DR. WOFSY: David Wofsy.

I have a couple of questions about the radiographs. We've seen two different ways of looking at the data. So, in the Centocor presentation, we saw a flat line in the patients treated with infliximab showing no deterioration in the scores. In the FDA presentation, we saw that, by certain criteria, close to 50 percent of people had progression. I'm trying to determine whether these are inconsistent with one another or entirely consistent.

That is to say, one interpretation of a flat line, if you look at patients, is that there will be sort of a noise scattering around it, half of the people above the line and half of the people below the line. If you define the half of the people above the line as progressors, then you get 50 percent progression, but really what you're looking at is noise. So, that's the question. Is this noise?

DR. SIEGEL: Let me take that.

The data are consistent. I think we would all agree that that flat line that you looked at, which was a median, does stay close to 0. Depending on which analysis, it could be plus or minus a half here or there. The reason that you'll see, if you look at a 0 cut point, 50 versus 80 progressing is, in fact, because there is scatter around

the median.

That then raises the question, though, you raised. Is that noise or is it real?

B if you look at the least detectable difference, which is a very conservative way for defining progression -- when you say "conservative," you better qualify it, for defining progression. If people progressed by less than 8.6, they're not counted as progressors. So, there only 6 or 7 percent versus 31 progressed.

But it's not at all conservative for defining nonprogressors. The conservative way for defining nonprogressors would be to count as a nonprogressor only people who improved by 8.6 perhaps, because those are only ones you can be sure didn't have a small amount of progression. There, as you see, the two curves are close to 100 percent progressors. There are very few on either arm. A few I think on the infliximab arm. It doesn't amount to a lot.

So, if you look at an individual patient who progressed by less than 8.6, it could be real or it could be noise, but one would guess, on average, that if you look at mean rates and you look at the 0 point, those 30 versus 80 that had higher numbers probably represent some significant element of reality, that smaller amounts of progression that you couldn't be sure exist in an

individual patient are occurring in some subset of patients on both arms.

DR. SIMON: We're going to come back to this in our discussion in a little bit.

DR. SIEGEL: It does link closely to the issue of the use of the word "prevention." In a group median you don't see any change, but there's scatter around, some of which is statistical and some is real.

DR. ELASHOFF: I'd like clarification of the FDA analysis of the HAQ because the slide says nothing about looking at change, but presumably you must have looked at change since some of the scores were negative. That's the number one part.

The number two part is to clarify the difference between the way you dealt with missing data and the way the company dealt with missing data.

DR. MATTHEWS: Yes, it is a change from baseline to week 54. So, it's a landmark analysis. The way that we handled data, although there's always some difficulty with that, if a patient had a missing data point at week 54, we carried their last value at the last visit where there was a calculation because it occurred at various time points throughout the trial. We carried that value forward.

As you heard from the sponsor, they did their

analysis for the area under the curve a little bit differently. My understanding is that they assigned them a change of 0, and in order to bring the missing data points -- and they also didn't do a range. They brought people who had a worsening of their HAQ score up to 0 as well. They assigned it a point of 0. So, it's a little bit different. Well, actually very different.

DR. WHITE: I just want to go over the first question again. If we could go to your slide, Dr. Mills, the last one that you showed. I just want to make sure I understand the data. In this slide, radiographic progression is defined based upon change from a particular cutoff value. So, if you go to the 8.6 at the top of this graph, then given the smallest detectable difference at 8.6, if the patients have a change of 8.6, the smallest detectable difference, then that would define progression. So, that slide of the curve defines progression and means that's 30 versus 8 percent.

Now, if you go to the other side of the curve, that defines not progression, given the smallest detectable difference?

DR. MILLS: What you're identifying across is that this is simply a demonstration of multiple cutoff points from 8.6 to a negative 8.6. We're not changing anything as we go across. So, if you used a negative 8.6,

virtually all of the patients in both groups, the all infliximab patient group, as well as the methotrexate/placebo arm, are defined and stated to have evidence of radiographic progression.

DR. WHITE: Right. I'm just trying to make sure I understand this. So, given the smallest detectable difference of 8.6, then a reading of 0 might not be different from negative 8.6. That's the smallest detectable difference. So, you would have to be minus 8.6 or beyond to be not progressing.

DR. SIEGEL: Let me try to reword this again. All points on this curve and those that would extend out from it potentially can be looked at to divide progressors from nonprogressors. You can look at the first curve and say -- and I think this is what you were asking -- those that we are rather certain on an individual basis progressed is 30 percent versus 6 percent. If you look at the other end, those that we're rather certain on a individual basis did not progress, it's 0 percent versus -- I don't know -- 2 or 3 percent. But that leaves 70 percent on one arm and 90 percent on the other arm who changed, most of whom changed. A significant number may have stayed exactly the same, but many of them changed, but changed by less than 8.6.

My personal bias is that looking at 0, if you

really want to dichotomize the population, gives you the best guess at proportions who change. But the variability in the data on an individual basis makes it harder to make firmer statements than that.

I don't think it would be right to say, though, for example, when you say that defines the numbers who progress, that 6 percent defines the numbers who progress on infliximab, but defines the numbers who progressed enough that we're sure that it's real, but a lot progressed by 2, 3, 4, 5 -- not a lot, but several did -- where it could be statistical but it could be real as well.

DR. SIMON: One more.

DR. FIRESTEIN: Thanks. Gary Firestein.

I would appreciate it if you could clarify some of the issues on malignancy in the safety database. There are a couple of ways of looking at this, and one that we've heard about is how the number of malignancies that have occurred in the treated group compare with expected number of malignancies based on historical databases.

But what I'm not sure we have heard about is how the number of malignancies that have occurred in the treated groups compares with the placebo group, whether that is statistically different, or if you look at subgroups, specifically the higher dose groups, whether or not there is a trend or there is statistical difference

among the groups at this point.

DR. MATTHEWS: Well, I think at this point the numbers are just too small to make that comparison. They did occur in the higher dosing regimens which sort of makes you think perhaps that might be a potential risk factor. But again, it's just too small. The occurrences were just too rare in this database.

DR. ELASHOFF: If you use a statistical test where you dose response across the five groups, then you do have statistical significance. If you just compare the five groups not paying any attention to dose, then I think it's like .07, something like that. I actually ran those with 0, 0, 0, 2, and 3 as the numbers in the cells.

DR. FIRESTEIN: So, there is statistical significance if one analyzed it in that way. Okay, thank you.

DR. SIEGEL: I'm not exactly sure on this question, but the data were presented both at 30 week and at 54 week, and the five cases would be, I guess, at 54 week. It gets somewhat confounded by differential loss to follow-up, so that there are more patients in the placebo arm who were lost to follow-up.

DR. SIMON: Thank you.

I'd like to take the chairman's prerogative at this time and switch the agenda around slightly. We have

four questions to answer. We have patient open statements to take. I'd like to take a 15-minute break at this point, come back, do the open forum, and then answer questions, and then go to lunch. After lunch, we have at 2:30 an open session for discussion with the FDA about x-ray outcomes. So, at this time we're going to take a 15-minute break and reconvene here at 11 o'clock.

(Recess.)

DR. SIMON: So, just to review with everybody again, what we're going to do now is the open patient forum. Then we are going to attempt to discuss the questions that are provided by the people on the FDA, and then we are going to do that hopefully leaving time for an adequate lunch and then for us to all reconvene for the afternoon session.

So, without further ado, I would like to recognize the open public hearing and to ask Mary Armitage to approach the microphone please for her 5-minute presentation.

MS. ARMITAGE: Good morning, everybody. My name is Mary Armitage and I live in a town called Richfield in the southwest corner of Connecticut. I retired from a job as an accounts payable clerk at a small company called the Institute of Children's Literature just over two years ago, and I know spend every other week looking after my two

grandchildren and sharing them with their other grandmother.

I have no financial associations with Centocor or Johnson and Johnson. I'm just a grateful patient who wished to add my voice on behalf of Remicade. I'd also like to add that I am one of the 428 original patients of the ATTRACT trial.

I was diagnosed with rheumatoid arthritis approximately 10 years ago. I was originally treated with standard anti-inflammatory drugs such as Placquenil and Clinoril. Over the years I also took nonprescription medicines such as glucosamine and chondroitin and something called Oxygen for Life, anything to try and fight this insidious disease.

I was, of course, very concerned about the progression of RA and how it affected my everyday life. My hobby and my passion for the past 18 years has been tap dancing, and I voiced my concerns to my physician. The RA had attacked my ankle and I had been forced to wear an ankle brace and was only able to wear the flat, sensible shoes. My doctor advised me to stop dancing and find a less damaging hobby. I said that that was not an option and that I had come to him for help to enable me to keep dancing and walking.

With each successive flare-up, the arthritis

would be worse and it spread to my knees and arms. My neck was so compromised that sleeping was difficult and my neck had to be supported. Driving was also difficult, having to turn my head. Dancing was also becoming increasingly hard to do and many times I was forced to sit and take notes. Once in a while I would have cortisone shots just so I could get on my show.

I was fatigued, depressed, and very concerned about the future, as was my family, because there didn't seem to be any hope. My husband feared that I would be in a wheelchair sometime in the very near future.

I was taking methotrexate but this time, one of the strongest drugs used for RA. But this also had failed to control the disease of seemingly the progression of the damage being done to my joints.

I was also on prednisone for several months but really hated the thought of being on such a destructive drug, as I have heard of the long-term side effects of this drug and felt that I had reached the end of the line in drugs used for the treatment of RA. At my request, my doctor took me off the prednisone and within weeks the RA was back.

At this time my doctor asked me if I could participate in a research project for infliximab. It had already been in testing and preliminary results were

encouraging. It was a difficult decision to make and I naturally was quite hesitant, but I had read all the literature and decided I didn't have many alternatives as so far as the disease had not progressed to the point where there had been any total destruction of my joints, all luckier than many RA patients. So, I agreed and then kept my fingers crossed that I wouldn't be given the placebo.

In November 1997, I had my first infusion in my doctor's office among friendly and familiar faces. Over that time I went back for my second infusion. Two weeks later, my morning stiffness had disappeared completely. I knew immediately that I was not on a placebo. Maybe I am one of the fortunate ones taking this drug, but I have never felt better and have not suffered any side effects, nor have I had any recurrence of the RA symptoms. I am now back to dancing as well as I was before the RA started, which is something I never thought could happen.

After learning about the hearings that were to take place this week, I felt obliged to appear before this committee and tell my story in the hope that others can benefit from my experience because that is what Remicade has given to me and my family. Hope for the future.

Thank you very much.

DR. SIMON: We'd like to thank you for your comments and we applaud you on your fortitude.

Can we please call to the microphone Regina VanDervort? I don't know if I pronounced it correctly, so I apologize.

MS. VANDERVORT: You did pretty good. Thank you.

Good morning. My name is Regina VanDervort. Thank you for allowing me to speak about my experience with Remicade. I come as an individual not associated with any business.

I'm a 40-year-old with chronic acute rheumatoid arthritis. I was diagnosed 15 years ago. I've had several surgeries and quite a struggle with it. Over the years, I've tried most of the NSAIDs, including also gold injections, Plaquenil, a lot of prednisone. I have pretty bad osteoporosis from it. At the time it was my only solution.

Seven years ago, I began methotrexate with good results. Two years ago, however, it seemed to lose its effectiveness. My doctor at the time added Plaquenil and Celebrex to it, but my arthritis continued to worsen until I had to quit my job as a surgical tech. Soon I needed help bathing and dressing and even help turning over in bed.

At this time I sought help from the doctors at Johns Hopkins. They added high dose prednisone which at

least helped me sleep more than the four hours a night that I had been.

In July of '99, I started Enbrel. Six months of this therapy yielded little results. I still needed knee braces, wrist braces. I couldn't lift a cup of coffee or climb a flight of stairs.

