeda Pharmaceuticals' Uloric (febuxostat) will be discussed.

Based on the agenda for today's meeting and all financial interests reported by the committee members and temporary voting members, no conflict of interest waivers have been issued in connection with this meeting.

With respect to FDA's invited industry representative, we would like to disclose that Dr. Mark

Fletcher is participating in this meeting as a non-voting industry representative, acting on behalf of regulated industry. His role at this meeting is to represent industry in general and not any particular company. Dr. Fletcher is employed by Pfizer, Inc.

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

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

DR. O'NEIL: Thank you. We will now have opening remarks from Dr. Jeffrey Siegel, of the Division of Anesthesia, Analgesia and Rheumatology Products at CDER, FDA. Dr. Siegel?

Opening Remarks

DR. SIEGEL: Thank you. Good morning and welcome to this meeting of the Arthritis Advisory Committee. We appreciate the committee's willingness to take your time to

participate in the process and to offer the agency your advice.

We are here today to discuss a new drug application for febuxostat for treatment of hyperuricemia in patients with gout. The particular issue that the agency is asking you to focus on has to do with the cardiovascular safety of febuxostat in the treatment of hyperuricemia. For that reason, we have supplemented the usual participants on the Arthritis Advisory Committee with members from the Drug Safety Advisory Committee and from the Cardiorenal Advisory Committee, and we appreciate your willingness to participate.

To address the issue that FDA would like the panel to focus on, we are going to begin the day with two talks. The first will be by Dr. Jack Cush. He will be discussing gout, the impact of gout on patients and current treatments for this disorder. The second presentation will be by Dr. Milton Packer and he will be addressing the issue of addressing safety issues when there are small numbers of events. As this is the case with the application today, this will be helpful for the panel to orient you towards the challenges that are faced in this situation and towards some of the approaches that can be taken.

Following these talks, you will hear presentations on the safety and efficacy of febuxostat by Takeda Pharmaceuticals. Then you will hear the agency presentation by Dr. Jane Gilbert. Her presentation will be focusing on safety, and in particular cardiovascular safety, because the agency does not have major differences with Takeda with respect to the efficacy of febuxostat.

Following the agency presentation, we will be asking you to address a series of questions regarding the safety and efficacy of febuxostat, the risk/benefit relationship, and ask you your judgement of whether febuxostat should be approved based on the currently available information.

Again, we would like to thank you for your participation. We look forward to hearing your discussion and your assessment about these important issues. Thank you.

DR. O'NEIL: Thank you, Dr. Siegel. Next we will hear from Dr. Jack Cush, from Baylor University Medical Center, who will review gout, the disease for which febuxostat is seeking approval.

Clinical Overview of Gout

DR. CUSH: Good morning, everyone. My charge this

morning is just to provide a launch point for this discussion today and to talk about some of the clinical features of gout.

[Slide]

This is an aged disorder, first described long ago by Hippocrates, and called the Aking of diseases and the disease of kings.@ It is a disorder of uric metabolism wherein the resultant hyperuricemia leads to deposition of monosodium urate crystals in joints and soft tissues and may, thereby, produce a myriad of clinical manifestations.

[Slide]

It is a prevalent disorder. In the U.S. and the U.K. the prevalence approaches about 1.0 percent or 1.5 percent in general. In some populations it is as high as 7 percent. In a Dutch study it was 3 percent. In general men are much more commonly affected than women, primarily because women are affected in the postmenopausal years, whereas men can be affected throughout their life span.

The National Arthritis Data Work Group had estimates recently that 3.1 million Americans have self-reported gout in the last year, and this may be as high as 6 million ever having gout in the United States. This is likely to be an overestimate because it is self-reported but

many feel that this is a reasonably accurate number in many ways. Earlier numbers were much lower, certainly lower than 2 million.

In >81 37 million lost work days in the United States due to this condition alone, and the cost to society is estimated by *Kim to be over $27 million annually in the United States. Again, this is also probably an underestimate.

[Slide]

So, again, its impact on society, on the workforce, on quality of life is significant. The peak onset of this disease is probably between ages 40 and 50, although as I mentioned, it can occur life-long, especially in men. In women it tends to be postmenopausal in its onset, primarily because estrogen is a uricosuric compound and may protect women during those years. It is estimated that only a minority of women have an onset prior to menopause.

The prevalence of the disease is certainly influenced by a number of factors, as I mentioned hormonal, but including geographic, racial genetic, dietary and even other conditions. Certainly males are more affected than females. There are certain populations where this disorder

is incredibly prevalent, as shown here, including Filipinos and Taiwanese males. In certain medical populations such as those on cyclosporine and those having renal transplants the incidence is markedly higher. Even patients with hypertension have an increased risk of developing gout. There is also an observed seasonal variance to the disease that some people talk about and, I must say, is probably not as important as we are led to believe by those papers.

[Slide]

What is interesting about gout is that its incidence in the population seems to be increasing. These are three different reports comparing two different time periods, a first era and a second era. You can see the NHIS study had an incidence of 5/1,000, almost doubling some 25 years later. In Rochester, Minnesota 45/100,000, almost 20 years later going to 62/100,000. Wallace, in the United States, also shows almost a doubling in just a 10-year time period.

[Slide]

Why is this so? There are probably a number of different factors. These would include general longevity of our population, increased prevalence of a number of different disorders that probably contribute to the onset of

gout, including hypertension, alcohol, obesity, the metabolic syndrome and increased number of patients with organ transplantation and cyclosporine use, and increased survival of patients with coronary artery disease and heart failure. Again, in this mix are the obesity and dietary trends as important considerations.

[Slide]

We do know that the prevalence of the disease does go up by age, and certainly in men and really in women after the postmenopausal years.

[Slide]

There are many known precipitants in association to gout. Those of us who manage this condition are well aware that these episodes in these patients frequently have histories of hypertension, obesity, diabetes, renal insufficiency, heart failure with or without diuretic use, alcohol consumption and alcohol abuse, lead exposure in a minority, and a family history in a significant number of people, and more recently association with sweetened soft drinks.

There are number of precipitants that may give rise to an acute gouty attack. This would include the use of alcohol; hospitalization; major surgery; certain drugs,

especially diuretics and cyclosporine and a number of others; total parenteral nutrition. Then, there are associations of gout in patients who have other arthritis*es, such as patients who have septic arthritis and/or reactive arthritis, and lupus which sort of may come down with gout on top of their already preexisting inflammatory articular conditions.

[Slide]

We do know from a number of studies, and Campion is an often quoted source showing that the number of gout attacks per year per 1,000 individuals goes up based on the serum uric acid levels. When you are in the normal uricemic range the number is quite low. However, with rises in uric acid we note a significant increase in the number, as high as 5 percent, in those who are over 9 mg/dL. The 5-year cumulative incidence of this may be as high as 22 percent.

Another study in a recent publication by Ken Saag and Choi shows also that, again, in a normal uricemic range the incidence of gout is relatively low. This is about 400 uM/L. This is around 7 mg/dL. But after this level the rate goes significantly up in the general population.

[Slide]

So, I mention again the number of clinical

associations between hyperuricemia and gout with some well-known and common disorders, and the net result of purine metabolism is uric acid. This can be contributed to significantly by the patient being an under-excreter, which is probably the larger group here. In patients with an impairment of renal handling of uric acid it leads to more hyperuricemia. Then there is a minority of individuals who are genetically determined over-producers of gout, the net result being hyperuricemia which can have a number of different manifestations, either asymptomatic hyperuricemia, gout with or without tophi manifestations, so gout and tophaceous gout; nephrolithiasis and other renal manifestations of gout. Then there is concern that there may be an association with a number of medical conditions including the metabolic syndrome, cardiovascular events, hypertension.

Is there an association between gout and metabolic syndrome? It is not surprising that the frequency of both gout and metabolic syndrome have both dramatically increased in the last few decades. This is probably being fueled by the increased obesity epidemic. We do know that patients who are obese have a significant increase of gout. As we will hear later on, gout imparts risk of cardiovascular

disease in men. However, whether there is a true association here as is suggested by this slide--the brown bars are patients without gout, and these are population percentages of metabolic syndrome. This is going up with age, but it seems to be more sizeable and maybe more premature in patients with gout.

However, other key elements of the metabolic syndrome are not really associated with gout. That would be resistance in dyslipidemia. So, whether or not hyperuricemia is an independent risk factor for cardiovascular disease, hypertension and renal disease has often been debated with growing amount of evidence in animal literature, which is very convincing, and in human experience, which is not as convincing, that there may be an association here but, again, I think the current state of affairs is that this is mainly a circumstantial relationship rather than a causal relationship. But, again, hopefully we will hear about that later on today.

[Slide]

A number of reports from Choy and colleagues in British Columbia have looked at the influences of gout in society. Many of the data are drawn from a survey of almost 50,000 males and looking at 700-plus new cases of gout.

Here, looking at the issue of diet and gout, and looking specifically at those on a high purine diet and those on a low purine diet, what we saw was that meat intake increased the relative risk of gout, as did seafood.

However, dairy products actually seemed to decrease the rate of gout. Not associated were purine vegetables and protein intake. So, again, the sort of low purine diet didn't seem to have much of an influence. And, there were some instances where, again, seafood and meats and beer and certain forms of alcohol could increase the risk of getting gout. But, again, our belief that diet drives this is probably not that well founded. Certainly, diet is one of the factors that influences the clinical expression of disease.

[Slide]

Choy has written a number of different articles that I think have been very helpful. We do know, based on some of his work and work of others, that alcohol is shown to have a dose-related increase in the risk of gout. In the Women's Health Study, and other sources have also shown this of course, obesity is associated with an increased relative risk of gout. More recently, the consumption of sweetened or fructose-rich soft drinks, especially in those taking

more than two servings a day, had an increased relative risk of developing gout of 1.85.

Again, there are many contributors to this disease. The net result is the precipitation of uric acid crystals in tissues, you know, above a certain level, certainly 6.8 or higher but varying according to the age of the individual and sex of the individual as to at what level the uric acid will be solubilized and may precipitate in tissue.

