moving ahead unless there is some consensus about something that's measurable.

DR. ABRAMSON: Dr. Hochberg?

DR. HOCHBERG: I think that is why I raised my point initially, is if we say that structure is a surrogate for a clinically-important outcome, then you can consider structure in the context of the joint as the organ, and one of the features of that is joint-space narrowing.

I mean, another feature could be the size of the largest osteophyte, et cetera, and it might differ based on different joints. I think it's different from the hip as well as the knee. Those need to be considered differently, but you know, I think joint space can be a surrogate for a clinically-important outcome.

DR. ABRAMSON: Okay. Question Number 2 with respect to pain and function. "If in the course of these studies, no worsening is required, how would no worsening be defined?"

Dr. Dougados?

DR. DOUGADOS: Just to remind that, within the Osteoarthritis Research Society, with Roy Altman, we are chairing a task force in which the objective is to propose response criteria for symptoms. That is, to have a composite index combining, as an example, responder would be considered if he or she fulfill the -- such as the ACR

criteria for rheumatoid arthritis internal responder, and we have a meeting in June, and we have the description whether or not we can take this opportunity not only to define responder criteria that is an improvement on symptoms but also the other posit to have a responder criteria, worsening criteria, and that at this time, there is no plan to conduct such a study within our standing committee, but otherwise personally I have no answer.

DR. ABRAMSON: All right. Let me just ask for a clarification here. Most of the studies that are ongoing, there are measurements of pain and function, the WOMAC, the VAS, and various other pain and functional studies. I guess for clarification, I'm assuming that one would continue to use those kinds of measurements going forward, and is the question put to us, should there be other things that we're looking at, in addition to what's built into the current OA-type studies? Or are those sufficient to just continue to follow and look for what the delta of worsening or improvement might be?

DR. JOHNSON: Well, I think this question actually was related to a previous one, in that if there's a sentiment that a substantial, or even maybe a small change, in joint-space narrowing in a certain subset of patients succeeds, and it's argued that the only test should be at approval, and that there should be no clinical

worsening because it's anticipated to take two more years to show a clinical benefit, I think it's a statistical test question. How do you define no clinical worsening? I don't think we mean no statistically significant worsening, but I also don't think that most people mean by that that it has to be statistically equivalent by some ignorable predefined delta.

DR. ABRAMSON: But the instrument that one would use to make that assessment would be the same instruments, I'm assuming, that are built into --

DR. JOHNSON: Right. Yes. The instrument itself is an analytic challenge because there's two or three components to it. I mean, the concepts are pain and function, but the perception is that you also have to address the patient global because I think my sense from the OMERACT proceedings was that there was just a lot of reservation and concern that if that were jettisoned, something would be missed, and as I mentioned in my slide, in some sense, you have to address what you know are important covariates.

So your assessment of symptoms becomes complex, too, but I don't think there was the suggestion that any new or additional dimensions be added.

DR. ABRAMSON: Dr. Witter?

DR. WITTER: I had tried to take a stab at that

in my presentation, and I don't mean to kind of lead any conversations, but perhaps it refers to clinical benefit as defined by the use of analgesics or, you know, NSAIDs, things like that.

I mean, can we define no worsening on clinical grounds in that regard, and if we would do something like that, then is it, as we've been alluding, as Ken was talking about, for example, is it equivalent to something or, you know, difference from? So I think that's part of what we want to hear.

DR. ABRAMSON: Yes, David Yocum?

DR. YOCUM: I guess a question here for me hearing Paul's comments earlier on comorbidity, which I think we all realize the psychosocial issues in this disease process, especially this group of patients, and Ken's comment about the earlier group of people with no change in WOMACs earlier on. This is either going to have to be a very complex functional analysis looking early on or, I guess, relatively gross and simplistic.

It doesn't sound like an easy process. Is that missing the point?

DR. ABRAMSON: Does someone want to respond?

DR. BRANDT: Steve?

DR. ABRAMSON: Yes, Dr. Brandt?