In January of 2000, I began Remicade therapy. Five days after my second infusion, I had an excellent response. Morning stiffness was completely gone. My energy level soared. Strength and joint function dramatically increased. This level remained for about three weeks and then it dropped just a little.

I've had five treatments so far, and my current level of functioning has stabilized and is very acceptable. I can take care of myself and my home. Last week I went on vacation with my husband and teenage daughters, and I actually hiked a mile a day for several days in a row.

During the Enbrel and Remicade therapy, I've remained on the same supplemental drugs, methotrexate, Plaquenil, Celebrex, and prednisone. Recently due to my Remicade response, I've been able to cut the prednisone dose in half. I hear people grouping Enbrel and Remicade together. I know they have a similar action, but in my own personal experience, I responded very differently to these two drugs. Therefore, I am extremely grateful that I had

the opportunity to get the Remicade treatments.

I sincerely hope that the FDA, Centocor, the insurance companies, and doctors will work together to make Remicade available and affordable to many others like me who need an effective option to fight rheumatoid arthritis.

Thank you.

DR. SIMON: We congratulate you on your response, and we thank you for your observations. Hopefully, all of us will work together to make access to therapy a very reasonable alternative.

At this point, we'd like to ask if there are any other people or persons that have any comments to make in this open public hearing?

(No response.)

DR. SIMON: If not, we will then move on.

The next session is going to talk about the questions that have been provided by the FDA. There are a couple of guidelines that I'd like to review first. Specifically that's relating to voting. The gentlemen on my left, unfortunately, although we are delighted that you're here and we look forward to open and honest and energetic discussion, can't vote. So, oh, well.

Secondly, I'd like to encourage everyone, as we discuss these rather lengthy questions, to take advantage of the expertise around us, including the company, to

ensure that we get answers to issues that might be coming up as we discuss some of the questions. Also, don't forget that we also have a large amount of expertise over on the FDA side that we'd like to take advantage of as well.

So, I'd like to draw everybody's attention to question number 1. Question number one has to do with the database itself, the size, the completeness, the numbers, the dropouts, what that dropout rate, for whatever reason, might do to our interpretation. We've heard both the company and the FDA present discussions that, in fact, highlight different aspects of that various different dropout rate and the implications of that.

There is a summary here that states that a total of 340 patients received Remicade in the ATTRACT study. Radiographic data from pre- and post-treatment x-ray films were unavailable in 16 percent of these patients. 10 patients were ultimately unevaluable for analyses of radiographic outcome because of a history of prior foot surgery. Despite these study limitations, a number of analyses clearly support the robustness of the data with regard to structural outcome measures. I would assume that almost all of us -- I am sure all of us -- are impressed with the robustness of the changes that we observe.

Now, what we'd like to know is a discussion on

the size and completeness of the database that we have been exposed to, and do we believe that the database is of sufficient size and, more importantly, quality to allow a determination to be made about the benefits of Remicade on radiographic progression in the patient population?

I'd like to point out two things about this. One is that this is a patient population that has failed methotrexate therapy, number one. Number two is the issue of progression. We've already had some discussion about progression, and it is critical for us to discuss this further because of what the sponsor has requested, which is the change in the label to reflect not necessarily a delay in progression, but actually a halting or lack of progression at all. Although that may be perceived to be splitting hairs, one just has to think about the possibilities of advertisements associated with a label that allows people to believe perhaps that we halt progression of disease. And it has to be in enough patients to make us feel comfortable that in fact that's true. This we will come back to again in the second question.

So, going back to this, do we believe that the data set provides enough information, given the dropouts and other vagaries of the data, to make us feel comfortable about the delay in progression of disease? Or

alternatively, we're not impressed with the vagaries of the data set and it's really not an important issue. We're very impressed with the robustness of the data, and we should move on to the second question.

DR. WHITE: I'm never one to be shy.

My concern with it, Lee, has to do with the issue of quality and it has to do with the issue of blinding. I remain concerned that, despite reassurances that the radiologists couldn't read soft tissue swelling and have a sense of who was on drug or not, I think it's still possible that the radiologists might have been unblinded to who was on treatment and who was not. And I think that might have skewed the data. When I think of the range of 8.6 in terms of what's interpretable as a significant difference and that the, quote, confidence intervals, if you look at them that way, might be as small as 0 on one end of the scale, if you were to throw into that some unblinding of patients, it gives me cause for concern.

DR. SIMON: Well, we have a large number, relatively speaking of radiologists or people interested in radiographic progression in our midst, some of whom have spent their lives doing this. Recognizing that there are limits to technology, we also have to recognize what is the best technology we have available to us right now.

Would someone care to comment from the table in the front?

DR. SCHWEITZER: Mark Schweitzer, radiologist.

I can understand a concern about seeing the soft tissue swelling and also even periarticular osteoporosis, and maybe that may go away and that's not something that was graded, but it may be something that someone perceives or even doesn't perceive consciously but affects the unconscious interpretation.

But I still think having the three x-rays together and not knowing the date of the x-rays, I think that there's probably some perception of a soft tissue swelling, but I don't think it would have unblinded them to a degree to make the data not usable in my experience.

DR. WINALSKI: Carl Winalski.

Basically you have three choices. You can either read all of the data together knowing the chronological order, and that will bias you towards progression of disease, or you can have them completely blinded and read completely separately, which will bias you towards showing no progression of disease. And this is kind of in between. There's no way to do them as a set and bias towards progression, which would vote against the claim they want to make, without having the potential of this unblinding due to soft tissue swelling.

So, I think it's getting back to the noise of the data. The 8.6 for the minimally detectable difference is what's seen in other studies. So, for me the data is as good as it can get, if you will.

DR. SIMON: Could you comment on the minimally detectable difference versus the minimally clinically important difference and whether or not you as a radiologist think differently about those?

DR. WINALSKI: I do. I think the minimally detectable difference is a measurement thing where you're going to compare your data. For a minimally clinically important difference, I don't think that has been defined, and I think there are so many variables in what causes patients' symptoms that a radiographic test is not going to be able to do that.

DR. SIMON: Since we're graced with the individual who did all the seminal work in the field of x-ray analysis, I'd like to ask Dr. Sharp just to make a comment about how he handles swelling on an x-ray as it relates to blinding of the x-ray system.

DR. SHARP: I never pay any attention.

(Laughter.)

DR. SHARP: I think that the quality of films has to be consistently extraordinarily good for this to be a factor. I think basically I don't look for soft tissue

swelling. Now, occasionally I observe it when it's pretty obvious, but I don't think it would unblind it. I can't imagine the circumstances that it would.

The films being randomized as to sequence and blinded as to sequence, plus the observer being blinded to treatment, assures that you've got the most objective read you can.

Now, in "the old days," if we saw definite progression in three or four joints, we knew which film was the first and which was the second. Since we've got more effective drugs where at least we have to consider the hypothesis that there can be healing, one has to keep in mind that even though you see a difference, if one film is worse than another, it does not necessarily give you the time sequence. I think anybody approaching a set of films today would be a little bit on rocky ground to assume that were the case.

I'll comment about the minimal detectable difference and the minimal clinical difference while I'm up. I think that minimal detectable difference is a conservative statistical measure of error. I happen to believe that any real progression in radiographic damage in a patient is clinically important.

There are a lot of people who say, well, you've got one new erosion over a year's time. My patient is

about the same, so is it really clinically important? If you have a machinist as a patient who loses one finger in an accident, depending on which finger it is, he may have very little change in his function. If he loses his whole hand or all fingers on one hand, he's got a real problem. Now, over time we anticipate that if you're having one erosion or two erosions in a year, in 10 years you're going to have 10 erosions or 20 erosions, and by then it becomes really important.

So, basically we're looking at what I call a footprint of the disease. The inflammation is the disease, but we're looking at an erosion, which is the consequence of inflammation over a period of time, and we're trying to predict what's going to happen over an extended period.

DR. SIMON: Since you brought up error and since part of that error, when you have two individual readers reading independently, that's going to be part of that. But isn't there error also in the sense of the extent of deformity and physical disease that patients manifest in the ability to reproducibly perform exactly the same kind of x-ray each time in setting it up? And if that's the case, how does one take into consideration that error, particularly in studies that go over one or two years which might in fact infer change over that period of time?

DR. SHARP: Change in position of a hand, with or without deformity, is a problem in comparing films, and I think only experience will teach you how to deal with it. I think some people are more conservative; some people are less conservative. I tend to, when I look at two films, ask myself could a change in position or extent of exposure or development of film, whatever, quality of film, account for this difference, and if I think it is, then I'm much more cautious about scoring a difference.

DR. SIMON: Our guest experts, do you have a comment about this?

DR. SCHWEITZER: Yes, I want to make several comments.

First off, usually they do standardize, and in a protocol, they standardize the film and the development and the screen. I usually believe they use a template for each patient, specific for each patient, to get rid of the error from potential changes in positioning, albeit if they have subluxations that wax and wane, then those templates don't work. And I understand that.

Getting back to Dr. White's question, I think also part of potential unblinding beyond the soft tissue swelling is seeing both hands and both feet together because you can kind of develop a gestalt for if the patient is progressing or not because you have a fair

amount of data there to look at. Kind of a pure model is just looking at one hand at maybe all three time points and being blinded for the time points, but just one hand by itself because then there's less chance of the other hand seeing some improvement and then looking more carefully for changes in the contralateral hand.

In reference to the first question about the clinical relevance of the x-ray findings, I kind of look at it as two different things. I would love x-rays to be clinically relevant and all imaging studies to be clinically relevant all the time, but the reality is that they're not. It's an anatomic way of looking at a person, and it is related to their function. It's related to their symptoms. It's related to their signs and the laboratory data, but it is kind of an independent function of what their anatomy is doing. In some cases, usually it does lag other clinical findings, and it's kind of interesting that in this situation it may not lag the clinical findings or it apparently does not lag.

DR. KATONA: My question is for the FDA colleagues. I just would like to introduce one more question to all this puzzle.

The proposed label indication reads as prevention of structural damage. We're living in the 21st century. Even if we accept that there is a change and a

significant difference of radiological appearance, to me as a rheumatologist, that does not mean that there is no structural damage. By the time you see something at the x-ray, a lot of structural damage occurred, and we've all seen MRI scans and other modalities.

I think this might be a moot point today, but at one point we need to discuss that structural damage might be better served if we would say radiological damage or somehow better define it because, to me as a patient, it would be very reassuring taking a drug and say that everything is going to stay as it is. I just don't believe that there is any drug on the current market which will do that.

DR. SIMON: Could we ask our MRI local expert to comment on the benefits, Carl, of MRI over x-ray at this technological development stage?

DR. WINALSKI: I would say at this technological development stage that MR is in its infancy compared to the radiologic and x-ray data. The error bars that we have on that for determining structural damage, though perhaps MR will be more sensitive for detecting early change or pre-erosion change, I have not seen longitudinal data to show that the MR findings do predict or herald true erosions and true bone destruction.

So, at this point I think it would be early to

be throwing MR into the mix, but hopefully there will be studies to bring it in because it's quite clear that you do see MR signal changes in the bone that are not evidenced by radiographs. I have heard some anecdotal data that some of those resolve without becoming erosions radiographically, but I think all of that is yet to be shown.

DR. MILLS: Lee, several things in terms of what I've been listening to. First of all, I've reviewed all of the x-rays in this study. I could not pick up the soft tissue change to be concerned that you would break any blind here.

I think, though, that Dr. Sharp's comment should be taken very carefully and listened to, which is that when you're looking at a series of x-rays at random time points and you see in one evidence of erosion and in another you don't see it, you begin to start to smooth your findings a bit in terms of raising the concern as to am I looking at a time point or am I looking at a resolution.