[Slide]

When it does so, this leads to an aggressive inflammatory response. Uric crystals are very, very proinflammatory and will incite the production of a number of proinflammatory cytokines. Again, I think the pathogenesis of gout has been expanded upon a lot in recent years. This is a very simplistic view where, because of these cells being taken up by macrophages and leukocytes, and what-not, there is an aggressive amount of chemotactic factors, enzyme production, proinflammatory cytokines, oxygen radicals, etc., which is why the significant inflammatory response that typifies gout is seen.

[Slide]

Again, these patients will tell you it is the

worst pain they have ever had and they can't put a sheet on their toe. It is very rare that they can walk on it. It is maybe more profound than other manifestations of inflammatory arthritis that we may see. Again, it may be because of multiple mediators being provoked by these uric acid crystals.

[Slide]

Gout has many different phases. We know that a significant number of individuals will have hyperuricemia but will never have clinical manifestations of gout, and at a certain point we need to treat hyperuricemia because it may predispose to gout nephrolithiasis and other complications that may ensue. But not all patients who have hyperuricemia need to be treated and, actually, that is one of the common clinical mistakes that are out there.

When it does become a problem we see acute manifestations which are in the form of podagra where the first big toe is involved. Gout in its acute form is an intermittent and recurrent disorder. It preferentially affects the MTP and lower extremities and then, with repetitive attacks, will ascend to involve the ankle, the knee, etc. Hence, it is unusual for gout to have its manifestations in the upper extremities at the outset.

When the disease quiets down we go down to an intercritical or interval period where these patients may have elevated uric acid levels and will continue to have tissue deposition but may not have any clinically manifested disease going on.

In its worst form, when the total body urate will become significant such that there is damage that occurs, we call this tophaceous gout that can lead to chronic inflammatory arthropathy, along with other complications as well. There are renal manifestations of hyperuricemia, mainly in the form of nephrolithiasis, but also gouty nephropathy and uric acid nephropathy which can sometimes occur.

[Slide]

So, at the onset, during the acute phase, there is a severe onset, often at night, of a warm, red, painful, swollen joint. Patients tell you they can't put a sheet on it. They can't walk. It is a disabling form of arthritis and often leads patients to consult a clinician right away.

Podagra is the primary manifestation. It is estimated that up to 90 percent of patients will develop podagra. Again, that is inflammation and swelling of the

big toe. Other joints can be involved, again, preferentially in the lower extremity, other joints, the MTP, the mid foot, the tarsus, the ankle and the knee. It usually is a monarthritis and less commonly oligoarthritis with a few joints being affected at the outset.

As I mentioned earlier, with repeated attacks the arthropathy will ascend and may become polyarticular late in the disorder. A polyarticular presentation is usually only seen at the outset in either the elderly or women usually who have renal insufficiency, are on diuretics, or those who have myeloproliferative disorders, or those on cyclosporine.

Because of the inflammatory nature of this disease and the enzymes, and the cytokines produced, it is not uncommon to see patients presenting with fever and impressive leukocytosis, high white count, high sed. rate, high C-reactive protein levels. Hence, it is not surprising that gout could easily be confused with septic arthritis. Again, untreated attacks last up to 14 days and it is estimated that those who have an episode of acute gout have almost an 80 percent likelihood of having a repeat attack within the next two years.

[Slide]

Nicholas Bellamy did an interesting study which I

don't know that we could do these days, but he studied 11 volunteers who had acute podagra and followed them for 7 days to look at the natural history of the disease. Two patients withdrew on day 4. Because of severe pain they couldn't tough it out. The 9 that remained showed some improvement, although it was gradual and over time. Either pain or swelling improved by day 5-7. However, only 3 noted resolution of their pain during the 7-day study.

The implications of this are important for those who want to get into acute gout and also for those who want to manage gout. We expect that without any therapy it would take maybe up to 7 days for patients to improve. We hope that our therapies will prevent this from happening or, if we are treating an acute attack, would lessen the duration of symptoms. Again, the symptoms here are significant, with extreme pain, swelling, redness, warmth, etc.

[Slide]

One of the common associations with this disorder when it is severe is the development of tophi, which are tissue depositions of uric acid. The incidence of tophi has decreased over the last several decades, presumably maybe because of better therapy and larger numbers of patients who are taking allopurinol. It is seen in up to half the

individuals, however, especially with long-standing disease. This is not something one would see early on in the disease process. Typical sites of tophi would be over the elbow, the olecranon, or in the bursae over the elbow, over the digits as shown here, and over the helix of the ear as shown on your right.

What is unique about these nodular swellings is that they may ultimately break through the skin and ulcerate, which distinguishes them in many ways from rheumatoid nodules which do not usually ulcerate, and this is not an uncommon manifestation of gouty tophi. It is, nonetheless, a palpable measure of total body urate load.

[Slide]

Clearly the most diagnostic test we have is the identification of monosodium uric acid crystals within the joint or from within a tophus. There are other forms of supportive evidence that can come from the lab. Again, we like to look at uric acid levels but, although you would think uric acid levels would be elevated in everyone with gout, it turns out that up to 50 percent of patients during an acute gouty attack will have normal uric acid levels, and the reasons for that are not entirely clear. Other manifestations in an acute attack are leukocytosis and other

acute phase reactants.

There are radiographic features that can support the diagnosis but, again, this would only be seen in patients with long-standing chronic disease. With soft tissue swellings and opacities that are seen on the x-rayB- this is shown from a paper by Dr. Schumacher, you can see these punched out nodule-looking lesions with an overhanging edge. This is sort of typical of gout. Patients, because of intermittent gouty inflammation, will have normal joint spaces and normal juxtaarticular bone mass which distinguishes them from rheumatoid arthritis. So, again, there is a typical appearance to the x-rays and that too can be useful in the diagnosis and management of these patients.

[Slide]

So, again, random hyperuricemia is not equal to gout so asymptomatic hyperuricemia does not always need to be treated. In the acute attack we may see normal uric acid levels in up to 50 percent of patients, and why that is, is not entirely clear. My belief is that actually inflammation promotes uric acid excretion, probably in many ways cytokine mediated. The diagnosis of this condition is a clinical one. It is certainly supported by the identification of uric acid crystals.

[Slide]

In 1977 American Rheumatoid Association criteria for diagnosis were complex, to say the least, and I think may be helpful in clinical trials but not practical in clinical practice. Nonetheless, the identification of uric acid crystals would be most important but then having any of the six following would be supportive. That would be one attack of acute arthritis; maximal inflammation within one day; erythema over a joint; podagra, history of podagra; unilateral tarsal involvement; tophus; hyperuricemia; asymmetric swelling on x-ray or exam; subcortical cyst without erosion; and then a joint fluid culture that is negative for infection.

[Slide]

On the other hand, in practice what clinicians often do is try to make a diagnosis that makes sense. So, a patient who presents with acute or recurrent inflammatory mono or oligoarthritis, especially if it is in the lower extremity and especially if it is repeated, would be presumed to have gout and that can be supported by, and confirmed by the presence of monosodium uric crystals, but when that is not present, to rely on other clinical features such as a history of prior intermittent similar attacks; the

evidence of hyperuricemia either during the acute or during the intercritical period; or x-ray evidence that would be compatible with a diagnosis of urate arthritis.

[Slide]

Here we see someone with a swollen foot with a big toe that is swollen. It is more of a big single one joint swollen, the MTP, and all this being due to deposition of those needle-like or uric acid crystals, in this case being phagocytized by a neutrophil.

[Slide]

The goals of that would be, obviously, to terminate the acute flare when it occurs, but then to protect against further flares and to reduce the consequences of crystal-induced inflammation and to identify patients in whom hyperuricemia needs to be treated to prevent such manifestations in the future.

[Slide]

I am giving you just two simple algorithms that could be used to treat gout. This would be to first treat acute gout. The first step would be to say that patients who have acute gout should receive nonsteroidals unless they are contraindicated, as shown here. However, if they were contraindicated for some reason, you would next want to use

a corticosteroid unless it was contraindicated. If you could use a corticosteroid, what form and how you use it would depend on the extent of involvement. If it was one joint you could use either PO or intra-articular steroids. If it was many joints you could use, again, oral or intramuscular or intra-articular steroids. If they are, however, contraindicated, then colchicine would be the drug of choice for the management of acute gout.

This is an algorithm which I think many support but, interestingly, when you survey most physicians they will tell you the drug of choice in the management of gout is colchicine, which is unfortunate because the goal of therapy with colchicine is to give enough colchicine to give you GI toxicity in the form of cramping and diarrhea and the patients are just miserable. If they were miserable from their joints, now they are miserable from their gut. I think we can do a little bit better with some of the other therapies that I have outlined here.

[Slide]

During the interval period or the intercritical period patients may or may not need to be treated, and that sort of depends on the number of acute gouty attacks and, again, it is up to the patient to decide whether or not he

needs to be on chronic therapy. So, patients who have two or less attacks per year may want to just be observed and to educate them on risk reduction, and just treat an acute attack if and when it occurs.

For patients who have more attacks or even a few attacks, you want to instruct them about stopping alcohol, diuretics and weight loss as those being potential factors which can modify their clinical course.

However, those who have a number of attacks may require some form of urate-lowering therapy. In that case the drug of choice would be allopurinol unless, of course, there were reasons that allopurinol could not be used. The choice would be uricosurics unless you couldn't use a uricosuric and uricosurics would be contraindicated in a situation of nephrolithiasis, with tophi present, with extreme elevations of uric acid, with evidence of renal insufficiency or patients who are over-excreters of uric acid.

Again, allopurinol in many cases is the drug of choice because it can be done in either situation, either with any of these or without these. Mainly what needs to happen here, you need to dose adjust allopurinol patients who have renal insufficiency.

[Slide]

So, again, there are some factors about gout that we need to recognize. It is treatable and preventable. It is a disorder that is largely managed and diagnosed and treated by a wide variety of clinicians out there, mainly in emergency rooms and primary care offices. Many of these clinicians will make this diagnosis without evidence of uric acid crystals because they won't or can't do arthrocentesis. There are very few who are in fact managed by rheumatologists. In most rheumatology practices we tend to see the most severe patients and we are not seeing those patients with acute or intermittent gout.