DR. BRANDT: Yes. I didn't mean to imply that

those were earlier disease. They're community dwelling, and they may just have better coping skills. These are patients who, for example, with WOMAC pain scores similar to what we might see in the clinic, in patients who are on NSAIDs, are not taking NSAIDs and are not going to physicians for care. They don't perceive a need to get treated for this. They may be treated for other things, and it's not that medical care isn't available. These are all Medicare-age people. They just have not seen the need to become patients with their knee pain and x-ray changes.

DR. ABRAMSON: Dr. Dieppe?

DR. DIEPPE: Yes, and that's a terribly important point, but just coming back to this question, when I read it, I wasn't sure I understood what it meant, but I presumed when I read it that what you were asking us is what's the variability, the natural variability in pain and function in the course of osteoarthritis.

Now, if that's the question, the answer is that it's huge, and there's all sorts of rhythms that have been demonstrated by Nick Bellamy. There's a daily rhythm of pain, and there's a weekly rhythm, and then there are clearly other rhythms of change over longer time periods where there's quite marked variability, and again Maxime as well as our group have some data that speaks to that, and I think the longer rhythms myself are related to these issues

of psychosocial factors which determine whether people are patients or whether they're not.

So this isn't an easy question to answer either. The other observation I'd chuck into the pond is that most people get better with osteoarthritis. Certainly clinic populations get better because they come to us when they're at the peak of a problem and just regression to the mean will get them better.

But I think everybody gets better by and large, except those few that progress to smashed-up joints and joint replacement, and I think they get better again in relation to what Ken's saying because of adaptation and coping strategies, and if you have a condition for 10 years or 20 years, where your pain relates to what you do, and the issue is function, and you're getting older, you adapt, and you cope with it, and you get better.

So you posed another impossible question, but let's just be sure we understand the complexity of the question.

DR. ABRAMSON: Okay. If I were to pose the question perhaps a little differently for the committee and the agency and simply said if one enters into a protocol, a doxycycline protocol being one, where there are going in outcome measures for pain and function with WOMAC patient and global assessment and whatnot, pain measurements, would

there be a reason not simply to continue to follow those instruments throughout the study, looking for deterioration in pain and function that might be unanticipated?

Is that not the best way to proceed in these discussions, and so I just put that out. What's the alternative to doing that that's practical or doable or preferable?

DR. YOCUM: I think that's very valuable because from our meeting yesterday, the group yesterday did not include functional analysis, and even though we're talking about a short term here, what I would hope in these studies, and the FDA would require doing these studies out there in the real world, that getting companies to follow patients long term is critical, and we often come back to it as second thought. Oh, yes, let's do this long-term analysis now that your drug's approved, and it would seem better that we incorporate early on the functional analysis of predetermined issues which I would try to include some sort of psychosocial measurements, affect scales of some sort, that are there, that can then be carried long term, so we can get long-term data, rather than giving it a second thought.

I think your comments are appropriate, Steve. I think it's important. The question is which are the best, and it should include psychosocial issues as well.

DR. ABRAMSON: Other comments? Dr. Johnson?

DR. JOHNSON: Well, we haven't included psychosocial, although a lot of these trials are ongoing. So these are pressing problems, and we have asked pretty uniformly to cover the other more standard domains.

The problem, I think, as it's posed, is an analytic problem. How do we get analytic assurance and approval for a structure that does not have symptom deterioration? I mean, it's easy to describe in words, but how do you translate that into an analysis?

DR. ABRAMSON: Is there a concern -- I guess what I'm not fully understanding that if you use these criteria that are put in place, whether it's for an NSAID or a structural modifier, for OA outcome, is there concern that you will not be able to differentiate from placebo deterioration using these instruments?

I mean, clearly they're put in place to measure improvements over placebo in the normal circumstance. Is the opposite not equally true?

DR. JOHNSON: Well, they may well not show a difference from placebo at the time that structure occurs if the presumption that clinical change is going to take much longer than structural change.

DR. ABRAMSON: So it's in Phase IV. The issue is more of a Phase IV continuation of pain and function.

DR. JOHNSON: The issue is what approval analytic reassurance can we conjure up, you know, vis a vis the absence of symptom deterioration? I think it comes around to this issue of an equivalent study, but that strikes me as too rigorous, you know, to demand that the drug be statistically equivalent.