Here we find that some of the evidence is represented as negative values for some of these patient responses. You have to be careful that some of this may be noise in terms of the interpretation at multiple different time points without knowledge of those time points. As a result, you may feel that you cannot see an erosion. You may feel that there is no erosion, but we also know that

looking at just standard posterior and anterior radiographs in two dimensions may disguise an erosion change, and you may not pick it up. So, again, you may have some softening and smoothing of this data. Some of the negative numbers that are being presented here are actually possibly related to this phenomenon of looking at random time points. As a result, you may be missing some of the findings, which if you were looking at them in a structured time point evaluation, you would identify.

Remember there was a comment about Dr. van der Heijde's own article stating that once an erosion, always an erosion. At this time, as we're looking at these, we're looking at them in a different time sequence, and Dr. Sharp said in the old days, which were only about two years ago --

(Laughter.)

DR. MILLS: -- indeed, an erosion was always there. So, if we didn't see it on the radiograph, we declared it still there. Now we're saying it may not be there, and part of this phenomenon you may be seeing is this miss in terms of the positioning and the use of only a two-dimensional radiographic evaluation.

I hope that the MRI will get us there, but indeed we don't have the data to be able to support it right now. So, we're limited in terms of our evaluation

model.

The other point was in terms of radiographic change. We have soft tissue, we have cartilage, and we have bone. For all the world, we can see the bone. I just told you you can't see the soft tissue, and the cartilage we kind of intuitively discuss. So, be careful in terms of how much you want to put in terms of this information because, having looked at the x-rays, I can sure see the bone, but I sure couldn't see the soft tissue, and I was intuitively talking about cartilage.

Thanks.

DR. SHERRER: Hi. Yvonne Sherrer.

I just wanted to comment on that because those are my thoughts as well, as Dr. Katona and you just mentioned. As a clinician, you can see structural damage apart from what you see on bones. There are tendon ruptures and so forth. We see that there's somewhat of a difference in the data in terms of the response looking at x-rays versus the clinical response because some of these patients only had an ACR20 response, which means they continued to have swelling and pain, and yet apparently those same patients did not go on to have progression in terms of bony changes.

Now, what does that continued inflammation for those patients mean to them and to me as a physician in

terms of my approach? To me, the way this is worded, it would suggest that I didn't have to be worried about structural damage in those patients who have ongoing inflammation, and yet intuitively I know that I probably do. So, that's why I would want some clarification here.

DR. SIMON: As chairman, I always tend to be sensitive to where the discussion is going, and we've moved into the discussion on number 2. I'm happy to do that. I just want to make sure that we all have gotten out our feelings about number 1 and have resolved it. We're not going to be taking a vote on number 1.

Before we go on to talk about Dr. Sherrer's comments, which I think are very appropriate, do we feel by consensus, just to settle the issue, that there's enough data here in this robust analysis to make us feel comfortable about discussing number 2, that in fact, there may be some issues regarding progression of x-ray damage? Is there anybody who is feeling that there isn't enough data here to achieve that?

(No response.)

DR. SIMON: Seeing no response, I will assume we can go on to number 2. Is that okay?

DR. SCHWIETERMAN: Yes, that's fine.

DR. SIMON: Question number 2 is the crux of the day to a certain degree, although there are issues

otherwise. There are several different things that are inherent to this question. It's important to recognize that to date we have several different therapeutic options that have been approved just in the last 18 months that have received a label suggesting a delay in progression of damage. This is the first time we are being asked to say or imply that there is no progression of damage with this therapeutic approach.

In that context, we have multiple different levels to consider, one of which is the scientific evidence. Is there evidence that shows there is no progression of disease over the time course studied? Is that no progression of disease in one year truly indicative of disease that lasts for a long time, since we all know that this is not a cure, and thus is this one window of opportunity and observation going to be reflective of 20 more years of Remicade therapy?

I think the third issue is partly related to the technology, partly related to our ability to reproduce the data, and partly related to the exact trial that we are discussing. Is this patient population actually extrapolatable to a degree to any other patient population? They're not asking for that. However, in our making a decision one way or the other, that will be done regardless. And is this patient population truly different

than patients who respond to methotrexate early on?

In that, I'd like to make one request. I am still confused about medians versus means, and since the data is expressed as medians and means, and the FDA seems to put more emphasis on means, at least the way it was presented -- it may not be currect -- and the sponsor is putting much, much more weight on medians, I just need to have some statistical analysis here to help me with that.

DR. SIEGEL: Actually I think the numbers that George read off most of those slides were also medians. I think we're in agreement. We put the means on the slides.

But the data are significantly skewed. There are some very large numbers, like some 61's. I think it was the highest number of progression. When you only have some 60 some odd people in the group and one of them is a 61 progression, one person influences the mean by a whole unit. So, if you're looking at a central tendency for this sort of skewed data, median is probably more informative.

I think also the actual analyses were nonparametric. Right? So, median makes more sense in the setting of a nonparametric analysis. We're actually in agreement there.

DR. SIMON: But given the range of change in each of the patient populations in each therapeutic group, lack of progression would imply that all, most, some don't

progress? How do you take that?

DR. SIEGEL: That's a different question certainly from the one I answered --

(Laughter.)

DR. SIEGEL: -- and not the one I understood. But I think it's an important question. It's a focus both of this discussion and of the discussion this afternoon.

I guess I would simply say, regarding your background comments, that I would slightly correct and say this isn't the first time we've been asked for a claim of prevention. Working together with this committee in the old days --

DR. SIMON: Was that four weeks ago?

DR. SIEGEL: -- the guidance document described such a claim and described under it issues such as slowing x-ray progression. But I think as you know, at a prior meeting, as we discussed what prevention means, there was some concern around the question you just asked. Does prevention carry the implication that nobody is progressing, that it isn't happening, that it won't happen, whatever? What are the implications of the use of the word? I'm not going to answer that question because we're asking it today.

DR. SIMON: Dr. Elashoff.

DR. ELASHOFF: Apropos of that question, the

data show that whatever changes we're looking at in the ones that are significant are significantly less in the treated groups than they are in the placebo group. There has been no analysis shown which addresses the question of whether in any particular group there is "progression" or lack of progression on the median or for individuals. So, they are entirely two separate things.

You can show that there's less for the placebo group based on the analyses presented. The whole question about whether there's some or none then has all these details of how one would address that question. Would it be on the mean? Would it be on the median? Would it be on percentages of people? Then you have all kinds of cut point issues. As a statistician, I don't like to get into cut points at all. So, that's my comment on that.

DR. SIMON: But didn't we hear from George that there were more than 40 percent, but not 50 percent, that did progress in each arm?

DR. MILLS: You're referring to the sensitivity analysis where we selected. Again, in that we're making an adjustment in the data set. The numbers across the board were approximately in the 40 percent range. I can pull that slide back up if you'd like to have them bring that up. It's the sensitivity analysis. It's the percent progression, the fourth one that we had there.

DR. SIEGEL: Well, I guess you put it in your sensitivity section. But, yes, if you're talking about the percentage of people who had a higher score at 54 weeks than at 0 weeks, right. That was between 40 and 50 percent on the Enbrel arm.

(Laughter.)

DR. SIMON: Infliximab arm.

DR. SIEGEL: Thank you. Sorry.

DR. SIMON: David?

DR. WOFSY: I'll identify myself again. I'm David Wofsy. And I identify myself because I'm really asking a question that comes from someone who's new to these proceedings. It is to some extent a semantic question.

The indication about signs and symptoms is given to an agent despite the fact that a third of the people don't have improvement in signs and symptoms. So, how does that apply to consideration of an indication to prevent bony erosions or structural damage, however you want to word it, if some people in fact have it prevented and others don't? Wouldn't that be the same as some people having responses in signs and symptoms and others not?

DR. SIEGEL: I think that's a very good question. It's true that it's probably true -- it's true to my knowledge -- that virtually no drug does what it's

intended or hoped to do in all patients. We give a lot of claims based on -- the question has to do, though, with whether the word "prevention" per se -- I think the question that was raised and discussed with this committee -- carries, perhaps not in any statistical or clinical trial sense, an implication to the consumer or the physician that there is an absolute effect.

As we were discussing this earlier, I noted, for example, that we have drugs that in treating heart attacks reduce mortality. So, there's a significant difference. But still people die, and we don't say that they prevent mortality due to heart attacks. One might be concerned that if you said they prevented mortality that people would think that if they took that drug, they'd have no chance of dying. There we say "reduce."

DR. SIMON: Bill?

DR. SCHWIETERMAN: Go ahead, Harlan.

DR. SIMON: Could you identify yourself?

DR. WEISMAN: Yes. It's Harlan Weisman and I'm from Centocor.

As a frame of reference -- and I intend nothing more than that. Jay, this is something that you and I discussed earlier. I'm reading from the package insert or the prescribing information for Actinel, which is a drug intended for use in patients with osteoporosis. Let me

just read two of the indications in the package insert.

Postmenopausal osteoporosis. "Actinel is indicated for the treatment and prevention of osteoporosis in postmenopausal women." That's one indication.

The other one is glucocorticoid-induced osteoporosis. The reading is "Actinel is indicated for the prevention and treatment of glucocorticoid-induced osteoporosis in men and women," and then it goes on to describe the population.

Just to clarify from the sponsor, it was never our intent to claim that Remicade works in all patients either for treating signs and symptoms or for prevention of structural damage.

In fact, Dr. Harriman tried to make a very clear point of what our operating assumptions were here. We looked at the guidelines. We designed a clinical trial to obtain indications according to those guidelines by defining three very clear clinical endpoints that you've seen. One of them was structural damage. We used the guidelines. We defined the primary endpoint, and we defined very clearly what the criteria were for assessing whether that was a positive result or not in concordance with discussions with a learned body of experts, who were our steering and chairman of the trial, as well as with the FDA.

I have to give credit to the FDA because many of the people at the table over there substantially contributed to the design of the ATTRACT trial and substantially contributed to the endpoint definitions, all of which we decided, before the fact of the trial and before the fact of the analysis, would constitute a positive trial, constitute demonstration of efficacy for the indication we were seeking.

The language we are seeking seems to be in accord with other language that has been used for other products such as osteoporosis, which does have analogies to rheumatoid arthritis because we're talking about x-ray evidence or at least density.

DR. SIMON: We applaud your hard work. However, I'd like to point out that the technology is very different in the two fields. Consensus has been achieved in the osteoporosis field about those particular areas. I would argue that we have achieved consensus upon applying this technology to either the idea of prevention as opposed to delay, as well as to healing. I think that that's where the rub lies.

Your observations have well outpaced our ability to develop a technology that helps us understand this. This has implications regarding a minus score, which is interpreted by some to mean something very positive, and

by others to not know what it means as a minus score. That's very different than in the osteoporosis field.

DR. SIEGEL: I should say with regard to this issue, though, of examples of how the language has been used in other labels, that the agency does not uncommonly label vaccines as for the use and prevention of disease, and for any of a number of those, there are some case occurrence rate in the vaccine-treated arm, and the lack of absolute prevention has not inhibited the use of that term, although most of them have a very high rate of reduction compared to control.

DR. FIRESTEIN: I think the points made on prevention versus delay are very important. Setting aside the statistical arguments, the main concern that I have has to do with the natural history of rheumatoid arthritis and the duration of this disease and what one year or even two years means in terms of truly preventing versus delaying structural damage. If one looks at the very interesting and exquisite data set from Dr. Wolfe, which was shown earlier, you see a 20-year evolution of this disease with regard to structural damage. There's not even a little tick mark at 1 year. It's too early to say that.

Also, there are a number of studies that were published in the last couple of years looking at radiographic evidence of damage on individual patients over

a 10- to 20-year history and how variable that can be.