There are studies that show that rheumatology referral and management leads to more accurate diagnoses, shorter symptom durations, less hospitalizations and lower overall hospital courses. A recent study by Krishnan and colleagues showed that more than two-thirds of gout patients are managed by primary care doctors, with only 1.3 percent being managed by rheumatologists.

There are a number of physicians who, unfortunately, do equate hyperuricemia with gout and will reflexively treat all hyperuricemia, and there are many who believe that colchicine is the drug of choice and, maybe

more importantly, there is still a significant minority of individuals who are using IV colchicine, which is a very dangerous drug and should be warned against. Again, there is a significant amount of inappropriate management and I think most of this rests with a lack of education about better ways of treating the disease.

[Slide]

I did a survey, and I am all for doing surveys of my colleagues. This is a survey of almost 500 rheumatologists, done earlier this year. When I asked them what are the major milestones in the treatment of gout, the number one major milestone they rated as being very important was the introduction of allopurinol, and 70-plus percent said that allopurinol was the most important thing. As you can imagine, crystal identification was also quite high; colchicine use, etc.

When I asked them what was most disappointing about the care of patients with gout they said the lack of new drug development in 70 percent of cases. Patient non-compliance, over 50 percent of responders. Treatment by non-rheumatologists, almost 40 percent. Then, the management or treatment of tophaceous gout they viewed to be problematic. At least a third of the responders thought

this was so.

So, I will end here and we will move on to I guess Dr. Packer. Thank you.

DR. O'NEIL: Thank you, Dr. Cush. Next we will hear from Dr. Milton Packer, from the University of Texas Southwestern Medical School, in Dallas, and he will review issues in the interpretation of safety data in clinical trials. Dr. Packer?

Interpretation of Safety Data in Clinical Trials

DR. PACKER: Thank you very much, Dr. O'Neil. I would like to congratulate the wisdom of FDA for inviting their two speakers from the same medical school, in the same city. Both speakers, of course, took their training in New York City.

DR. CUSH: And we both like Mexican food.

DR. PACKER: One of us does.

[Slide]

The topic that I am going to cover today has nothing to do with arthritis and nothing to do with gout but has a great deal to do with the generic issue, and that is our ability to reach conclusions about safety when most clinical drug development is focused on efficacy. This is an enormous generic problem. It is a problem that many of

us have been worried about for quite some time.

I head the Department of Clinical Sciences at UT Southwestern and one of our major research focuses over the past several years has been to develop a partnership between clinical investigators and statisticians to try to tackle what is an enormously vexing problem. Part of what I am going to show you today is a progress report, an interim progress report of some of our progress in this area. What I am going to present really represents a joint effort between myself and Dr. David DeMets from the University of Wisconsin, who is head of statistics there. We have both worked together for the past 15 to 20 years to try to reach a handle on this dilemma.

[Slide]

I think it would be fair to say that most of you realize that the clinical trials that you see submitted by sponsors for approval for a specific indication are largely focused on efficacy and when they design a statistical plan and they identify primary endpoints and secondary endpoints they specify a limited number of analyses. They specify an analytical approach to evaluate endpoints which are prespecified at the start of the study. There is almost always adequate statistical power to find an effect if there

is one. And, there is rigorous control of type I and type II error.

All of these are highly desirable characteristics in trying to reach conclusions as to whether there is an effect or there is not an effect. Almost all of these enormously beneficial characteristics are absent in the evaluation of safety. There is generally an extremely large number of analyses if you look at the safety listings in clinical trials, 200, 500 potential adverse events. You can group them together in a variety of different ways. There is generally inadequate statistical power to find an effect. There is inadequate power to rule out an effect and, therefore, there is no meaningful control of type I or type II error in the evaluation of safety.

[Slide]

That means that we have to interpret what we see in the evaluation of safety very, very carefully. We actually think we know what a p value means when type I and type II error are rigorously controlled. But it is very difficult to interpret p values regardless of whether they are greater than 0.05 or less than 0.05 when type I and type II are not controlled. So, p values for safety are very hard to interpret.

[Slide]

So, I want to give you my top four list of things to worry about when analyzing the incidence of adverse events. Let me emphasize that this is not a David Letterman type of list. If I felt that we had the time I could exceed ten but I am only going to give you four.

[Slide]

First of all, there are hundreds of adverse events and, therefore, statistically we have an issue that we are performing hundreds of comparisons. Theoretically we could be calculating hundreds of p values. I would strongly recommend you not do that, but it is certainly possible to do that.

[Slide]

A typical large-scale clinical trial may describe as many as 500 individual terms describing adverse events. If, in fact, a p value were calculated for each pairwise comparison, then one would expect, by chance alone, about 25 events to have a p value less than 0.05 and 5 events to have a p value less than 0.01, assuming that there were sufficient statistical power for each of those events which, of course, there is not.

[Slide]

Second, remember how adverse events are reported in a clinical trial. They are generally spontaneous reports and almost always are non adjudicated. There is no systematic way that each individual investigator uses the same judgment to report an adverse event.

[Slide]

These are done at the discretion of the investigator. They are translated into standardized terms using a variety of dictionaries. These dictionaries, frankly speaking, are awful. They are arbitrary. The individuals who are responsible for the dictionary have an impossible job. They keep updating it. I don't know how they think they can make a judgment as to whether their efforts are improved or not. The real difficulty is that the uncertainty increases when the event is in a field remote from an investigator's focus. So, I am a cardiologist. You would not want me to make a judgment about the reporting of arthritis. Therefore, it is equally hard I think for rheumatologists who are principal investigators to have a uniform definition of what is a cardiovascular event, myocardial infarction or unstable angina.

[Slide]

Some have said that I know how to fix this problem. What we are going to do is take the events, certain events that we care about and we are going to adjudicate them. Now, ideally, one has to understand that the adjudication process does not, in fact, result in the truth. An adjudication process results in a reduction in the degree of imprecision. Therefore, one applies uniform criteria. Depending on the criteria one applies, one could, in fact, get closer or further from the truth and it is particularly worrisome when the adjudication is done post hoc.

[Slide]

Let me just emphasize that the post hoc adjudications in general have very high quality control. They are done blinded. But the rules guiding post hoc adjudication are inevitably influenced by the knowledge that somebody is worried about something. And, because you know someone is worried about something you can create definitions for events that set a high bar or a low bar and they can magnify or dilute any treatment effect.

The other thing to remember is that adjudications apply not to every single adverse event. It is applied only to adverse events of interest. For example, if you are

interested in acute myocardial infarction you send to the adjudication committee only, for example, the serious adverse cardiovascular events. Well, it could be that there are acute myocardial infarctions coded in the database as something other than cardiovascular. So, that results in some uncertainty about this particular process.

[Slide]

Third, the analyses that are done in large part also depend on the grouping of events and those groupings can be subject to bias. Let me give you an example. The reason, by the way, is that the groupings are frequently created by individuals who have already looked at the data and can subconsciously decide which events they feel should be grouped and which events shouldn't be grouped.

[Slide]

Let me give you an example. Just suppose you had a concern about cardiovascular adverse events, thrombotic cardiovascular events, and you were looking through a list of adverse events and you saw this imbalance on myocardial infarction. For the next couple of examples let me just assume that the ratio here is 1:1 randomization.

Here, you know, gee, this doesn't look good. It seems worrisome, 20 events versus 5 events. That doesn't

seem very good at all. Well, that seems a little bit clearer than the next example.

[Slide]

That is, just suppose you had some imbalances. You didn't reach a p value, assuming that you wanted to. So, now you start grouping myocardial infarction and stroke because you think they may be biologically related.

[Slide]

Well, you could even get more creative if you wanted to and you could notice that, well, maybe you are worried about thrombotic events and you actually find that the risk of myocardial infarction and the risk of stroke on active therapy is very similar to what is seen on placebo. But you still want to worry and so what you do is you go to other cardiovascular events, for example unstable angina and transient ischemic attack which are not myocardial infarction and not stroke but are clinically and biologically related to these, and you see there is an imbalance there. So, depending on how you group these you could say there is a problem or there isn't a problem. If all of these groupings are done after someone looks at these tables, they are subject to considerable error.

[Slide]

So, it is best to develop a uniform definition of a group before classifying events because, if the process of developing a group definition is started after a concern has been raised, those creating the definition have frequently already looked at the data and know subconsciously what kind of definition is needed to capture the events of interest.

[Slide]

But the most important concern that exists in the interpretation of safety data is that in general the trials that we see result in the reporting of small numbers of events that result in extremely imprecise estimates.

[Slide]

Let me give you an example. In a clinical trial you are looking at major adverse cardiovascular events. You have 25 events on placebo, 26 events on drug. Almost every clinical investigator would look at this and say, well, you know, there still could be a problem. Twenty-five and 26 doesn't rule out a significant problem. You can see the confidence interval is here. The upper limit goes up to 1.79. The absence of an observed difference does not rule out the existence of a true difference.

I think most people realize that if you compare the frequency of adverse events in two groups and there

isn't a significant difference, it doesn't mean there isn't a difference. But what is more interesting is that seeing a difference doesn't mean there is a difference. This is what becomes really perplexing.

[Slide]

Here is an example of a drug where the risk ratio is 2.5. Confidence intervals don't embrace 1.0. There is a nominally significant p value. You would think there is a difference. But the problem is that these estimates are terribly imprecise.

[Slide]

And you might say, well, what is wrong with imprecise estimates?

[Slide]

Well, imprecise estimates are fine if the intent is to withhold judgment until more precise data are collected to make the estimate more precise. But imprecise estimates are highly problematic if the intent is to stop and reach a conclusion.

When calculated in the conventional manner the 95 percent confidence intervals and the associated p value of an estimate primarily have meaning in the context of a completed experiment.

[Slide]

But the adverse event data generated in a typical trial is not the result of a completed experiment. Viewed from the amount of data needed for a precise estimate, the adverse event data in a single study represents a snapshot in an ongoing experiment to characterize the safety of a drug.