DR. ABRAMSON: Right. So at the time of approval, if there's structural gains by whatever criteria are established, and no worsening, one could go theoretically forward, but the issue is often, as the case in front of these committee discussions is, what is the mandate for Phase IV study, and what should that consist of in terms of measuring pain and function out two and three and four years?

DR. JOHNSON: Well, I think we know how to do that. I mean, at least we've done it in short-term trials. The question is what does no worsening mean analytically? I think that's one of the challenges.

DR. ABRAMSON: Does anybody have a comment? Dr. Hochberg?

DR. HOCHBERG: Well, the question is, do you have data or access to data from trials which have been submitted as part of other NDAs in either OA or possibly in RA, if you can extrapolate, to determine what is the clinically relevant difference in some of these continuous

measures with an anchored Likert scale outcome?

For instance, in some trials in the past, we've had a patient response to treatment measure, you know, global assessment of response to treatment, where the patient could say they were worse, no change, minimally/moderately or markedly improved, and is there a bank of data which would allow you to assess what is the mean and the variability of the change in the VAS scale which is anchored to that in order to make some estimate as opposed to looking at a comparability or just a statistically significant, which might not be a clinically important change?

DR. JOHNSON: Again, you're asking that to make the Phase IV decision more rational. No?

DR. HOCHBERG: No. You could apply it to the Phase III where you've collected data on -- for instance, if you're going to use a VAS pain scale, let's say, then you want no worsening in pain. Well, a one-millimeter change in a pain scale, which might be statistically significant if you have large enough numbers for the structure study, may not be clinically important, but a 10 millimeter may be clinically important.

DR. JOHNSON: Well, but the issue is no difference compared to placebo. I mean, you've got a control.

DR. HOCHBERG: But what if you have a difference compared to placebo, but it doesn't fall within a range which is felt to be clinically important?

DR. JOHNSON: Okay. I see what you mean. The only database we have would be the non-steroidal world, and I think the decisions you would make from that database arguably might be inapplicable to structure-modifying drugs or the sponsors might argue that anyway. But I think it's a good idea.

DR. ABRAMSON: Dr. Anderson?

DR. ANDERSON: I just want to take up the point. Kent Johnson, you said that you thought the equivalent would be too stringent perhaps, and it seems to me that pain and function as measures are generally so much more sensitive than the joint-space narrowing is likely to be that having enough power to detect, to prove equivalence, shouldn't be a problem in this context, I don't think.

DR. JOHNSON: Is that assuming that you're going to -- are you making it not a difficult test by prespecifying a relatively large delta that can be ignored essentially? Is that what you're proposing?

DR. ANDERSON: Well, you have a certain delta that can be ignored, but I would think that it could be quite large before it would be the limiting factor. I

mean, quite small before -- you know, you wouldn't expect it to be the limiting factor in the powering of the study, I don't think, but I don't have any exact calculations obviously on that.

Also, you know, in the principle of do no harm, surely you wouldn't want to allow there to be any worsening. You want it to be equivalent to, in terms of pain and function, to the placebo group.

DR. SCHWIETERMAN: Dr. Abramson?

DR. ABRAMSON: Dr. Schwieterman?

DR. SCHWIETERMAN: This is a very difficult question. I think I may agree with Dr. Dieppe that it's an impossible question, but let me try to give the thrust of this.

In our earlier discussions in the agency, I think we generally agreed with the committee about the difficulties of using a surrogate endpoint, such as joint-space narrowing, to predict long-term outcome.

An inevitable question that then arose from that discussion is, given our skittishness, nervousness about this particular endpoint, to what degree would we not go forward with an approval if there was some evidence that there was clinically worsening, whether that be trends in admittedly clinically-debatable relevant differences or whatnot, and, of course, the question involves what's

practical, what's not practical, and so forth.

The point isn't that we have to simply demonstrate that something is clinically worse. The point is to what degree do even small trends in the worsening in the clinical signs and symptoms which, by definition, are clinically-relevant endpoints, bear upon this question, given the complexities of joint-space narrowing?