The long and the short of it is that we're really looking at a very narrow snapshot and that 1 year of a lack of progression doesn't necessarily mean prevention of progression. It means just that, that within the first year of a 20-year disease, you're not seeing radiographic changes but certainly on the order of 3 to 5 years are needed to make some definitive statement about that.

DR. SIMON: Carl?

DR. WINALSKI: With regard to changes on radiographs, the measurement even for a single joint is an average of all sorts of changes around the joint. If you watch one erosion disappear either because of positioning or because it really did heal, but the formation of a new erosion, that joint may not have progressed radiographically. I think that's an important difference in trying to say it's preventing structural damage. If you go a long with what Dr. Sharp said and every new erosion is like losing a finger and eventually you lose the use of the joint, then the appearance of a new erosion, even with the regression of another erosion, to me means that there has been some structural damage. So, I think perhaps we should be talking about radiographically measurable or radiographic progression rather than actual structural damage.

DR. WHITE: I would like to make additional comments along the same lines that we're hearing. I think it's very important to convey in the change of wording exactly what was observed and to not convey more than was actually observed. I agree about the issue of time. Prevent is defined by how much time you have to follow them. You can prevent for a year, but you may not prevent for five years.

So, I don't have real problems using "prevent" in language, but I would feel more comfortable if a time period were included in that kind of a prevent statement because that's the truth and that conveys what was actually observed and is less likely to be open for misinterpretation.

I would also feel more comfortable if, rather than a global term, whatever is more appropriate were used in the wording. "Prevent radiographic progression" for a year might more accurately what was observed than "prevent structural damage."

DR. SCHWIETERMAN: This has been a very helpful discussion. I'd just like to point out that actually we spend weeks and weeks and weeks of time at the agency over the appropriateness of particular label claims, recognizing that words have a great deal of impact. So, I'm not sure to what extent diminishing returns will come in here, but

suffice it to say that the fact that the RA guidance document was written at a time before any of these agents had been developed is somewhat culpable here.

The second part is that, in fact, the word "prevent" has a great deal of charge, cachet, because there are a number of competing products in this and because -- and I think this is really the underlying issue here -- it's really not erosions and x-ray damage we're talking about. It's functional disability in the long term that we're trying to prevent. To the extent that the prevention claim at all connotes a long-term benefit -- and perhaps that was Dr. Firestein's comment -- that can or cannot be misleading.

I think all those things need to be considered in the label when we write this. I just want to make the point that we need not necessarily resolve what this word ought to be now but, rather, to point out the pros and cons of it.

DR. SIMON: In discussing this further, I'd like to go on to the first highlighted question. We are talking about a modification of the total Sharp score, which has components, erosions and joint space narrowing. We've heard allusions today that the biology of damage related to erosions and the biology of damage related to joint space narrowing may be slightly different or may be

significantly different in how it's carried out, the cytokines and other chemicals that are associated with it, as well as whether or not it progresses exactly the same in each person or in the same disease.

So, the question has to be now that we've seen such data, is it still important to have a measure be a total joint score, a total Sharp score or its modification, or do we also want to consider important aspects of components of the total Sharp score or its modification, such that could someone present data on statistical improvement in erosions but yet have no statistical improvement in joint space narrowing? And that would be statistically important, but because the joint space narrowing is not important, perhaps the total score is not important statistically but yet they still have important changes in erosions.

How do you all feel about the component in a secondary analysis of the components of the total score?

DR. ELASHOFF: Relevant to that, I would like to ask if either the company or FDA has information on the individual correlations between erosion and joint space narrowing scores or between changes in these two scores across patients because how highly correlated they are ought to be relevant to whether you want to break them down or not.

DR. SIEGEL: We've seen in another setting with a different drug, that I won't get too specific about, which hopefully will remain nameless --

(Laughter.)

DR. SIEGEL: This time, right. Exactly.

Some apparent differential, in comparison between treatment arms -- you know, whether they're statistically significant interactions I can't speak to, but what would appear to be a difference, as treatment arms were compared, in relative effects on joint space narrowing versus erosion. So, at least those data would suggest whether it's a result of the particular patient, the stage of the disease, the mechanism of the drug, or whatever, there might well be dissociation of the two.

DR. ELASHOFF: But the question of what the correlation is on an individual basis is answerable and should be answerable in connection with this kind of question.

DR. SIMON: I suspect we have someone to present something like that.

DR. HARRIMAN: We would like to just say that we're going to try and get that data for you and hopefully we'll have that as soon as possible, perhaps after the break.

But in the meantime, I think Dr. Wolfe might

have something that he would like to say in this regard.

DR. SIMON: The group recognizes Dr. Wolfe.

DR. WOLFE: They are very highly correlated, of course. In looking not at this particular data set but the one that you saw on the slide, we had the opportunity to compare both the Larsen and the Sharp measures. With appropriate standardization, they are correlated at about .9 something. My remembrance is that the correlation is about .8 for the two separate measures here.

They do separate out. We've looked at this in terms of doing a rash analysis and plotting all of these points. They are separate in that sense. They contribute additional information.

The reliability of the measure is much higher. It depends on the size of the data set, but reliability of the Sharp score in the set that we presented previously is well over .9. The reliability coefficients for each component individually is significantly less than that. So, there's an advantage of using both of them together.

DR. SIMON: Thank you, Dr. Wolfe.

Gary?

DR. FIRESTEIN: With regard to this specific question on the secondary analyses and the various components, first of all, you're quite right. The current thinking is that there are distinct but overlapping

mechanisms for bone destruction versus cartilage destruction, and it's not worth going into the basic science of that now. Professor Maini presented a little bit of that.

But we don't really understand which ones are more important with regard to later functional disability. One can try to guess, for instance, in knees which are not looked at in this analysis. But in knees erosions are nearly as important as joint space narrowing or loss of cartilage. In hands it may be erosions with ligamentous laxity that cause deformities.

So, until we have some notion in terms of how each specific component of the radiographic scores impact functionality, it's hard to say that we should just be looking at the total score or whether we should continue to ask people to look at individual components. So, I think for future analysis, we still have to look at that.

Then there are these very interesting questions in terms of why in other studies with drugs that might have similar mechanisms, you would get different types of results. And I don't have an answer for that at all.

Finally, the issue of how patients cannot have clinical improvement but still have evidence for lack of progression on radiographs is on the surface surprising, but actually there's a very long and distinguished history,

looking at the distinct inflammatory and destructive processes in rheumatoid arthritis, going all the way back to corticosteroids in the '50s through nonsteroidals and many other agents.

I think really the results of this study underscore that, and that is that you can make people feel better and not have an effect on their radiographic progression, or now you can potentially improve outcomes in terms of radiographic progression but not make them feel better.

DR. SIMON: So, I guess then the information is as we've discussed it. Bill, is more to be gotten out of this particular question for you?

DR. SCHWIETERMAN: If there's no consensus on this, that is to say, that there's no data really to support these things, perhaps that's the answer.

But we have, with different products, looked at different outcome measures and so forth and have, in fact, by virtue of those outcome measures, incorporated them into our analyses of the overall safety and efficacy of this particular product. Is it my understanding from this committee that that's something that's worthwhile doing, or is that something that we really shouldn't be broaching given the fact that the data aren't there?

DR. SIMON: Looking to discover and learn, I

would highly recommend creating labels. Until we understand what is going on, I would not. I could imagine, in the not too distant future, a targeted therapy that may actually target erosions. It may have no effect on joint space narrowing. Maybe. And under those circumstances, you would cripple the ability to develop that drug if you weren't willing to look at the secondary analysis in that regard.

DR. SIEGEL: You said secondary, but I guess part of this question would be if that sponsor came to us and said they want their primary analysis to be erosions, because that's what they're targeting, that's their treatment targets, would this committee find that problematic when we then presented the data to get a claim based on that as the primary analysis? That's part of what's in this question I guess. Or joint space narrowing or any other.

DR. SIMON: Dr. Katona, do you have a comment?

DR. KATONA: I think the discussion was going that we really do not know clinically which one is more relevant, how does it correlate with symptoms. Until we do, absolutely we really need to look at both, as well as it's nice to look at the two together. So, at the current time that our knowledge is, I think it's absolutely important that the agency request both.

DR. SIEGEL: Well, yes, we'll get both. And either can be prespecified as primary? Is that what you're saying? Since we don't know which is important. Or are we saying the total should be or it doesn't matter?

DR. SIMON: Why don't we expect that, if it ever happens, it would be incumbent upon that particular sponsor to demonstrate the functional correlative outcome related to their particular intervention. In that manner, we actually may advance both regulatory science and real science.

DR. FIRESTEIN: But that's a much more difficult proposition because that's a 5- to 10-year study as opposed to a 1- to 2-year study.

DR. SIMON: The peanut gallery has a comment over here?

DR. JOHNSON: A backdrop to this whole conversation is what does an x-ray assertion mean. Period. I think when we put together the document -- it may have been five years ago, but I think it still holds, that we don't know for sure what they mean. That's why we wanted a clinical correlate to go with it. So, these are all kind of dependent claims. They're contingent claims. So, I think that should be kept in mind. I think that's the point that a number of people have been making, as a matter of fact.

DR. SIMON: But, Kent, that also raises the issue, as per the sponsor's comments before it, my first question. If then we're going to link these outcomes, then the functional outcome needs to be linked in a way that's doable, and we've accrued more and more data on several products that in fact functional outcome changes can be measured in a shorter period of time than we previously thought. Now, whether or not these functional outcome data, like the x-ray data, are only just snapshots of outcome really remains to be answered.

DR. SIEGEL: Well, in that regard, Kent of course is right. These are contingent claims, but as the guidance document is written, they're not contingent on long-term functional or disability outcomes. They are contingent on clinical benefit outcomes, which is to say our current approach is we would not give somebody a claim based solely on radiographic changes if it weren't accompanied by evidence of clinical benefit. However, we will approve a drug, as has happened with drugs in this field, based on signs and symptoms data and then look at the radiographic claims without having what you're asking about the long-term functional disability data in hand. That's where we are at present, in any case.

DR. SCHWEITZER: Also, trying to think about the future with MR imaging or ultrasound or even

scintigraphic imaging, you may have a claim for decrease in effusions, decrease in synovial proliferation. That may, in turn, be a better marker than the x-ray.

DR. SIMON: Or not.

DR. SCHWEITZER: I'm kind of a splitter in that way. I think erosions and narrowing really should be separate. I think everything should be separate because there may be some agents that affect joint fluid, some agents that affect synovium, some that affect erosions directly. I think I'm kind of a splitter in looking at each individually.

DR. SIMON: Last comment. Carl?

DR. WINALSKI: I was just curious because it seems to me that the drugs that have been approved for osteoporosis, getting back to that analogy, are being done just on a radiographic, or is that also to show that you have decreased fractures?

DR. SIMON: Both. It depends on the claim that they go for, and there are stringent criteria for both x-ray fracture, which is arguable but, nonetheless, stringent, and clinical fracture, as well as densitrometric change, which actually has a World Health Organization imprimatur of diagnosis. So, prevention of osteoporosis is based on not achieving a densitrometric diagnosis of osteoporosis. So, that's where that all comes from.

DR. SIEGEL: From a legal and policy perspective, the agency does not absolutely require clinical benefit to approve a drug. We will approve a drug on a surrogate endpoint that is validated to predict clinical benefit, sometimes rigorously, sometimes based on historical data and presumption. A lot of, say, antihypertensive drugs are approved on blood pressure rather than on stroke and mortality data.

We will also, in certain cases, especially for new and improved therapies for serious diseases, give an accelerated approval based on a surrogate that's not fully validated but reasonably likely to predict benefit.