So, essentially although you can characterize the efficacy of a drug in a single trial or two trials or three trials, you are essentially characterizing the safety of a drug in every single study which is done with the drug, and the assessment of safety is done at the end of this accumulated experiment. Therefore, performing an analysis of adverse events data in a single study is akin to performing an interim analysis of primary endpoint data in an ongoing clinical trial.

[Slide]

Now, in the cardiovascular field we have a large number of big studies, outcome studies, event driven, and we do interim analyses in these studies. Many of you may know that what we do is look at the information time on the X axis and we look at the treatment difference, usually displayed as a Z score, on the Y axis.

[Slide]

We plot what the results are. For example, at the beginning we have no information and we have no differences at the very beginning of the trial, and during the course of the trial we get some information. We see some difference; more information, whatever difference we see. And, we try to interpret these differences in light of the amount of information that we need to make a judgment.

[Slide]

Now, many of you may think that reaching an answer in a clinical trial is a linear process, that if you plot information time and the Z score it will be a straight line.

[Slide]

Let me tell you it is never a straight line. In real life the ups and downs of the differences over time can be quite striking. At any given point in time in a clinical trial you can have results that are significantly different from the final result both in a favorable and unfavorable direction. Usually a lot of this fluctuation in differences occurs when the information is very small. So, when the information is small these differences really do not reflect the precision of the estimate and the apparent truth that is seen at the end of the study.

[Slide]

So, what some statisticians have done is to set up certain rules, certain boundaries up front in order to guide the interpretation of differences in observed frequency of events in safety data. This is just one such rule. This is a rule by Len and DeMets. The way this rule works, very, very simply, is to say that you can conclude that an effect is seen at the end of the trial at a p value of 0.05. But when you are earlier and earlier in the trial you need much larger differences in order to feel confident that you are minimizing your type I error and keeping it at the 5 percent level.

[Slide]

So, for example, here you would conclude there is a treatment effect because the p value is less than 0.05. But here this same difference with an observed p value of 0.05 would not result in the conclusion that there is a difference. The difference would have to be much more striking and cross the boundary in order for you to declare that the difference was real.

[Slide]

Depending on how much error you are willing to accept, these boundaries can look very, very different from

each other.

[Slide]

So, reaching conclusions from data derived in an underpowered trail, and this is generally what we see when we are dealing with safety data, raises the same concerns as reaching conclusions based on an underpowered interim analysis in a definitive, adequately powered trial. Basically, when you look at safety and safety data are sparse it is like looking at the very early part of a large-scale clinical trial.

[Slide]

Let me give you some entertaining examples of how wrong we have been in the area of cardiovascular disease. I am going to pick heart failure because it is my primary field of interest, and I am only going to give you a few examples because the number of examples is too numerous to mention.

Here is one early example. This is a trial called the IMPRESS study. The comparators here are not important. But the sponsor did a Phase 2 trial looking at the effect of their drug, omapatrilat, versus a conventional ACE inhibitor and they found, wow, a 47 percent reduction in the risk of this very important endpoint. By the way, this is

what we care about in heart failure, death or hospitalization for heart failure. Only 39 events but a 47 percent reduction. The p value is 0.53. We were really excited.

[Slide]

Boy, it is time to go and do a big trial and replicate this, and they did and there was nothing. Risk ratio of 6 percent reduction in risk, not statistically significant.

[Slide]

Here is a trial of a drug called amlodipine, also a cardiovascular drug. This is looking at all-cause mortality. All-cause mortality, I mean how more important or objective can be than death? Here are the results of the initial trial, phase 1. This is in patients with non-ischemic cardiomyopathy, and 74 deaths on placebo, 45 deaths on amlodipine, 45 percent reduction in the risk of death; p value of 0.001.

[Slide]

You would think this is real until we tried to replicate it in a trial which was four times larger and showed no significant benefit.

[Slide]

In case you are getting depressed, this is probably the most striking example we have. I am going to show you two trials. One is vesnarinone, a drug for heart failure, versus placebo. This is in the top panel. The second is losartan versus captopril. Again, the specifics of these drugs are not important.

I am going to show you the results of an initial study with both drugs. You can see here is vesnarinone versus placebo and this study had 450 patients in it, and a 33 to 13 split on the endpoint of all-cause mortality; 62 percent reduction in the risk of death; p value, 0.002.

Here is losartan versus captopril. About the same number events. About the same treatment difference, 46 percent reduction in risk; p value of 0.039.

[Slide]

You would think these differences are real until larger scale studies were done. What is scary is the top panel. In the top panel you have the effect of a small pilot study with a 62 percent reduction in risk, statistically significant at 0.002. I know many of you would say at least you could conclude there is no harm. But that wouldn't be a valid conclusion.

When this trial was replicated in the same

population with the same drug and the same dose there was a 22 percent increase in the risk of death which was statistically significant in the opposite direction. The same kind of pattern was seen in the losartan versus captopril trial.

[Slide]

We can show you many, many examples, results of trials of magnesium in myocardial infarction. Look at the number of events here. We are not talking about small numbers of events both in a trial called LIMIT-2 and in a meta-analysis significant benefit and reduction in risk of death with magnesium, intravenous magnesium in myocardial infarction.

[Slide]

And then a large-scale definitive trial which, in fact, went in the opposite direction. So, we don't know what we know.

[Slide]

It can go in terms of safety as well. Here are the effect of a drug called metoprolol in heart failure. Beta blockers in heart failure are established treatments for this disease. The results of a trial called RESOLVD there was a three-fold increase in risk with metoprolol

versus placebo on heart failure hospitalizations, statistically significant.

[Slide]

And, when the definitive trial was done, of course, the risk was significantly lower in the metoprolol group compared with the placebo group.

[Slide]

So, one has to realize that to achieve statistical significance in an underpowered analysis the effect size must be extreme and the estimate must be precise.

[Slide]

This is the best way to look at it. In order for a small number of events to reach statistical significance, because the number of events is small the confidence intervals have to be very, very large. And, in order for this confidence interval to not overlap 1.0, this point estimate has to be shifted all the way to the right, which means the only way you are going to get statistical significance with a small number of events is when the calculated treatment effect is extreme.

[Slide]

Let me tell you the more extreme the effects, the more imprecise the estimates; the less likely the result

will be reproduced in definitive trials.

[Slide]

I am going to try to life the depression by, in the last few minutes, trying to address the question of what to do.

[Slide]

The most important first step is to develop an approach to analyzing data in trials with small numbers of events which accurately reflect the true imprecision of the treatment effect estimate and its statistical significance.

[Slide]

I am not going to go into this. This is actually a pretty interesting topic. What we essentially try to do, and I am just going to race through this for the sake of time, is we change the way the confidence intervals are calculated because, in fact, when the number of events is small it is not appropriate to use nominal 95 percent confidence intervals and one should, in fact, adjust the confidence intervals because of the fact that it is an interim analysis in an ongoing study.

We have developed the concept of using boundary Z scores and boundary-adjusted confidence intervals but all this does is make you feel more insecure. All it does is

tells you that what you think you know you don't know. So, why don't we try to come up with something that actually could resemble a solution? One, some people would say believe in observed differences that are biologically plausible.

[Slide]

God, I hope we don't do this because you have to be wary of differences that are deemed real based on biological plausibility because physicians can always be relied upon to propose a biological mechanism to explain the validity of an unexpected and potentially preposterous finding that happens to have an interesting p value. I don't think anyone in this room would challenge this.

There is another way that one can approach this, and that is to look for patterns. One pattern that many have relied on is the concept of a class effect and, therefore, one interprets imbalances and the frequency of adverse events with a new drug by looking for the presence or absence of similar imbalances in other members of the same drug class.

[Slide]

For example, suppose I began with experience in a drug and I had a very imprecise estimate. This is what I

observed. There was a nominal increase in risk but I really don't know. I am not very, very confident about it.

[Slide]

I look at other members of the class and I notice that in other members of the class there is no risk. So, what I essentially do is adjust these data so that now the point estimate is shifted towards neutrality. But, unfortunately, the confidence intervals are still quite wide.

[Slide]

It is more interesting if you actually know that other members of the class have had a problem. When other members of the class have a problem, then effectively you are taking the information from the other class members and narrowing the confidence interval even though the number of events you have with that particular drug are very, very small. This is something that we all do. In fact, the FDA does this routinely in the absence of data. If you have an ACE inhibitor it is presumed to produce angioedema even though angioedema has never been reported, and seems reasonable. So, you don't even have to have an estimate here in order to use class as a way of informing the decision-making process.

[Slide]

The problem here is that we don't know how to define a class. Do we define a class based on chemical structure, physiologic action, mechanism of action? We don't know how to do this and this makes life interesting.

[Slide]

You can look for confirmatory evidence in other studies with a drug. Obviously, one needs to avoid being selective. You can actually do a cumulative meta-analysis. The problem is the amount of data you need to make these estimates reliable is very large.

[Slide]

Of course, everyone has the perfect solution. Why don't we ask every sponsor to carry out a definitive trial with the adverse event as the primary endpoint, powered to detect meaningful treatment differences? This really is the perfect solution but it is unbelievably expensive. Most of the trials will show that there isn't anything to worry about because most of these will represent false differences. It does significantly delay the entry of drugs into the market.

[Slide]

But none of those should really matter because

sponsors get excited about looking at encouraging trends for important endpoints and, in fact, if sponsors are excited about looking at encouraging trends maybe we should encourage them to look at discouraging trends and provide some real issues. A definitive trial can fix a lot of the problems that we have already talked about.

[Slide]

But here is the problem. Just suppose you saw an imbalance of events; just suppose you were actually worried that a drug increased cardiovascular risk and you wanted to do a definitive trial to show that the drug increased cardiovascular risk. So, you go to a patient. You give him an informed consent. You say we are worried that this drug could kill you but we actually don't know. We would like to find out. Can you please sign here? It would be difficult.