Now, having said all that, I'm not sure that there's an answer to it, but I hope that that at least clarifies things.

DR. ABRAMSON: Dr. Dougados, and then Dr. Elashoff.

DR. DOUGADOS: I think that, to try to give a practical proposition, that we have to keep in mind that we need placebo-controlled trials in order to show deterioration, yes or no, because we don't know the natural story very well, and we cannot anticipate that we will bring improvement because of the regression to the mean, as Paul emphasized.

So placebo-controlled study in order to avoid the one-millimeter difference in the VAS of pain or functional impairment or in global assessment. One possibility is to use a composite index. That's the one that we propose within the Osteoarthritis Research Society, responded yes or no, and to compare the index to responder

in the placebo, and as an example, to consider that a 10-percent difference between the placebo and the toxic drug can be accepted. That is a delta which has been proposed.

DR. ABRAMSON: I'm going to do Dr. Elashoff. I'd ask you to make the final comment. Then I'm going to ask that we break for 15 minutes right after this comment. We'll continue afterwards.

Dr. Elashoff?

DR. ELASHOFF: I think there are two different issues from a statistical point of view in defining no worsening. One would have to do with whether you want to define no worsening in terms of a mean not being below zero or whether you want to define it in terms of the mean not being worse than the placebo change, and then the other has to do with, however it is defined, you have to take not only into account the mean but something in terms of the confidence interval for that mean, and whether you go explicitly to an equivalence formulation or not, certainly confidence interval has to be part of the formulation.

DR. ABRAMSON: Thank you. We will continue this when we return. We'll take a 15-minute break. At 11:00, I'd like to begin promptly with the open public hearing. Thank you.

(Recess.)

DR. ABRAMSON: The next portion of the meeting

is the open public hearing, and I just want to make an announcement, a housekeeping announcement, about the schedule.

We're probably going to break for lunch around 12:30 or quarter to 1:00. We'd like to get through the questions through the duration on the design, and hopefully we'll do that before we break for lunch.

For the open public hearing, we have two registrants, and I'd like to call on Dr. Peterfy from the Department of Radiology, Stanford University, to give the first brief presentation.

DR. PETERFY: I think you got them mixed up.

DR. ABRAMSON: I'm sorry. Oh, I'm sorry. Go ahead. I'm sorry. Dr. Peterfy, Chief Scientific Officer, Synarc, Inc., San Francisco. I apologize.

DR. PETERFY: Thanks very much.

You'll have to excuse me if this presentation goes a little awry. It didn't seem to work with my PC. So I transferred it over to Phil's computer, which is MacIntosh. So anything, I think, could happen, but you may not be surprised to hear that I won't have any dramatic conclusive evidence to support the argument for cartilage loss as a surrogate for clinical and functional outcomes, and it's mainly because I think only recently have we really learned how to properly use radiographs even for

clinical trials, and only recently have we learned to take MRIs, a tool which was adapted really originally for clinical service, and then adapted for the different priorities of clinical trials research, and that it really brought to within reach the possibility of finally evaluating the joint as a whole organ consistent with many of the discussions that we've been having today.

So what I wanted to do really was to review what we've really learned over the last few years about joint space with measurement, with radiography, and the practical and theoretical issues surrounding that, and contrast that against also what the current status is of cartilage imaging with MRI and introduce some of the opportunities for whole organ scoring with MRI and then just indicate what the remaining questions are and what studies we currently have ongoing that will use these updated approaches from imaging to answer some of these questions, and how quickly we can expect results from those.

One thing that we have learned about radiography of the joint-space width is that knee positioning is very critical, and that's because in fact only a very small region of the articular cartilage is evaluated, that portion that is in direct contact, and in an incongruent joint like the knee, that's a very small

area, and which portion of the femur is articulating with the tibia at any one point really depends on the degree of flexion of the femur, and so on the one hand, the technical appearance of the joint-space width is dependent upon how much flexion there is in the knee.

But, in addition to that, it raises the question -- and excuse the overlap. This is part of the problem I was talking about earlier. But it really asks the question of which location is the most sensitive to change. Is that the same or different from the location that is most significant to clinical function and pain?