So, as you go to other diseases, from the perspective of the law and the policies and the regulations that guide the agency, you have to look separately at the extent to which, say, if you're talking about osteoporosis, a radiographic change is felt to be predictive of and an effect on it is felt to be validated to be predictive of a clinical change. And that can differ from indication to indication.

DR. SIMON: Yvonne, last comment.

DR. SHERRER: In terms of comparing prevention with osteoporosis versus in this setting -- and maybe we as clinicians interpret that wrong. As I relate to osteoporosis and/or infections, you're preventing the

development of disease in somebody who does not have disease given the right exposures or the risk factors. Whereas, here you have people who already have disease. You're not taking a healthy person and preventing them from developing disease. You're saying in somebody who already has disease, you're preventing that disease from progressing, which seems to me is saying something different.

DR. SIMON: So, in saying something different, what clinical trials would this committee like to see to help us understand that further? What advice can we give the FDA about future data accumulation in this area to clear this up?

DR. WINALSKI: So then my understanding is at this point radiographic progression is not a surrogate endpoint for preventing clinical symptoms or signs. If that's the case, then it seems to me we need a long-term study showing whether or not the addition of more and more erosions leads to joint disability.

DR. SIEGEL: Well, question 4 in this set will ask you to address for us to some extent whether radiographic progression in the absence of clinical benefit can be taken to support a claim, which would imply that it could be accepted as a surrogate.

But, yes, your statement I think correctly

reflects where we are, where we have been, what our guidance says and the way we've been practicing. Radiographic changes alone would not get a drug approved that is not approved.

DR. SIMON: But we'd like to extend that a little bit farther. It's not clinical benefit in the context of what you're measuring; it's signs and symptoms. Clinical benefit is yet to be defined. Although we presumptively think about that as a functional benefit over time, signs and symptoms may or may not reflect that.

DR. SIEGEL: Let me say that the way I was using the word "clinical benefit" -- and this is more a semantic thing -- would include signs and symptoms. If somebody has less pain, they've benefitted, but that's more semantics than science.

DR. SIMON: Dr. Katona?

DR. KATONA: I think we are discussing two different things at least. We're discussing how to design the trials, but you're also discussing the labeling. I think what came out of that labeling for one class of drug and the terminology might be very different than for our drugs. I think to change the labeling or the philosophy about labeling -- what do we call prevention -- I think that might be a much quicker thing what we could fix because we have to be very fair to the sponsors. I think

they have to use the technology, whatever is available today, and I think we need to help you to design the labeling, but we just can't pick up something what works for osteoporosis and it doesn't work for us. And these are chronic diseases. I just would like to underline what Dr. Sherrer was saying, that this is a very different clinical setup, what we're dealing with.

DR. SIMON: Bill?

DR. SCHWIETERMAN: I think that was well said. The first part of the question is actually most of the afternoon's discussion, and unfortunately or fortunately or inevitably, we're getting into a mixture of where do we go from here with these radiographic outcome claims, including the ones that are on paper, albeit it perhaps poorly written on paper, and new ones coming down the road.

Again, I would just reiterate to this committee, we need not do the labeling at this particular meeting because there are more considerations than simply the science here. There are precedents involved. There is the context of the labeling itself, which isn't to say we should stymie this conversation. I don't want to necessarily go all the way with that.

DR. SIMON: We prepared to be able to help each other in that regard, as this meeting went on, to make sure that we wouldn't get stuck in certain areas.

I'd like to recognize Dr. Wolfe for a minute.

DR. WOLFE: Yes, just for a minute.

Again, using some of the patients you've seen presented previously from our data set, we'll present, at the ACR meeting this fall, long-term outcomes based on the rates of radiographic progression in which we show that the rate of radiographic progression is significantly associated with the rate of work disability and total income of individuals, after controlling for all of the variables. It seems to me that it's important to understand that when you're looking at radiographic progression, you're looking forward to preventing some event that occurs in the future, and that's an important functional outcome that has now, I think, been shown.

DR. SIMON: Thank you, Dr. Wolfe.

Since we're going to discuss this afternoon other ways to study this issue, I'd like to go to the one question that I'm advised we're actually going to take a vote on. That's the third part of this number 2 where they're going to actually ask you to put up or shut up. I'm allowed to paraphrase this, so that's what I'm going to do.

Do the data support the sponsor's claim that Remicade prevents progression of structural damage in patients with rheumatoid arthritis? Now, remember, there

are many caveats to the patient, because this is a specific patient population.

And the second part of this then is, to what degree, if any, can that benefit be extrapolated with this data set to patients with either earlier onset disease, less severe disease, or disease-modifying responsive diseases? Thus, patients that were not studied in this trial.

So, let's ask the question first. Do the data that you have seen and we have now grappled with support the claim for preventing progression of structural damage in patients with rheumatoid arthritis?

Barbara?

DR. WHITE: I just would like a point of clarification from the FDA before I vote, since you wrote this question. I would like for you to give me your definition of prevent. Should we vote based on the definition that you worked out with the sponsor?

DR. SCHWIETERMAN: The definition we worked out with the sponsor comes from the guidance document, and if you read the guidance document, there are two or three different ways of measuring that. So, we were viewing the prevention of structural damage claim in the context of changes in Sharp scores, but never with the specificity that I think you would like me to answer with here. The

sponsor is certainly acting in good faith with this.

I'm less interested in this being a referendum on prevention or delaying because I think that therein lies an afternoon's discussion, rather than in a vote on how this committee feels if there has been a demonstrable effect upon this agent. And we can continue to have a discussion about the actual wording if you like. I just don't think that this is the time to actually try to sort that out.

DR. SIEGEL: Let me second that comment. I didn't actually personally word this question, although I probably looked at it.

(Laughter.)

DR. SIEGEL: I always look at the questions. Did I actually word this question? Maybe I did.

(Laughter.)

DR. SIEGEL: I'm sure I looked at it. I do look at questions.

I think in light of the discussions we've had, a lot of the issues that are raised by that are addressed by this discussion. I think, as Bill has pointed out, there are many other issues that will go into what is in the label.

If I understand what you're saying, I would agree entirely that what we really need -- and maybe we

could even reword that question there -- is a vote on whether the data support the -- well, delay wasn't even the sponsor's claim, but whether they support a claim that Remicade has a favorable effect or has demonstrated a favorable effect on progression of structural damage. Then we can integrate your advice and other factors on how to word that.

DR. SIMON: Would you prefer us to restate this question?

DR. SIEGEL: I think that would be better. That will make it easier for you to vote on what we need to know.

DR. SIMON: That's fine. I'm entirely happy to do so.

So, correct me if I'm wrong. The question then stands, in the evidence presented this morning, does the committee feel that there's enough evidence to warrant the claim that the patients did better with infliximab than they did otherwise?

DR. SIEGEL: Regarding structural damage.

DR. SIMON: Regarding structural damage. You'll notice that I chose an incredibly gentle word that you can then grapple with on your own.

So, again, the males on the left side are not able to vote. So, we begin with Dr. Sherrer.

DR. SHERRER: In light of that modification, yes. Yes, it does show that the patients on infliximab and methotrexate did better than those on methotrexate alone.

DR. SIEGEL: Could I rather suggest -- I think that we don't need to ask. The data show that the patients did better.

If I might revert that wording to saying, do they support a claim that Remicade had a favorable effect on progression of structural damage. That would be the question. With the understanding that such a claim would then go into the labeling, but the wording is not fully decided.

DR. SIMON: Only modifying one question. What does "favorable" mean?

DR. SIEGEL: Well, it means that -- aha.

(Laughter.)

DR. SIEGEL: Simply that it goes in the right direction, not to imply that we know that that has a clinical implication.

DR. SIMON: Not to make this into tort reform --

DR. SIEGEL: Now, you can see why Bill says we spend weeks discussing the wording.

DR. SIMON: Then can I then restate the question one more time and then really rely on the delay in

progression as opposed to asking the question about prevention? Because in fact that's what we're saying. Is there data in this data set that at least shows there was delay in progression? I think that we can answer that in structural progression.

DR. SIEGEL: If the committee is comfortable with that. Some have suggested that delay is an issue because we don't know what happens after the trial. Others have said, however, that it's not an endpoint that directly measures delay in this trial.

Maybe what we should say is reduced or less progression. So, do the data support a claim that Remicade treatment resulted in less progression of structural damage? And we can take it from there.

DR. SIMON: I think then that requires us to say less progression in a year, because that's what the study was. Does the data show that? Then that's exactly what Dr. Sherrer agreed it showed. It was beneficial in that regard. I really would urge you to consider the use of "delay" because, in fact, that's what we're talking about.

DR. SIEGEL: Okay. I'm comfortable with that.

DR. SIMON: Now, we can add in "delay for a year."

DR. WEISMAN: Wait a minute. I think the

sponsor is not comfortable with that. Come on. I think we should have some opportunity because I think, first of all, we didn't test whether the product delayed. That was not the primary endpoint of the trial. It wasn't --

DR. SIMON: We recognize that the sponsor may have issues with that, but that is a discussion that the sponsors have with the FDA in the labeling discussion.

DR. WEISMAN: But you're modifying the question, and I guess the sponsor should have at least some opportunity about whether the rules have changed in the last 5 minutes because of Jay's equivocation. That's what I'm protesting is the equivocation of what the questions are here.

DR. SIEGEL: I'm trying to get advice on what we need advice on and trying to accommodate a lot of thoughts on that. It sounds like the sponsor is concerned about the word "delay" because of the wording in the labeling. If that's the case and they want specific advice about wording in labeling, I'm not sure that we're too happy with that. But we could have --

DR. WEISMAN: I guess what we were saying is we were happy with the vagueness.

DR. SIMON: Could I ask the sponsor to take his turn? Thank you.

DR. SIEGEL: Perhaps if you're concerned about

"delay," we could have two separate votes, first on preventing and then on delaying, and that would accommodate all interests.

DR. SIMON: Ms. Malone?

MS. MALONE: I just thought the sponsor was looking for something more definitive because obviously in marketing, if you can say "prevent" as opposed to "delay," the consumer is going to want "prevent."

DR. SIMON: Barbara?

DR. WHITE: Again, I am going to have trouble voting on either one because of really lack of firm definition. I don't feel I have been provided with a firm enough definition of prevent if it's different from what was in the document. Or delay. I don't know how we could do delay on this one because we don't have follow-up to then show that it comes up. So, delay implies that it was down for a while and then it comes up. We don't have those data, so how could I vote on delay?

I know what difference I saw for a year's period of time. I could vote on that.

DR. SIEGEL: Do you think that a 1-year study should not, at least in the future, be sufficient to give a claim? Because, see, we developed a guidance, in consultation with this committee and others, that said 1 year was long enough? Are we now saying 1 year is not long

enough? It's very hard to do controlled trials. You see there's missing data.

DR. WHITE: I think that it's reasonable from my viewpoint to stick with the 1 year. That's what we set up. That's what the sponsor was working under. We don't have hard data right now to tell us otherwise. So, I would feel a little uncomfortable right now right here changing that definition that was set up in the absence of data that I know of that says it's wrong.

DR. SIEGEL: Right.

DR. SIMON: David, then Mark.

DR. WOFSY: As I understand it, the labeling wording is a topic for a different time and a different group, and it seems to me that what we're talking about now -- and the word has been used before and I'd like to resurrect it -- is "reduce." We have a data set here that claims that there has been a reduction in radiographic progression of disease. That reduction may be 100 percent, which some would call prevention, and it may be less than 100 percent. We're not being asked to judge at this moment what the percent is. We're being asked does the data set support the claim that there's been a reduction. I don't see what's the problem with wording it that way.

DR. SIEGEL: I'd propose reduction or has a favorable effect on progression or reduction of progression

as votes that would be informative to us, leaving open the door for labeling.