[Slide]

Lastly, I just want to briefly say that there are many who would say this isn't a symmetrical issue, that we need to be conservative when it comes to safety and that the evaluation of safety and efficacy really reflect two different standards. Well, realize that we are strict in reaching conclusions about efficacy because saying that there is a benefit when there is none means millions will be

treated unnecessarily and be subject to side effects and costs.

Now, some may advocate being less strict in reaching conclusions about safety, but saying that there is an adverse effect when there is none means millions will be deprived of an effective treatment. I actually think of these as being symmetrical.

[Slide]

So, in conclusion, the findings of controlled clinical trials are most easily interpreted when they represent the principal efficacy endpoint of a study. Safety data is subject to many interpretations, difficulties, including ascertainment biases, inflated false-positive rates due to multiplicity of comparisons and the imprecision of estimates inherent in analysis of small numbers.

FDA, industry and academia remain in a quandary as to how to respond in a responsible fashion to observed differences in reported frequency of adverse events when, in almost all cases, the frequency of such events is small.

Questions from the Committee to Speakers

DR. O'NEIL: Thank you for that uplifting review. Next, the panel is asked to explore questions with Dr. Cush

and with Dr. Packer. Dr. Cush?

DR. CUSH: I would like to ask Dr. Packer a question. I like the analogy or the information time axis and how you showed that each of these is sort of an interim analysis until we get more data. But then you showed examples where we increased the size of our populations, as little as a four-fold increase or, actually I think one was a two-fold increase and with as much as a 30-fold increase in longer studies we got the answer, maybe a more appropriate answer that was different from the initial analysis or interim analysis. Is there a formula that one can use, you know, that would actually get us to that right answer, and should we look at all clinical trials as a cumulative process to get to that point?

DR. PACKER: Well, philosophically, of course, the collection of human knowledge never ends. You don't actually stop at any point in time and say that you know something, or you shouldn't. Therefore, there is no point in time when you actually have 100 percent information time. Even when drug development ends there is still information, useful information that accrues.

But the one thing to keep in perspective is not the knowledge of whether you are at 100 or whether you are

at zero because knowing that you have no information is easy, and you should never be at 100 percent information because you should never stop asking. But knowing whether you are closer to zero or closer to 100 is really important. Sometimes someone asks me the question, well, you had a trial that had 50 events and you got a certain answer. You now have a trial that has 500 events and it has a different answer. How do you know which one is right?

Those are not data with equal weights. The precision of the estimate with 500 events is very narrow. The precision of the estimate with 50 events is very, very wide. So, remember that the search for truth in the human experience is not for accuracy; it is for precision. So, we are far more confident about the estimate from a 500 event trial than a 50 event trial and we imagine that it gets us closer to the truth but the process never stops.

DR. O'NEIL: Dr. Gibofsky?

DR. GIBOFSKY: I want to compliment Dr. Packer on an excellent presentation but I still have one fundamental problem, and that is I am not quite sure how to deal with the notion that the report of an event in a trial means that the active agent caused the event. Risk does not equal cause. So, how do we deal with the notion that one can see

events reported which may be serendipitous or in no way the result of the drug itself?

DR. PACKER: One thing I did not use during my whole presentation was the word Acause.@ Let me just emphasize that I think it is impossible, if the only thing I am interested in if I am looking at events in a trial, is a comparison of the frequency of an event on active therapy compared to the control.

I really think one has to be very, very cautious in putting any weight on whether the investigator thought that the event was related or not. Of course, there is no way an investigator can make that determination in a clinical trial. So, although physicians really think they know whether something caused something in clinical practice, the number of spontaneous events in the absence of therapy is very, very meaningful, just as you said.

So, I think what you are stuck with is comparing the frequency. If you don't have a comparator group it is very, very hard to make a judgment as to whether any individual event has occurred with greater or lesser frequency than expected in that population.

DR. O'NEIL: Dr. Furberg?

DR. FURBERG: Well, Milton, you spent your time

describing and illustrating the problems we are facing and I think they are overwhelming. You didn't spend much time on solutions, other than saying that we need more data and that is clearly what we do. But then the issue is what are our options? Do we request that information prior to approval? Or, would it be okay to do it post approval? And, what type of study do we set up? Could you comment on that?

DR. PACKER: Oh, I was actually asked by FDA to describe the problem rather than the solutions. I was very grateful for that because if I had to focus on the solutions it would have been a very brief presentation.

I think what we are left with is two driving principles. The first is pattern recognition which is, unfortunately, a very subjective process. I think we have to engage in pattern recognition but, frankly speaking, it is replete with error and I have no problem in engaging in pattern recognition as long as we admit it is an error-filled process.

The second is everything has to be risk/benefit. So, you have to think how important is this drug to what we need to do for patients compared to the level of uncertainty we have about its safety. But a very important thing that I think you just emphasized is that the information gathering

process absolutely should never stop at the point of approval because, you know, no matter how precise your estimates are, you need to get more and more information over time.

So, I think part of the problem is that we actually think of pre approval and post approval as being binary. There are lots of drugs where the information about safety comes out after approval even though the trial started pre approval. So, I don't have an answer. But I think it is the combination of pattern recognition in the context of risk/benefit.

DR. FURBERG: I agree with that. I think that is critical. Also, you said to consider what the available options are. I mean, if we have no treatment options we may lower the bar for approval but I am not sure it is part of the regulatory process for us to lower it under those conditions.

But there is one other thing that has troubled me, Milton, and that is the time to an adverse event. There is no guarantee that problems start on day one. There are some adverse effects that clearly are initiated very, very early on and we see them very soon, and then others will take some time.

DR. PACKER: Yes.

DR. FURBERG: And the case we have today, for example, is where the trial goes on for six months. Well, that tells me what happens over six months but not a word about what is happening down the road. And, if we have therapies that are intended for chronic, long-term use I don't think six months is adequate. We need to follow people much, much longer, either pre approval with longer studies or post approval where we get the information about the long-term adverse effects.

DR. PACKER: Yes. Curt, let me underscore that. First of all, I wanted to just make the point that I didn't want to deal with anything related to today's application in my presentation.

DR. FURBERG: That is fine, sure.

DR. PACKER: But the best way to get safety data is to maximize the information for every single patient in a trial. Getting patients into trials is a very challenging, very expensive proposition. So, if you can take a patient that is supposed to be in a trial for one month and you can make that six months, that patient now provides more safety information which is obtained far more efficiently than if you had to recruit another patient. Then it is the point

that you made, which is some safety issues are entirely time dependent. You would never see them in a trial of a month or two months, or whatever.

So, it is both. The safety issues sometimes are time dependent, cardiovascular issues being a classic example. But the truth is that it is actually better for the sponsor to take patients and make each patient more efficient in terms of the information that is derived from each individual patient by following them longer.

DR. FURBERG: Thank you.

DR. O'NEIL: Dr. Harrington?

DR. HARRINGTON: Milton, I want to also congratulate you on a very nice overview of these complicated issues. I want to follow-up on a couple of Curt's points.

You made the comment that if you have an adverse event it would be difficult to go and approach a patient and say, hey, we see this adverse event; would you be willing to enroll in a clinical trial? But turn that around and say that you have an adverse event which seems to occur with an increased frequency in a certain group of patients who might take a therapy. Is it not appropriate to do clinical trials where you enrich the population, so to speak?

DR. PACKER: Oh, absolutely. One of the other issues which I thought Curt was going to mention but you are mentioning instead is the fact that it is very hard to get safety data when you put patients who can't experience an adverse event into a trial, or are unlikely to experience an adverse event into a trial.

So, the maximization of safety dataB-let me take what I said and expand it--would be to put patients with all sorts of risks, including high risk patients who would normally get the drug in real life, and follow them for a longer period of time so you are maximizing the information in each patient and in each trial. And, that would greatly enhance the event rate and, therefore, make the estimates more precise.

But what I thought you were going to ask was just suppose you didn't do that and you have what you have, and you have an imbalance, and you are worried about it, and you really feel strongly it ought to be pursued, how do you go about doing that in a trial? In other words, if you know you are stuck with going forward, what do you do? And, one step which I think you are suggesting is do another efficacy trial but in a high risk population. Is that about right?

DR. HARRINGTON: That is certainly one approach.

But the follow-up to that would also be, as Curt said, you did not offer solutions and I can certainly appreciate why you didn't, but your views on non-randomized safety collection. You avoided that topic and I understand because it is much easier to make comparisons in randomized data sets, but what is your view in general in terms of how observational, non-randomized data contributes to a safety base?

DR. PACKER: Oh, I am not a skeptic when it comes to observational studies and here is the reason why. Everyone or a lot of people, and certainly it is taught frequently in school and this is really an important point, create this hierarchy of evidence, with randomized trials up here and meta-analyses of randomized trials, by the way, even higher, and observational studies here and descriptive studies, you know, case reports, here.

This hierarchy is entirely appropriate as long as each one is carried out to an optimal degree. Here is the thing I worry about. You have a clinical trial which has no precision. It essentially generates early in the trial random numbers. And, you have an observational study and there are observational methodologies which are fantastic. They are really good. You have to identify the confounders.

You have to correct for them. It is much better if you are following a cohort over time. The nice thing about these observational studies is that they actually correct the deficiency that we had before because they have a large number of events. So, now you have clinical trials with a small number of events but they are randomized. Observational studies have a very large number of events but they are not randomized.

Which one is more important? Thankfully, we don't have to choose. And, what I am hopeful about is that this hierarchy between randomized trials and observational studies disappears because they both bring different things to the table.

DR. O'NEIL: Very quickly, Dr. Glasser, you had another question?

DR. GLASSER: Well, we have skirted around a lot of this issue I am going to bring up, but I guess I will finish since I guess time is short now, with this.

I was involved in two large meta-analyses of the FDA database on the safety of placebo-controlled trials in angina and hypertension. In so doing, we had to adjudicate adverse events from case report forms. Some, I might mention, case report forms from studies I was involved with

making notations on the adverse events. And, it was very clear, and you mentioned this, that when we set up efficacy trials without outcomes we are very precise in the definitions of the outcomes that occur. We are not so precise in terms of the adverse events.