DR. SIMON: Mark and then Dr. Katona.

DR. SCHWEITZER: To me I think the phrase really is prevents the progression of structural damage. Of course, they've shown in that group that the progression was arrested. So, it really prevents the progression.

DR. SIEGEL: Or reduces the progression.

DR. SIMON: Dr. Katona?

DR. KATONA: I just would like to second what was just said except with the change specifying radiographic damage. So, I think for the sponsor it is important that "prevents" stay in. They showed data. They prevented the progression of the radiologic damage. I think to me that's what the data showed and that's what I feel very comfortable with. Any which other way we phrase it, it's going to be very difficult.

DR. SIMON: Dr. Katona, I have a problem with the word "prevent" because I have no idea what prevention means in the context of this technology. I think we're looking at numbers that are just arbitrary and constructed. I'm very concerned about what the implication of prevent means. I would have a very difficult time even voting on that particular question, regardless of whether the sponsor is unhappy or not about this particular discussion. I

think in fact we have to remove ourselves from what the sponsor wants in this context to actually reflect on what we have seen and what the numbers actually mean in consensus of our profession. And I am not aware that there's consensus at all that we know that these numbers mean stop, and I think that's what prevent implies. I don't know that we know that it stops.

DR. WHITE: I would disagree with you a bit on that one, Lee. I think that I could feel comfortable voting for a prevention claim if I had the specification that it was radiographic changes rather than structural damage. That gives me a bit more comfort in what I'm saying, a little more restriction. We've seen the statistical analysis of the data. Everybody will make their own judgment of the statistical analysis of the data, and whether you want to use the minimal determinable difference, whatever that stands for, the SDD, or the means or the medians, we all have to make our judgment of that.

But I think I could personally feel comfortable using radiographic damage, using the term prevent in terms of it meaning a statistically significant -- and what I also will throw in perhaps might imply in the future clinically meaningful difference -- if I only have to do it for a year. That's all we're talking about, a year, based on the guidelines. That's been given to us. We don't need

to change that.

DR. SIEGEL: The other thing that's given, if you take it as a given in the guideline, is that there are claims for structural damage that are based on x-ray findings, not claims for x-ray. That would be a claim that isn't mentioned in the guidance and hasn't been used before for radiographic changes. It might be wise, and longer studies might be wise, and dropping prevention might be wise. I don't want to necessarily take anything off the table. I think we're all a lot wiser than we were two or three years ago.

DR. SIMON: How about two or three minutes ago? Bill?

DR. SCHWIETERMAN: Well, perhaps I can't add too much except that the guidance document actually is somewhat inconsistent in the definition of prevention of structural damage in that you can either show a slowing of x-ray progression, which many people would not think would connote prevention, or prevention of a maintenance-free state, which is closer to I think some of the sentiments that were stated here. To rely simply on the guidance document and what's stated there, therefore is problematic, which is why I was trying to be a little bit circumspect about that, and therefore not vote on the wording of a claim.

Nevertheless, we can take these into consideration, the thoughts that people have here, when we write the label. We have some latitude about the data that we put into the clinical trials section and the indications section and so forth about these outcome measures.

DR. SIMON: Carl, did you have a comment?

DR. WINALSKI: Two things. One is as far as what they've shown, I could say either prevention as measurable by radiography or reduction of structural damage.

But one semantic thing is there is no radiological damage being done here. It's radiologically measured damage.

DR. SIMON: Thank you for that very appropriate term.

(Laughter.)

DR. SIMON: Bill, Jay, you've heard a significant discussion and some, I think, very strong opinions one way and the other. I'm not sure that you'll get any further benefit in taking a vote on this particular question at this particular time.

DR. SIEGEL: Let me say -- we'll have more discussion of this aspect of the issue later -- that I'm intrigued by the fact that while some people like the term "prevention" and some don't like the term "prevention,"

we've heard three quite different reasons among those who don't like the word "prevention," some because there might be progression after the 1 year on study, some because there might have been some people who progressed on study and it wasn't 100 percent effective, and it might imply that in some whether there might be subtle progression not measured in the Sharp score. So, it's interesting, just from a semantic point of view, how many different ways people can view the same word and its implications.

That said, I would say this. What we were looking for in this question was not the right wording for the label, but I was trying to reword it to get a consensus as to whether this was, from a regulatory standard, adequate data to support a labeling regarding this indication.

Unless I'm mistaken, I'm hearing almost everybody operate under the assumption that it is and quibble more about what the label should say. If in fact there is a general consensus that the answer is yes, there ought to be some labeling about this, but we can't agree what it exactly should be, then that is a significant part of the advice I need. Then we can move on from there to have further discussions that might help us in the determination of what it should be.

DR. SIMON: I think that clearly from the

discussion I would say everybody feels that the evidence is exactly as you stated it. The question is how do you describe it, and that's a very different question.

DR. FIRESTEIN: We could vote on Dr. Wofsy's rather benign way of putting it if we wanted a formal vote, and that is just that there is significant -- how did you say it -- reduction. And that's the question that you're really asking for right now, and I think that's a fair vote.

DR. SIMON: Would the committee agree to vote on that question, and would you feel comfortable with that then? Okay.

So, then restating that with trepidation, the evidence that we have seen this morning would suggest that infliximab --

DR. SIEGEL: Let me state it because a vote that evidence suggests something doesn't really help us from a regulatory point of view.

Do the data support a claim that -- no, because then if it's a claim, then that would --

DR. SIMON: The evidence demonstrates that there was a reduction in radiographic-measured structural damage with infliximab and methotrexate as opposed to methotrexate and placebo treated patients. Can you live with that, Jay?

DR. SIEGEL: Given the consensus, I think I can live without a vote or with a vote on that, yes.

DR. SIMON: Well, I think the committee would like to take a vote. It seemed to me that's what they'd like to do. So, that's the statement. Is that acceptable to you?

Okay, so now, Dr. Sherrer.

DR. SHERRER: Yes.

DR. SIMON: Dr. Katona?

DR. KATONA: Yes.

DR. ELASHOFF: Yes.

DR. PUCINO: Yes.

DR. WHITE: Yes.

DR. SIMON: Yes.

DR. FIRESTEIN: Yes.

MS. MALONE: Yes.

DR. SIMON: Thank you, committee.

(Laughter.)

MS. MALONE: I have a question. The patient is looking for the ability to function normally, and so structural damage -- they're not aware all the time of what's happening radiologically unless they feel the effects of it. Can you tell me why you don't have the HAQ scores for after 102 weeks? None of that is listed for physical function.

DR. SIMON: You mean the HAQ scores at 102 weeks?

MS. MALONE: At 102 weeks.

DR. SIMON: Or the ACR20's or the function outcome.

MS. MALONE: Yes. You don't have any of that information.

DR. HARRIMAN: Yes. Actually Dr. St. Clair in his presentation showed the HAQ data through 102 weeks. Again, as I indicated in my presentation, some of this data were just recently obtained because the 2-year endpoint and the trial were completed just recently, and everything hasn't been fully analyzed. But we have the HAQ data through 102 weeks, and that was shown in Dr. St. Clair's presentation. We'll show it again.

DR. SIMON: We can certainly pull up the slide, but you have to admit that it was not the complete data. It was just a smattering, a taste of that 102-week functional outcome data. Is that correct?

DR. HARRIMAN: What the data were that were presented was the change in HAQ through 102 weeks. It was the primary endpoint in the study for 2 years, and admittedly, all the data have not been fully analyzed. But the primary endpoint for 102 weeks was the HAQ.

DR. SIMON: So, unfortunately, it's very

difficult to interpret in that context, if you're asking for a full assessment of outcome at 102 weeks.

Did that answer your question?

MS. MALONE: Yes.

DR. SIMON: I'd like to move on to the actual third part which has implications regarding dose. We've heard multiple times people comment on what the dose should be and we've seen data at higher dosages and have seen different responses both from an effectiveness point of view, as well as a point of view of safety. The agency would like to have some discussion about how we feel about the dosages that we saw. Is this the 3 milligrams per kilogram, 2, 4, 6, then every 8 weeks thereafter the right dose? Or should we be looking at a different dose that implies more efficacy?

DR. PUCINO: Yes. I have some question. Since there's an association with the concentrations in the plasma and the clinical effects, will there be commercially available assays? And has anyone looked at correlations with plasma concentrations and adverse effects, not just a dose response, and a test for trend with those different items?

DR. HARRIMAN: We're not aware of any commercial assay for assessing infliximab concentrations. As I showed in my presentation, there is a correlation

between the trough concentrations of infliximab with both clinical parameters as well as laboratory parameters, such as CRP, which allows one to look at the effects both clinically and laboratory measurements with regard to the trough concentrations. But there are not any commercial tests available for infliximab concentrations.

DR. ST. CLAIR: Just one other point, though. It's important to realize that patients that had undetectable trough levels at 30 weeks still had clinical responses. So, there's not a complete correlation of trough levels and response. So, you still have the opportunity to get good responses with a lower dose in some patients.

DR. PUCINO: And that would be the concern. Right now, as it looks, you're either saying to double or triple the dose, and if you have someone who's already 5 to 10 mics per ml is that going to add more therapeutic benefit?

In terms of a safety perspective, having assays available at least for individualized patients could be beneficial.

DR. ST. CLAIR: I think most rheumatologists would not choose to use the assay but rather treat based on clinical symptoms. That's, in fact, how most rheumatologists decide how to treat their patients.

So, to me it makes more sense to start patients out at the lower dose and capture what responses are going to happen there. Then if the patient's response wanes after maybe 14 to 30 weeks -- if you remember the PK slide that Dr. Harriman showed, you could see the trough levels starting to come down between 14 and 30 weeks. If that patient happens to show a waning of response then, then I think that's the time to increase the dose. For me, I would just increase it by a vial, 100 milligrams.

DR. WHITE: Could I just ask Bill? That's one approach, but for example, if we look at another drug, cyclophosphamide, we actually don't measure drug levels, but we measure something associated with it. And when we don't have a benefit, if the neutrophil count is 2,000, we're not likely to just add a little more because -- maybe we won't have it here, we know that adding a little more, in the setting of not working, might give us just really an unacceptable risk-benefit ratio.

DR. ST. CLAIR: That's an important comment. Then you have to look back to the safety database at the different dosage levels. Even at the top dose, it's not clear to me that there's any increase in toxicity at the highest dose compared to the lowest dose. But it's just medically prudent to use the lowest dose that would be effective in that particular patient.

DR. SIMON: Dr. Elashoff?

DR. ELASHOFF: With regard to the safety issues, I think all the adverse event data should be re-analyzed looking for a dose-response trend across the five groups to see if there is one significant and not using the far more conservative approach of saying there isn't anything if there isn't an overall effect. So, the analyses have already been done. That could be added to that.

In addition, in fact, since you have trough levels on everybody, you could use logistic regression kinds of approaches to see if the actual levels are predictive of adverse events. So, the data that are here could be examined much more carefully to address these questions even before one thinks of additional trials.

DR. SIMON: Dr. Katona?

DR. KATONA: I would like to ask the sponsor whether they have any pediatric data on pharmacokinetics, whether the same dosing regimen applies for children, as well as whether they have seen the same relationship between serum levels and efficacy.

DR. HARRIMAN: Yes. A couple of things.

First of all, we have a study that is planned, following discussions and a commitment to the FDA to perform a study in patients with juvenile rheumatoid

arthritis. So, that is planned.

Secondly, we have performed a study in pediatric patients with Crohn's disease, a pharmacokinetic study, looking at different doses of Remicade from 1 to 10 milligrams per kilogram and have found that the pharmacokinetics in those pediatric patients was similar to what was observed in adult Crohn's patients.

So, we have both a study that will be done in the future, as well as the study that has been done in the Crohn's pediatric patients.

DR. KATONA: In the Crohn's patients, was that the repeated dosage schedule or just the one dose?

DR. HARRIMAN: In that study that I indicated, it was a single dose pharmacokinetic study looking at full pharmacokinetics, but it was not multiple dosing.

DR. SIMON: Any other comments about this? David.

DR. WOFSY: I guess I do with some trepidation.

I think there are some reasons for safety concerns. Nobody claims that this agent, any more than any other agent, is entirely safe, and in all likelihood, higher doses will come at a higher price. I think we have some evidence of that here. We have evidence in the form of statistically significant, more frequent minor infections, upper respiratory infections and sinusitis. To

me, it would be hard to make the case that this will increase the risk of minor infections but not more severe infections. I think the reason we don't see it statistically significantly in more severe infections might at this point be a reflection of the smaller numbers. In those areas where we have bigger numbers, that is, minor infections, we are beginning to see it. So, I think there are some concerns about infection, even excluding the infrequent serious opportunistic infections that have been seen.

We heard this morning, somewhat as a surprise to me, that at least looking at this in one way, these data can be analyzed to show a statistically significant association with malignancy.

Now, this is early to make those comments, and that's why I speak with trepidation about this because I think we're really at a very early stage of understanding. If what I just said is true, will it be supported as this becomes used more widely? And to what extent one should be concerned about it. Barbara has made the point that we use a lot of drugs like cyclophosphamide which are known to be associated with strong risks of malignancy and infection, much stronger apparent risks, and yet in some individuals we make the decision that the benefit for that person is worth the risk.

It seems to me at this point that's what we're dealing with here. We have a dose that has been approved, is reasonably safe, and is reasonably effective. We now have some evidence that suggests occasionally there may be somebody in whom the potential benefit of going up is worth the additional risk. I think that is sort of what we're dealing with now.

I think prudence at this stage of development would certainly support the kind of approach that Dr. St. Clair mentioned, sort of routinely starting at the low dose and then considering whether there are special circumstances in which the severity of the disease and the potential benefit warrant these possible significant risks.

So, I don't know how that translates, but my own view looking at this is, yes, higher doses look like they're more often effective and maybe more potently effective, but I think the whole picture at this point would caution us to stay away from them in the majority of instances.

DR. SIMON: Before we go on, I just want to point out that all of that I agree with as well. The dilemma, of course, is that the question inherent in this discussion is should the higher dosages be labeled. Given the cost of the product, without having some kind of labeling, it's sometimes very hard to get the managed care

organizations to then allow you to use such dosages, even when they're appropriate, given the risk-benefit relationship. So, unfortunately, it does take us into a realm that we usually don't discuss but, unfortunately, will need to because of that reason.

DR. WOFSY: Can I respond to that, Lee, on the same point I raised?

DR. SIMON: Yes.

DR. WOFSY: I don't know a great deal of what precisely goes into labeling, but it would seem to me labeling could take into account the things I mentioned that says "in usual cases" or some such thing.

DR. WEISS: I just wanted to say -- and this is somewhat, I guess, inherent when we get down to part (d) of that question -- is that thus far the data that we have do not address those patients specifically that start out at 3 and then are increased subsequently, should they not achieve the response. So, that was part of why we asked question (d), which will come up a little bit later, that specific scenario that people seem to speak about and seem to have some sense that it might be beneficial.

DR. WHITE: I would like to speak in favor of what I think David said. I think given the data that we've seen, that it does look like, by a variety of different measures, that higher doses may have a higher likelihood of

being associated with benefit. I think knowing that, we ought to give that to the practicing physician and the patients. They ought to be cautioned, but I don't think that we should not use those data that we have to benefit the patients.

DR. SIMON: Gary?

DR. FIRESTEIN: On the other hand, unlike with cyclophosphamide and several other drugs but like cyclosporine, for instance, we can use blood levels in order to help us make decisions about dosing. Specifically, although there's not a great correlation certainly at the higher levels between response and blood levels, there clearly is a group of patients that have nondetectable trough levels and don't have a significant response to the agent.

So, I would propose that the most rational way to do it -- and rationality doesn't always come into play in clinical practice -- is that nonresponders have a trough level check, and if the trough level is low, then that provides a rationale for going to a higher level. If the trough levels are not low, then there's not much point in going to higher levels. And that those types of assays be available to the practitioners.

DR. SIMON: Dr. Katona?

DR. KATONA: I would like to come to the

question from a different point of view. Basically we could look at this preparation as an immunoglobulin and look at half-lives and so on. If one looks at the graphs, there is a very different serum level if you give the drug every 4 weeks versus every 8 weeks. Every 4 weeks gives you a very nice and even distribution, and the 8 weeks gives you a high peak and then it comes back. Basically the 10 milligrams every 8 weeks eventually will level out to the one that you were at 3 milligrams every 4 weeks. The question is in the labeling.

To me, as a clinician knowing this background, if the 3 milligrams every 8 weeks doesn't work, what would make the most sense to go to 6 weeks and then 4 weeks versus getting these high levels because I would be wondering that if we dose high serum levels -- I wouldn't worry about the low ones, but I would worry about the high ones, whether those are the times when I'm inducing the malignancy, those are the times when I am interfering with all the defense mechanisms and have the infections and so on. So, I think that would be very important to take into consideration.

The other thing, since the trial was done at 8 and at 4 weeks, I wonder whether in the label you give freedom to the physicians that they could use some other timing in between. It doesn't necessarily have to be

either 8 or 4. It could be anything in between.

DR. SIMON: Dr. Katona, would you then make one step further and say a few clinical studies that you'd like to see proving your proposition?

DR. KATONA: I think that that's actually a very, very good idea to have dosages between 3 and 10 -- we might not have to go up to 10 -- as well as looking at timings between 4 and 8 weeks.

DR. SIMON: Would these be safety or efficacy or both?

DR. KATONA: Long-term safety. I think that's something that we could collect the data. But definitely efficacy. That would be number one, and long-term safety.

DR. SIMON: Yes?

DR. SIEGEL: Regarding the last two comments, I would like to note that while there appears to be -- and there's data to suggest it -- a correlation between dose regimen and efficacy and perhaps some suggestion regarding safety, that speculation about the relationship of trough or peak levels is just that, a speculation. It may well be attractive, but to say we know it's the ones who have a lowest trough who would benefit from a higher dose, well, we don't know whether they would or whether people with higher troughs would benefit from a higher dose. I'm not suggesting that levels wouldn't be important, simply that

we don't have that information.

But as to the last comment as to whether it makes more sense to go to 3 q 4 rather than 10 q 8, I would note that while both of those had the same troughs, the ACR rates were substantially higher on 10 q 8 then 3 q 4. 3 q 8 was 42 percent. This is the ACR20. It went up to 48 percent on 3 q 4, but it went up to 59 percent on 10 either q 4 or q 8. So, it may be the peak that's more relevant. I'm not saying we know that. I'm saying we don't. We just have a suggestion that giving more of this, whether it's more often or at a higher dose, does seem to improve the response rates.

DR. SIMON: So, it suggests that the sponsor, if they're interested in having other dosages be approved by managed care organizations, should come in to you with suggestions for other studies that would answer those particularly questions to allow you to label it more fairly, so to speak, based on responsiveness and accessibility.

Dr. St. Clair?

DR. ST. CLAIR: Let me try to shed just a little bit more light on this. I think that you can assess whether the patient is going to be a responder or not while their trough levels are relatively high. I want to take you back in your mind to the figure that Dr. Harriman

showed where the patients received infusions at week 0, 2, and 6. Those trough levels went up. It was only when they went into the maintenance phase, every 8 weeks -- we're talking about 3 q 8 -- where they started to come down. Recall too that patients respond rather quickly to this drug.

So, when you're taking care of the patient and you start the patient and give them the induction regimen, what in effect really happens is that you do see an initial response in the patient, if the patient is going to respond, but it's later, between that 14 to 30 weeks, where the response might wane. That's where you might want to adjust the dose upward. It may be that in that particular patient their serum levels are dropping down.

I think the 1 microgram per ml, using that as a strict criteria for clinical efficacy, is taking the data way beyond what we know. There is another figure that has been shown too, but I'll just quote the data. Patients with trough levels of less than 0.1 at week 54, there were still 13 out of 28 ACR20 responders. So, I think we're getting too tight on these antibody levels.

But I still think that it's important for the clinician to have the option of increasing the dose up in certain patients, as Dr. Wofsy suggested. I think the safety issue is a little bit open at this point.

DR. SIMON: Furthermore, I think that I'd like to point out that we have to remember -- and that's the fourth question here -- that we did have a discrepant response rate, meaning where we saw patients who had x-ray evidence of benefit -- maybe that would have been the way to ask that question -- but, nonetheless, they had no clinical response or minimal clinical response. So, Bill, there are still people we won't be able to measure clinical response in acutely and yet over time have a structural response which may be important, and that may be only attainable by doing a blood level perhaps. I don't yet know until we do the trials.

DR. SIEGEL: I was just going to add to that last comment, though, that yes, most patients respond early. Yes, one of the issues is, as you move to that 8-week dosing, so that at week 14 they've been 8 weeks without a dose, you may see loss of a response in a patient who had responded. But there are also patients who don't respond at first who respond later, and there are more of those patients in higher dose than in lower dose.

So, there are both questions of potentially of using higher doses and dealing with people who have responded and lost a response, but also -- I don't have the numbers, and I don't think we've seen them presented here -- of people who haven't responded, potentially looking at

higher doses. One of our questions was should there be a study to look at whether higher doses are useful in people who have not had a response at a lower dose.

DR. SIMON: Carl?

DR. WINALSKI: I was just wondering how much does the addition of methotrexate add to the noise here in trying to figure out the safety and how important is methotrexate for the efficacy? It seems to me that if you have a lot of baseline noise, it's going to take a lot more patients and a lot longer to sort out the safety of just one drug versus the two added together.

DR. SIMON: Perhaps the FDA could answer the question as to why you've labeled this drug to be used with methotrexate.

DR. SIEGEL: It's the only way it has been studied. Go ahead.

DR. SCHWIETERMAN: It's the only way it has been studied.

(Laughter.)

DR. SIEGEL: We would note that there are a lot safety and efficacy issues that one could theoretically hypothesize as to single use. Immunogenicity may be different alone. Interactions. It could be better, it could be worse. We just don't have any information.

DR. SIMON: If I'm not mistaken, there is a

study pending that's single use alone, right, that you described? Right.

Barbara?

DR. MATTHEWS: I would just like to point out that's how it's labeled for rheumatoid arthritis, but infliximab is also licensed for patients with Crohn's disease. In those cases, there's no labeling saying that it has to be given in conjunction with methotrexate.

DR. HARRIMAN: I just wanted to let the committee know about a study that we're doing. It's called the ACCENT study which is in Crohn's patients, a fairly large study, 579 patients. In that study, we are looking at dose titration in patients who do not respond at a lower dose, crossing over to a higher dose. So, there will be some evidence learned from that study with regard to dose titration.

DR. SIMON: Have we achieved the goals of your number 3 series of questions? I think we've addressed each of those issues.

DR. SCHWIETERMAN: Have you talked about (c), about initially starting at higher doses? I heard Dr. Wofsy and I think he made some very good points, but is that the consensus of the committee that it's probably not worth it at this time starting initiation of therapy at those doses, rather to concentrate on treatment failures at

the lower doses and/or differing regimens in those patients?

DR. WHITE: It just depends on your point of view. It might be that if you started with a higher induction dose, then maybe you would get better responses.

DR. SIMON: That may be true. I think, again, it raises the question of how it's being studied and what is being studied at. It's interesting to note that in another product, many of us have complained that we don't have dose-response curves to understand how that product should be used. In this context, at least we have two separate dosages given at different times that give us some insight into the various different relationships of dose. So, at least we have that. But I would agree with Barbara that perhaps further studies in that particular realm would be useful.