So, at least one of the partial solutions perhaps is to be more precise in identifying the adverse events so at least there is some consistency.

DR. PACKER: In order to make life not totally miserable for the investigators, of course, you might do that for certain types of adverse events. But, Steve, the point you have raised here is terribly important. Everyone puts the concept of adjudication at some high level. You know, if you have lousy data it is very hard to shine it up with adjudication. It is very, very hard.

You know, everyone here has been part of an adjudication process in cardiovascular disease and sometimes you have nothing valuable to adjudicate. So, you are putting your blessing on rather imprecise information. You want to call the investigator on the phone or send them an email saying what really happened here? Good luck!

DR. O'NEIL: We are running a little bit late but I still want to get Dr. Stine's comment because his area of

expertise is pertinent to this.

DR. STINE: I just wanted to mention that I think oftentimes when we talk about data in one context we forget about methodologies that have been developed for data in other contexts that are very, very similar. The whole field of pattern recognition, machine learning, has, indeed, made huge strides in the identification and recognition of statistically significant patterns and very, very difficult dimensional data problems which are very similar to the safety issues that we are speaking about here today.

Even stepping away from the machine learning community, you can also look at methods related to the false discovery rate, and such, that are designed for multiple comparison sorts of settings. They are not quite the same as the method of DeMets that you are talking about here, but they are designed to control the opportunities for false positives, putting the weight of evidence on belief in the null hypothesis and only rejecting that when you have a preponderance of evidence.

There are ways to do that when you have many, many thousands of drugs, or whatever, that you are looking at. All one needs to do is look at the literature in genetics where it is common to have a sample of 50 observations and

have 10,000 measurements per subject.

So, you can do that there. You can do that here. I don't see any limitation on taking those same ideas that have been developed in that context and bringing them into the safety area. You are looking at many, many millions of combinations of factors and those methods carry over here. You are not going to have a hell of a lot of power because you don't have a hell of a lot of data.

I don't share Dr. Packer's comment on clinical trials and randomization. I am quite a fan of randomization. I don't like meta-analysis very much but I do believe in the quality of information provided by randomized studies, and I don't think that they produce random numbers early on in the trial either. If I sit here and toss a coin 20 times in a row and I get heads every time are you going to be willing to bet me at 50/50 I won't get heads the next time? I don't think so.

And, I think we have to step back from these kinds of comparisons. A p value is a p value, is a p value and if you have done it right it has a uniform distribution on 0/1. I don't care how many observations you have, if it has been correctly calculated it has that meaning. That is the whole point of how it is calculated.

Now, the other kinds of issues here I think go back more to the science of what is an adverse experience; how is it classified; do we have complete data; are there missing cases. I think the science in my sense, speaking as a statistician, trumps the statistical issues in this context and are going to always supersede concern over is this p value less than 0.05 or not. I think we have to understand the science and make sure we are measuring the right kinds of information over the right time horizons in the appropriate populations.

I have it written down. I am writing a book and so this example where a p value goes from statistically negative to statistically positive, I have to find out why because it is my belief that it is not the statistics that are goofed up; it is the study change between one phase and the other. That is my suspicion as to what actually happened, and I have to go back and find out. It was a great classroom example.

DR. O'NEIL: Thank you. We will now take a short, 10-minute break. Panel members, please remember that there should be no discussion of the PMA during the break among yourselves or with members of the audience. We will resume promptly at 10:10.

[Brief recess]

DR. O'NEIL: Now we will hear presentations from the sponsor, Takeda Pharmaceuticals North America, Inc. The first speaker this morning will be Dr. Nancy Joseph-Ridge who will present an introduction and overview. Dr. Joseph-Ridge?

Sponsor Presentation

Introduction

DR. JOSEPH-RIDGE: Thank you. Good morning.

[Slide]

Panel members of the Arthritis Advisory Committee, members of the FDA, ladies and gentlemen, my name is Nancy Joseph-Ridge. I am currently the president of Takeda North America, research and development. I am a rheumatologist by training. I have been involved with the febuxostat program previously as a medical director and also as area head in the therapeutic area.

I have been involved with the design of the original studies, including the original Phase 3 studies. I work closely with clinical investigators and key consultants.

[Slide]

Today I will provide a brief overview. This will

be followed by medical need. I will follow this with efficacy and safety but Dr. Becker will be doing the medical need presentation. Dr. William White will present cardiovascular safety and I will return to present on risk/benefit.

[Slide]

Febuxostat was evolved to address a growing population. We see approximately three to five million individuals in the United States with increase in incidence and prevalence. There is a need for a more effective urate-lowering agent compared to current therapies, and there have been no new gout therapies approved in over 40 years.

[Slide]

Our current proposed indication is for the treatment of hyperuricemia in patients with gout at a dose of 40 mg or 80 mg given once a day. We are recommending 80 mg for those patients with higher serum urate and those patients with tophi.

[Slide]

Our initial febuxostat NDA was filed in December of 2004. In there we had one Phase 2 and two Phase 3 studies of over 1,900 patients. The doses were febuxostat 80 mg and 120 mg. At the time we submitted, we received a

response from the FDA requesting additional information. At the time we were requested to further examine the safety profile of febuxostat. It was noted that there was a small number of cardiovascular events and an apparent imbalance.

[Slide]

We then submitted an amendment to the NDA with an independent evaluation of all potential cardiovascular events in the Phase 2/3 studies. We then also committed to conducting a Phase 4 clinical study outcome. FDA then responded to this submission requesting additional information. We were asked to more clearly characterize the potential cardiovascular risk of the 80 mg dose and an additional safety and efficacy of the lower dose of 40 mg.

[Slide]

We conducted an additional Phase 3 study. This was of over 2,000 patients. This was larger than both the prior Phase 3 studies combined. We prospectively designed to evaluate all cardiovascular events and enroll subjects with renal impairment. We had a data monitoring committee that evaluated safety throughout the trial, and we had a cardiovascular endpoints committee that adjudicated all events in a blinded manner.

The doses studied at that time were febuxostat 40

mg and febuxostat 80 mg. The 120 mg dose was not included in the study since it was evaluated in a prior Phase 3 study and had a similar safety profile as the 80 mg dose. We will evaluate the 120 mg dose at a later time, after examining the medical need and the best utilization for this dose.

[Slide]

Today we will discuss if there is a medical need for a new treatment of hyperuricemia in patients with gout. We will demonstrate that febuxostat did not show an increased risk of CV events relative to allopurinol; that our clinical program is reflective of the gout population; there was no plausible biological mechanism for these cardiovascular events; and that a new large Phase 3 study did not substantiate the previously observed apparent cardiovascular imbalance. We will demonstrate that the benefits of febuxostat outweigh the risks and support approval for the proposed indication.

[Slide]

I would like to ask Dr. Michael Becker to present.

Gout: Disease Burden and Unmet Patient Needs

DR. BECKER: Thank you, Dr. Joseph-Ridge.

[Slide]

Members of the advisory committee, FDA personnel

and ladies and gentlemen, as Dr. Cush has so nicely explained this morning, gout is a common, acutely disabling and often chronically destructive disease that is increasing in incidence and prevalence both in the United States and worldwide.

Estimates of the prevalence in the United States have focused on approximately four to five million individuals and the incidence appears to be rising, especially in the population of individuals over 60 years of age.

Gout symptoms result from their response to the body to urate crystal deposits in tissues. These crystals arise from body fluids that are saturated for urate. Hyperuricemia, defined as a serum urate level in excess of 6.8 mg/dL, is a mirror of the saturation of extracellular fluids and the invariable risk factor for the development of gout.

[Slide]

Shown on the previous slide was a polarized microscopic examination of joint fluid from a patient with acute gout. This is shown as well in higher power in this slide, in which the monosodium urate crystal has been ingested by a neutrophilic leukocyte.

As Dr. Cush mentioned, crystals can interact with resident cells in the joint lining, and the result of this interaction is the generation of proinflammatory cytokines and chemokines which call forth the type of cells seen in this slide and, upon ingestion of crystalline material, result in further inflammatory events--

[Slide]

-Bwhich we see clinically in the symptoms of patients with acute pain and disability of acute gouty arthritis and the classical clinical signs of warmth, swelling, tenderness and acute disability.

[Slide]

Patients who develop gouty symptoms arise among a much larger population of individuals with asymptomatic hyperuricemia. Asymptomatic hyperuricemia exists in approximately 15 to 20 million people in the American population. Hyperuricemia is a necessary but not sufficient state for the development of gout, and it may persist for days or for a lifetime. Best estimates are that approximately 25 to 30 percent of individuals over a lifetime ever develop the symptoms of gout, which is a disease state as opposed to hyperuricemia, which I consider a biochemical curiosity.

The onset of gouty arthritis is the most common means that gout declares itself. Again, as described by Dr. Cush, approximately 80 percent of patients affected by a single attack of acute gouty arthritis will have a second attack within two years. A few patients, perhaps 25 to 30 percent of the gouty population of four to five million, will have infrequent attacks that justifiably can be maintained by treatment directed to the acute attack rather than chronic urate lowering. The majority of patients, however, are likely to go on with an accelerated rate of acute attacks of gouty arthritis or the occurrence at intervals of urinary tract stones of uric acid or even calcium oxalate and, over a period of time amounting to perhaps 10 years or more, these individuals could develop the type of chronic arthropathy and tophaceous gout that will last a lifetime.

[Slide]

The development of chronic tophaceous gout is a result of urate crystal aggregates that can become expansive, compressive and destructive. The ones shown here are in synovium but can expand locally into cartilage and bone as well--

[Slide]

B-giving rise to the kinds of findings shown in this slide in the hand and the elbow of a woman whom I attended for a number of years who was unfortunate enough to have a misdiagnosis of rheumatoid arthritis for a substantial period of time and then, by the time that the correct diagnosis of gout was made, had significant renal insufficiency that, in fact, precluded getting full doses of urate-lowering medications that were then, and are currently now available.