Frank?

DR. PUCINO: And, if in fact, 75 percent of people within 2 years of diagnosis will have irreversible changes, it would be nice to have these additional studies.

DR. SIEGEL: In response to your question, Lee, as to have you given us the information that we need, let me try this. Let me state what I understand to be a consensus of this committee, although perhaps not unanimous. If I understand it and if we all agree, then

we'll know that we understand the consensus, which is that without again getting too highly specific about the labeling, it sounds to me -- and this is perhaps the most controversial part -- that most of the discussion has suggested that 3 q 8 ought to be a starting a dose and that labeling ought to allow for the fact that dosage might be made more frequently or higher, within the ranges studied. That would, therefore, lead us to present the data for all the doses studied, if we did that, and then put in the dosing section a range, perhaps not being too highly directive as to the best mechanism for titration of the dose.

Is that more or less what people are thinking is the right thing to do? I'm seeing a lot of head nods.

DR. SIMON: And tacking on the fact that it's possible that at higher dosages, there may be more problems with safety.

DR. SIEGEL: In the safety section, we would indicate those concerns about the infection and malignancy and the theoretical concerns.

DR. SIMON: Would everybody on this committee kind of feel comfortable with that as it was stated? Any dissenters?

(No response.)

DR. SIMON: See, consensus. It's almost

unanimous.

I think we've addressed each of these questions.

DR. SIEGEL: The other part is additional studies, and I'm generally hearing that everybody thinks it would be nice to know more. I haven't heard anybody say that it's compelling that a particular study be done.

DR. SIMON: Then we'll move on to the number 4 question which is going to have some reflection on this afternoon. For those of us who have been on this panel for some time, we've had any number of different discussions about separation of structure, function, and signs and symptoms as outcomes.

In this data set, there are patients who did better from an x-ray point of view than they did from a signs and symptoms point of view, depending on how one weighs that. I find that very interesting.

If that's the case and if everybody finds it interesting, the question at hand is, is there any basis to support a claim or -- I like this term -- "belief" -- it brings us into religion and teleology which I think is appropriate here -- that patients treated with Remicade who do not experience improvement in their ACR20 but show improvements on radiographic measurement findings, e.g., no x-ray progression, have benefitted from therapy?

So, can we say somebody who's done better by x-ray measurement, if that gets done in a clinical sense, has benefitted from therapeutic intervention where they don't feel better two days later or three days later?

Dr. Elashoff.

DR. ELASHOFF: I just wanted to comment that ACR20 has a variety of arbitrary cut points and that if you use one of the other ones, you'll get a different answer here. So, making one thing yes/no is always going to make it harder to agree with something else. I would just say from a statistical point of view, this kind of question is difficult to deal with.

DR. SIMON: We appreciate that.

David?

DR. WOFSY: This is a very important question, but it does seem to me that this is a question that has to be answered in long-term studies and not easy studies to do. I would love to see the sponsor take them on, but I don't know how to address this in any other way. You need to have willing patients who have not had a good clinical response get randomized to continue this for decades maybe.

As Fred Wolfe has said, it may take the 10 or 20 years to actually see that changes in the x-rays predict hip replacements 15 years down the line or some such thing. I think that's possible. I think it would be a very

important observation. I don't think we have any evidence to allow us to guess that that would be the case. If we did guess that that would be the case, what we would be saying is everybody with rheumatoid arthritis should be on this as sort of background therapy.

So, I think it's a very important question, and I hope some courageous patients are willing to participate in that kind of a study, to actually subject themselves to the risks of this agent without substantial, obvious, short-term clinical benefit. But I think in the absence of that kind of information, we don't have anything here to suggest that treatment for that indication would be a wise thing.

DR. SIMON: Any other comments about this issue?

It has actually major ramifications in osteoarthritis in particular, not that it's not important here. But it really has major ramifications in assessing outcome in osteoarthritis.

Carl?

DR. WINALSKI: I guess perhaps wanting to remain relevant in medicine, as a radiologist I'd like to believe that what we do detect is predicting long term what would happen. To make another perhaps flawed analogy, if you were to say, well, I don't have any proof that treating

blood pressure will decrease the risk of stroke, that's also a very long-term thing, which has been looked at with cross-sectional studies. I think if you took some of the data which has been mentioned and said, do people who have bad radiologically scored disease feel worse than those that don't, I think that that's some good cross-sectional data that the radiographic progression is perhaps a reason to be treating patients.

DR. SIMON: This may be poor solace as an observation to the sponsor, but this whole discussion and the importance of this discussion is predicated on the fact that you have such robust data in the context of such terrible technological outcomes, relatively speaking. They're the best we have. If we understood more about the technological outcomes and we had consensus about that, it may be as easy as it is in osteoporosis, which it isn't. It's because your data is so good that has caused us to have this kind of discussion.

I'm sorry. Dr. Johnson over here has his hand up. One more comment.

DR. JOHNSON: I couldn't resist this one. Just in light of the surrogate question and the blood pressure question and so on, blood pressure as a matter of fact, had a monstrous epidemiology and it still does, and it had some clear-cut interventional trials, the first of which I

happened to just review. Because, lo and behold, they used the worst case scenario. They randomized people to placebo versus treatment if your diastolic was between 115 and 130. A pretty impressive maneuver. This was the first VA study. There were 27 bad outcomes out of 143 patients, 2 versus 25. If you took the 10 lost to follow-up patients and put them in the bad outcome category also or switched the outcomes, like Desiree did, it still won by .001.

So, that's the two parts of the surrogate question. One is the epidemiology if and when we get it. We have some of it. But the second part is does your intervention which affects your surrogate translate into a clinical outcome, which has been proven in blood pressure, at least with some subsets of blood pressure medications.

DR. WINALSKI: I had a feeling it was a flawed analogy.

DR. SIMON: Dr. Katona?

DR. KATONA: This question just reminded me that question number 2(c) we did not answer half of the question. I think it's somewhat related and that part of the question was that to what degree the benefits what was seen from the studies which were done on patients who had longstanding moderate or severe rheumatoid arthritis could be extrapolated to patients with early onset, less severe, and DMARD-responsive disease. I think this "prevent" word

is very important because I think as clinicians we're going to be always confronted with this. So, I don't know whether the chairman would like to discuss it now or in the afternoon. I just wouldn't like to forget about it.

DR. SIMON: I'm very happy to discuss it now for one second. In that, we need to remember that it's not that these patients had long-term disease; they had nonresponsive disease to methotrexate. We have other studies that we're going to hear about this afternoon. So, I think actually this part of the discussion would do better this afternoon, if the FDA would agree.

So, have we achieved the point in this meeting where we have answered the questions you've come to the table with? And are there any other questions that you might have for the committee regarding the infliximab presentation?

DR. SCHWIETERMAN: From my standpoint, I think we have answered all the questions. I don't see any other heads. So, thank you very much.

DR. SIMON: At this time then we are going to break for lunch.

I'd to thank the Centocor sponsor for coming in and giving us such an excellent presentation. I'd like to thank all the speakers.

We are going to return at 2 o'clock for

continuing our afternoon discussion. Thank you very much.

(Whereupon, at 1:08 p.m., the committee was recessed, to reconvene at 2:00 p.m., this same day.)

AFTERNOON SESSION

(2:05 p.m.)

DR. SIMON: I'd like to welcome everybody back to our afternoon discussion. It's entitled, according to official ruledom, Discussion and Consideration of Proposed Radiographic Outcome Measures for Investigational Agents for the Treatment of Rheumatoid Arthritis.

I think that it's very critical to recognize that what we're going to be doing here this afternoon is actually initiating the entire discussion, perhaps again to some of us, about the issue of radiographic outcomes. This has important implications for the guidance document in rheumatoid arthritis. As a result, the discussion I hope will be lively. Certainly with any evidence from this morning, it should be more than lively. I would like the committee to feel comfortable in discussing any issue related to this.

We are going to have Kathleen Reedy present a statement, following which we are going to have several speakers. Then we have questions that are in your packet for us to discuss. Kathleen?

MS. REEDY: The conflict of interest statement for the Arthritis Advisory Committee, July 12, 2000, for general discussion.

The following announcement addresses the issue

of conflict of interest with regard to this meeting and is made a part of the record to preclude even the appearance of such at this meeting.

Based on the submitted agenda for the meeting and all financial interests reported by the committee participants, it has been determined that all interests in firms regulated by the Center for Drug Evaluation and Research present no potential for an appearance of a conflict of interest at this meeting with no exceptions. Since the issues to be discussed by the committee during this portion of the meeting will not have a unique impact on any particular firm or product, but rather may have widespread implications with respect to an entire class of products, in accordance with 18 United States Code, section 208(b), each participant has been granted a waiver which permits them to participate in today's discussions.

A copy of the waiver statements may be obtained by submitting a written request to the agency's Freedom of Information Office, room 12A-30 of the Parklawn Building.

In the event that the discussions involve any other products or firms not on the agenda for which an FDA participant has a financial interest, the participants are aware of the need to exclude themselves from such involvement, and their exclusion will be noted for the record.

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

DR. SIMON: Thank you.

I'd just like to establish some firm ground rules for the discussion. We will certainly ask for expertise to be brought in from people within the room, as well as from the committee, but the discussion is predominantly for the committee to discuss with the FDA about the questions.

Dr. Schwieterman, the next speaker is not yet here. There she is. I couldn't see her. Thank you.

So, I'd like to introduce Dr. Barbara Matthews from the FDA to initiate the discussion.

DR. MATTHEWS: Well, in some respects I feel that my presentation will now be kind of anticlimactic, given this morning's discussion. However, I think at the same time it will summarize and hopefully congeal the points that were discussed this morning and touched upon and hopefully lead into the afternoon presentation.

But I was asked to discuss the guidance document, which is the guidance to industry, clinical development of programs for drugs, devices, and biological products for the treatment of rheumatoid arthritis,

particularly as it applies to the recent increased development of new therapeutics and some of the questions that have been raised to the agency as these products have been undergoing development.

As you know, the document resulted from a collective effort on the part of academics, industry, and the regulatory personnel. It was published not that long ago really, even though it was the last millennium. It was published in February of 1999. I would say that it reflects the standard of patient care and our scientific knowledge as it stood in the mid- to late 1990s. However, as you know, medicine is a very dynamic science and consequently we need to continue to reassess the ability of the guidance to meet the needs of good therapeutic development.

What I'd like to do in this brief presentation is provide some background summary of the claims section of the document and then present points for present and future consideration.

So, what are the claims or indications that are discussed within this guidance document?

Well, first, there's the claim for the reduction in the signs and symptoms of rheumatoid arthritis, and for this claim, the guidance document discusses the need for -- well, 6-month trials are

encouraged, but within known pharmacological classes, trials as short as 3 months may be possible. However, the need for long-term treatment or long-term trials, because of the chronic nature of rheumatoid arthritis, is a consistent theme throughout the document as it discusses the issue of claims.

This section of the document also discusses the landmark analysis versus analysis of the patient's response over time, and it also provides examples of acceptable measures, namely the ACR20 or other well-accepted indicators of signs and symptoms.

A major clinical response is one that the patient achieves an ACR70 for 6 continuous months. A complete clinical response is one where the patient's response is greater than an ACR70 for 6 continuous months, and a remission requires a response both by ACR criteria and also radiographic arrest for 6 continuous months off of therapy. None of the products that we've seen recently in the last 3 to 5 years have achieved either of the last three claims.

The prevention of disability was really intended to encourage long-term trials, and the guidance document gives some guidance on the duration that they were thinking of, namely 2 to 5 years.

Validated measures to be used in such trials