[Slide]

In this slide the hands of a man whom I also have attended for the past dozen years is shown. This unfortunate gentleman presented with recurrent urinary tract stones preceding his gouty arthritis. When he was initiated with treatment on allopurinol to lower his serum urate levels he developed a skin rash which recurred during desensitization and, because he is a uric acid over-producer and has had many, many stones, has been unable to be treated with currently available medications, resulting in a great deal of disability. You see the presence of tophi in multiple places and his quality of life has suffered.

[Slide]

Another point made earlier was the significant

association of hyperuricemia and gout with important comorbidities, taking the form of impaired renal function and even the complete metabolic syndrome and cardiovascular diseases which include the thrombotic complications of myocardial infarction, stroke and peripheral artery disease, heat failure and hypertension.

[Slide]

Current urate-lowering management of gout is aimed at the achievement and maintenance of serum urate in the sub-saturating range, usually less than 6 mg/dL. The purpose of this goal is to reduce the body urate pool, dissolve crystals, prevent and reverse gout symptoms and the progression to disability and impaired quality of life.

[Slide]

At least the dissolution of crystals and the reversal of symptoms and progression can be achieved as suggested by this composite slide. On the left-hand side is the reduction in an acute manifestation of gout, that is, acute flares during the second and third years of treatment, as shown on the Y axis, as a function of the average serum urate concentration during therapy on the X axis. At serum urate concentrations below 6 mg/dL there is a very low rate of recurrence of flares in comparison to individuals who do

not achieve such control of serum urate.

In the right-hand panel is a manifestation of chronic gout, that is, tophus formation is shown here with regard to reduction of tophi, the Y axis showing the serum urate concentration, the X axis the velocity of tophus reduction.

What can be seen here is that the lower the serum urate level, the greater the velocity of tophus reduction, once again with almost an asymptote at the level of 6 mg/dL, or an inflection point at 6 mg/dL. And, 6 mg/dL or less as a goal is really an empiric observation or suggestion, but operationally appears to be very useful in the management of patients with gout on a chronic basis.

[Slide]

One problem in the management of gout with urate-lowering agents is what is called treatment initiation flares. There is an increase in the risk for gout flares that occurs early in urate-lowering management with any agent that has been used to date. These treatment initiated flares have a significant impact on patient adherence to therapy and the satisfaction of patients and, of course, their caregivers as well.

The mechanism is speculative but it is suggested

that preexisting crystals that have been inactive, perhaps because of their surface coating with non-proinflammatory immunoglobulins or lipoproteins, are then activated and the surface coating changed to immunoglobulins which become proinflammatory, or else the crystals actually migrate from small aggregates setting forth a proinflammatory response.

[Slide]

One positive note, however, is evidence that co-administration of prophylactic therapy during acute urate-lowering with allopurinol can successfully reduce significantly the number of treatment initiated attacks of gout.

In this slide we see the results of a randomized-controlled trial of 43 patients who were all started on 300 mg of allopurinol a day for their gout. Half the patients were randomized to receive colchicine at 0.6 mg twice daily and the other half received placebo as co-administered therapy. Over the period of the first 6 months, divided into 0-3 months and 3-6 months, there is a significant reduction in the number of flares experienced in the individuals who received colchicine. This parallels clinical experience of most of the practitioners in the field of rheumatology with whom I have spoken.

[Slide]

The backbone of urate-lowering therapy for gout for the past four decades in the United States and in most other countries in the world is xanthine oxidase inhibitor allopurinol. Allopurinol reduces urate formation by inhibiting the last two steps in the formation of urate acid. Allopurinol is approved at doses from 100 to 800 mg per day, but 95 percent or more of allopurinol dosing in the United States is at 300 mg per day or less. In addition, less than 50 percent of gout patients reach goal serum urate levels of less than 6 mg/dL at 300 mg per day.

Because of concerns that allopurinol and oxypurinol metabolites will accumulate and play a part in some of the toxic reactions to allopurinol, dose reduction in treatment with allopurinol is recommended both in the package insert and is widely known in algorithms provided to primary care physicians and rheumatologists. This dose reduction has a side effect however in that it lowers the likelihood that the patients will be able to reach a serum urate lowering towards the goal that we have outlined.

Finally, there is minimal randomized-controlled trial evidence for safety and efficacy of allopurinol at doses greater than 300 mg per day.

[Slide]

Intolerance to allopurinol occurs in up to 20 percent of patients at currently used doses. Most allopurinol intolerance is mild, reversible and easily managed, but rarely, likely less than 1/1,000 patients, allopurinol hypersensitivity syndrome or severe cutaneous reactions such as Stevens-Johnson syndrome or toxic epidermal necrolysis occurs that can be life-threatening or fatal. Such a patient is shown in this slide.

[Slide]

Among the characteristics that I think a new urate-lowering therapy should have are that, first and foremost, it should be safe and have clinical efficacy in all patients with gout, whether the patients have normal or reduced renal function; whether the patients are over-producers of uric acid or under-excreters; whether the patients have tophi or don't have tophi. It would be very useful to have agents that do not require dose reduction in patients with mild to moderate or even more extreme renal functional impairment which is very common in the gout population. Finally, it would be very helpful to maintain the convenience and increase the compliance through use of once a day dosing.

[Slide]

So, in conclusion, gout is an increasingly common, often progressive and disabling disease. Although available urate-lowering therapies benefit many patients with gout, there is a documented need for safe new urate-lowering options to prevent unnecessary acute disability and long-term disease progression in the broader range of current gout patients. Thank you. Dr. Joseph-Ridge?

Febuxostat Development Program

(Efficacy and Safety)

DR. JOSEPH-RIDGE: Thank you, Dr. Becker.

[Slide]

Regarding efficacy, I will review the pharmacokinetics and pharmacodynamics, provide an overview of the clinical program with clinical results in regards to serum urate, flares and tophi, provide results from our long-term extension studies and summarize the efficacy conclusions.

[Slide]

This is simply to demonstrate the difference in chemical structure between allopurinol and febuxostat. Allopurinol and its metabolite oxypurinol have a similar structure to purine hypoxanthine and, therefore, they are

referred to as purine analogs. This is in contrast to febuxostat which is a non-purine like structure.

We have noted that febuxostat inhibits xanthine oxidase both in its reduced and oxidized forms. This may account for the increased potency we see relative to allopurinol.

[Slide]

Febuxostat has a good pharmacokinetic profile with extensive absorption and dose proportional increase, and has a half-life that allows for once a day dosing without accumulation. The clearance is mainly by hepatic metabolism and renal elimination is minimal. Clinical studies have shown that febuxostat is unlikely to be involved in drug-to-drug interactions with those drugs commonly used in patients with gout.

[Slide]

Febuxostat is very effective in lowering serum urate. The pharmacodynamic effect is rapid and usually reaches steady state within one week. We saw that there were no clinically relevant differences in the pharmacodynamics with regards to food, age or gender, mild or moderate hepatic or renal impairment.

[Slide]

The primary objective of our clinical program was a reduction of serum urate to a level of less than 6 mg/dL. You have heard from Dr. Becker's presentation about the importance of that target. New treatment guidelines have established this level as a target level. The goal of treatment of the clinical manifestations of gout is correction of the underlying hyperuricemia. In maintenance of serum urate a level of less than 6 mg/dL is associated with reduction of tophi and gout flare.

[Slide]

Our clinical program studied doses ranging from 10 mg to 300 mg daily. We have 25 Phase 1 studies, six Phase 2/3 studies with doses of 40, 80, 120 and 240 mg. The 240 mg dose was a safety dose that was requested to be provided in our clinical study. That was two times the previous highest dose studied of 120 mg.

[Slide]

The studies that we will discuss today include the Phase 2 dose-ranging study; the Phase 3 FACT and APEX. Those were the prior Phase 3 studies. And, CONFIRMS which is a new trial. The long-term extension study for Phase 2 was FOCUS and for Phase 3 EXCEL. Those subjects who were in either FACT or APEX were allowed to enroll into EXCEL.

[Slide]

Let me spend a little bit of time on the study design for the Phase 2 since it is similar to that of the Phase 3 studies. These are parallel, double-blind, randomized-controlled trials. We have a 14-day washout period for those subjects who were on prior urate-lowering therapy. We provided prophylaxis for those treatment initiated flares. In this study it was colchicine 0.6 mg twice a day. The treatment arms in this Phase 2 dose-ranging studies were placebo, febuxostat 40 mg, 80 mg and 120 mg. The study duration was 28 days.

[Slide]

The results from this Phase 2 are as follows: We are evaluating the serum urate less than 6 at the final visit and the percent of subjects who achieved that level. The colors are indicating the different dose groups. Those colors will be used throughout the presentation. For the febuxostat 40 mg we have a slightly yellow color. Febuxostat 80 mg will be in orange, and 120 mg will only be presented in the dose-ranging study but that has the darker color.

What we see is that for the Phase 2 study we had a dose-dependent response with 40 mg, 80 mg and 120 mg,

showing a response of 56 percent, 76 percent and 94 percent of subjects achieving that target level, all of which were statistically superior to placebo.

[Slide]

Our Phase 3 studies: The first of the two original Phase 3 studies is APEX. The study design is similar, as I stated, to the Phase 2. The treatment groups were placebo, febuxostat 80 mg, 120 mg, 240 mg and allopurinol, either 100 mg or 300 mg.

Please note that the febuxostat dose of 240 mg and the placebo were randomized with half the number of subjects than the other treatment arms. Subjects with serum creatinine greater than 1.5 or less than 2.0 were randomized to the lower dose of allopurinol as is recommended per the label. We provided prophylaxis for an 8-week period of time for this Phase 3 trial with either nonsteroidals at a low dose or colchicine. The study duration was 28 weeks.

[Slide]

The second of the two original Phase 3 studies in is the FACT trial. In this study the treatment groups were febuxostat 80 mg, 120 mg and allopurinol 300 mg. Prophylaxis was also given for an 8-week period in the original Phase 3 trials. Subjects in this trial had to have

a serum creatinine of greater than, or less than, or equal to 1.5. This study had a duration of 52 weeks.

[Slide]

In the third Phase 3 study, which is the new study, the treatment groups were febuxostat 40 mg, the lower dose, febuxostat 80 mg and allopurinol either 200 mg or 300 mg. Based on the previous results from the APEX study and using new renal guidelines, subjects with creatinine clearance of 30 cc to 59 cc were randomized to allopurinol 200 mg. Prophylaxis was also different in this trial. Prophylaxis was given for a 6-month period of time. This was based on the literature that Dr. Becker reviewed from Borstad that was available after the original Phase 3 trials that showed that we could prevent treatment initiated flares for a longer period of time by extending prophylaxis. The study was a 6-month study.

[Slide]

For all of the Phase 3 studies the enrollment criteria were similar. The subjects had to meet the ARA criteria for gout. They had to meet the serum urate level of greater than or equal to 6 mg/dL at baseline. The exclusion criteria were that they not have secondary hyperuricemia and they were to be in stable medical

condition.

[Slide]

Our population is primarily male, Caucasian, age of approximately 50 years, a BMI of approximately 30 kilograms/meter squared. The presence of tophi we noted in approximately 20-25 percent of our subjects. Mean years of gout was approximately 11 years, and mean baseline serum urate level was approximately 20 mg/dL.

[Slide]

With regards to medical history, we see that the renal function with mild to moderate renal impairment was approximately 50-60 percent of the population; hypertension in about 50 percent; hyperlipidemia in about 30-40 percent; and diabetes, atherosclerotic disease, use of low-dose aspirin in about 10-15 percent.

[Slide]

Our primary endpoint for these studies for the proportion of subjects who achieved serum urate less than 6 for the FACT and APEX were at the last 3 visits. For the CONFIRMS trial it was at the final visit. This was changed because the last 3 visits were stable and we noted those in the prior Phase 3 studies.

The secondary endpoints for FACT and APEX were the

serum urate less than 6 at the final visit; proportion of subjects requiring treatment for gout flares; percent reduction in primary tophus size for the CONFIRMS trial the proportion of subjects with renal impairment who achieved the target level of serum urate of less than 6 mg/dL.

[Slide]

The results for the Phase 3 are as follows using the serum urate of less than 6 at final visit. Let me take a moment to orient you to this graph. It is given as percent of subjects who achieved that level for each of the trials that we conducted. There you have CONFIRMS, the APEX, the FACT trial and we have also included the dose-ranging study to compare the 40 mg dose.

You will note that I will only discuss the 80 mg dose in the APEX and FACT trials. The other dose and dose-responses could be found in the briefing document. Here and from the other presentations, green will be denoted for allopurinol.

The results of these trials are as follows: For the CONFIRM studies 40 mg was not inferior to allopurinol. It had a similar response as the allopurinol dose group. The 80 mg had a statistically significant response of 67 percent compared to 40 mg which had a response of 45 percent

and allopurinol that had a response of 42 percent. When we look at the APEX and FACT trial we see that both studies showed that 80 mg had a statistically superior response compared to allopurinol, with approximately 30 percent more additional patients having their response. In the dose-ranging study we see the 40 mg having a 56 percent response and the 80 mg having a consistent response as seen in the Phase 3 trials.

So, overall we see 40 mg having a similar response as allopurinol and 80 mg being statistically significant in achieving this target level compared to 40 mg of febuxostat and to the allopurinol dose group.

[Slide]

An important subgroup in our gout patients are those with renal impairment. So, we addressed that target with those patients who had mild or moderate renal impairment and we saw for the CONFIRMS trial that there was a statistically significant response with 40 mg compared to the allopurinol group and 80 mg was superior, again, to both 40 mg and the allopurinol group.

When we reviewed the APEX and FACT trials, the prior Phase 3 studies, using the same creatinine clearance calculation of greater than 30 cc to 89 cc per minute, we

see again that 80 mg maintains a statistical significance over allopurinol.

[Slide]

An important secondary endpoint was the evaluation of achieving this target in those patients who had a high baseline serum urate, which we defined as a serum urate level of greater than 10 mg/dL. In these subjects, across all 3 studies, we see that 80 mg had a statistically significant response to both 40 mg and the allopurinol dose groups with, again, approximately 30 percent more subjects having a response.

[Slide]

When we look at those patients who had tophi at baseline we see the 80 mg response being statistically superior to both 40 mg and allopurinol.

[Slide]

Now let's turn our attention to gout flares. You have heard from Dr. Becker regarding treatment initiated gout flares. These flares occur as a part of initiating treatment with a urate-lowering therapy. What we noted in our studies is that these flares occurred in all treatment groups. There were more flares with the higher doses of febuxostat. However, over time these flares decreased.

When we looked at the APEX and FACT trials when prophylaxis was given over a short period of time there was an increase in gout flares after the prophylaxis was withdrawn. This is in contrast to when we were giving prophylaxis for the entire 6 months in the CONFIRMS trial where we see a low level of flares.

[Slide]

This is depicted graphically. For the APEX trial you see that prophylaxis is given; the same for the FACT trial. Post prophylaxis, we see an increase in gout flares for both studies. However, you note that after time you start seeing a decrease and these are by dose groups where placebo is in grey and the different colors as I stated previously, with the darker, sort of diamond shape for 240 mg. Allopurinol is in green. When we look at the FACT trial we notice that close to one year we see that there is a difference that starts to almost emerge in a one-year trial. These are not statistically different between the treatment groups, but you do see a decrease in gout flare over time.

[Slide]

When we look at the percent of subjects who have gout flares over time for the CONFIRMS trial when

prophylaxis is given during the period of time, you only see a few subjects, a low percent of subjects, with treatment initiated flares and over time those rates of flares are low and maintained at that level.

[Slide]

When we turn to tophus, we looked at reduction of tophus. Twenty percent of all our subjects had tophi. We noted that reduction in tophus size was noted in all treatment groups, with 6 months having approximately 30-50 percent reduction in size but after one year we see approximately 50-80 percent reduction in tophus size.

[Slide]

One of the things that we wanted to evaluate was did achieving a serum urate level of less than 6 affect either gout flare or tophus size. This was a non-prespecified analysis, however we felt this was an important parameter to look at.

We grouped subjects by average post-baseline serum urate levels of either less than 6 or greater than 6 regardless of treatment. We then summarized the proportion of subjects with flares and the percent reduction in tophus size. What we noted was that there were fewer flares in those patients who achieved a serum urate of less than 6 by

weeks 20 to 24, and by weeks 48 to 52 in the FACT trial there actually was a statistically significant difference. We also noted that there was a larger reduction in tophus size with those subjects who achieved a serum urate of less than 6.

[Slide]

This is a graph of subjects with gout flares, percent of subjects with gout flares. In the light blue are those subjects who have greater serum urate, and this is regardless of treatment. The dark blue line are those subjects who achieved a serum urate of less than 6. What we note is that over time you start seeing a separation between the two groups where those who achieve a serum urate of less than 6, in the dark blue, have a much better response and reduction in gout flare.

[Slide]

Our long-term open-label studies, FOCUS, included approximately 116 subjects who enrolled from the Phase 2 dose-ranging study. That treatment duration was for approximately 5 years. For the Phase 3 studies, APEX and FACT, the enrollment to the EXCEL study, and we had approximately over 1,000 patients enroll, and treatment duration was about 3 years. The aim of the long-term study

was to maintain a serum urate of less than 6.

[Slide]

The study design for the Phase 2 study is as follows: All subjects were randomized initially to 80 mg of febuxostat. They were prophylaxed with colchicine for the first 4 weeks. During week 4 to week 28 they were allowed to titrate between 40 mg and 120 mg in order to maintain their serum urate level to a level of less than 6. After week 28 they were to remain on their dose until the end of the study.

[Slide]

For the EXCEL study, which is the Phase 3 long-term open-label study, we had two designs. One is the original protocol and then the protocol was amended. In the original protocol all subjects began on 80 mg, similar to the Phase 2 study. In this Phase 3 open-label extension study they were prophylaxed for a 2-month period of time, such as in the FACT and APEX trial. During the time from month 1 to month 2 subjects were allowed to switch to 120 mg to maintain their serum rate to a level of less than 6. After month 6 they were to remain on stable dose.

[Slide]

The study was amended to the following: We were

asked by the regulatory agency to add an allopurinol treatment arm. So, the amendment was as follows, subjects were then randomized post-amendment to receive either 80 mg, 120 mg or allopurinol upon entry into the open-label extension in a 2:2:1 manner.

During the first 2 months, as with the other study, they were given prophylaxis. During months 1 to month 6 subjects were allowed to switch between the febuxostat doses or switch treatment to allopurinol, or vice versa, to maintain their serum urate to a level of less than 6. After month 6 they were to remain and maintain stable treatment until the end of the protocol.

[Slide]

The results from the long-term extension studies are as follows: 80 percent maintain a serum urate of less than 6 while on febuxostat. The majority of those subjects were on the 80 mg dose. We saw that approximately 50 percent of subjects switched from allopurinol to febuxostat because they were not able to achieve and maintain a serum urate of less than 6. In contrast, only 5 percent switched from febuxostat to allopurinol.

Tophi resolved in approximately 50 percent of subjects after a 2-year period of treatment. With regards

to gout flare, what we noted when we looked at gout flare over time in the FOCUS, which was the Phase 2 study, and in the EXCEL study with all treatments because they were all titrated to maintain their serum urate of less than 6, we see a reduction in gout flare over time to almost no flares after a period of time. And, this takes about 2-3 years. You start seeing a total reduction and almost elimination of gout flare.

[Slide]

So, in conclusion, febuxostat 40 mg and 80 mg effectively lower and maintain serum urate level to less than 6 level; 80 mg was superior to both 40 mg and allopurinol, including those subjects with high baseline serum urate level and/or tophi. Both 40 mg and 80 mg were effective in subjects with renal impairment without dose adjustment. And, maintenance of serum urate of less than 6 demonstrated decreases in gout flare and tophi resolution.

[Slide]

The safety agendaB-I will review the exposure in the clinical trials; the discontinuations; adverse events and serious adverse events. The areas of interest include cardiovascular safety. Dr. William White will present on this area. I will present the renal, hepatic and