And we've also learned what other things must be standardized in image acquisition. I've mentioned the degree of flexion of the knee, weight-bearing in both the knee and the hip has to be controlled as well as rotation which is external for the knee and slightly internal for the hip, beam centering and alignment on the joint space, magnification and, of course, exposure settings. These are quite simple, but the others have been overlooked to a large extent until recently, and perhaps one of the pioneers of this, the most recognized one for standardizing image acquisition was Christopher Buckland-Wright, who really emphasized, and you can't see this here, I'm afraid, the importance of fluoroscopic imaging to position the knee in flexion, use of foot maps, and magnification markers to

correct for that.

The problem with this approach was that it needed special equipment, special training. The foot maps and markers were cumbersome to use, and ultimately this approach leaves the knee only in a slight degree of flexion, not necessarily in the most sensitive region, but there's been some potential improvements made along the way.

Foot map and that cumbersome approach could be replaced with frames which are standardized and allows the feet to be positioned quickly and reproducibly both for the knee and the hip, and the degree of flexion which in the past was aligned while supervising on a fluoroscope can now be replaced with non-fluoroscopic methods of fixing the knee in flexion. One can position the toe and the knee and then press the thigh up against the wall, up, for example, a frame, and then this will fix the degree of flexion of the knee, and then it's a matter of determining the right angulation, and in a study that we conducted in both normal and OA subjects, we found it to produce in direct comparison with fluoroscopy the same reproducibility error for joint-space width measurement which was .2 millimeters with the manual measurement method, and by turning the knee around like this and bringing it close to the radiograph, one also minimizes the degree of magnification and

stabilizes it across serial studies.

And then digitization and automated computer measurement further improved the precision, speed and capacity of joint-space width measurement which you can see is down to approximately 0.1 millimeters standard deviation and also enables these measurements to be audited.

Similar techniques have also been developed for the hip and improve on the approximately 0.3 millimeter reproducibility error for manual measurement with approximately .2 millimeters for automated measurements along with all of the other advantages.

So there's been, I think, considerable advance made on how radiography should be used for joint-space width measurement. However, it still leaves some unsurmountable limitations that are fundamental to the technique.

First of all, of course, it only provides an indirect visualization of the articular cartilage, and for that matter, other joint tissues, the major limitation to whole-organ evaluation, and then only a small region of the articular surface is actually covered by the technique, and that's owing to the projectional viewing perspective.

MRI carries a number of advantages. Direct visualization of the cartilage and other joints simultaneously really enables for the first time to perform

a whole-organ evaluation consistent with an organ failure model of the disease.

It provides full coverage of the articular surfaces because of its tomographic viewing perspective, and it provides compositional as well as morphological parameters, such as collagen matrix, bone marrow edema, and of secondary significance to this discussion, but there is considerable advantages, at least on a theoretical level.

Now, here is a beautiful correlation of an MR image with a scanning electron microscopy of the articular cartilage given to me by Doug Goodwin that simply illustrates the accuracy, the morphological accuracy that MRI gives relative to the articular cartilage, and, in addition to high-resolution specialized techniques like this, conventional MRI techniques are available on virtually all clinical magnets that are in use today, have also been looked at quite thoroughly.

Here you see an example of a three-dimensional technique that illustrates the cartilage as a high single intensity band, and you can see in this example, on a nine-month follow-up of a patient who's had a meniscal surgery, you can see a focal defect clearly illustrated, and several groups have looked at this and found it correlated with arthroscopy. That MRI with this technique is very accurate, both sensitive and specific for focal defects.

We applied a seven-point scale to cartilage evaluation and viewed this in 15 regions of the knee and found a very high interclass correlation coefficient for two trained readers blindly evaluating the same osteoarthritic joints.

In addition, we still don't know what the sensitivity range is for this scale or what its dynamic range is for clinical outcomes obviously, and, in addition to those subjective parameters, thickness mapping quantification is also possible now. There's been considerable work done with it in Germany. This was a study done by Zohara Cohen in Van Miles' group, and you can see that the root mean square standard deviation for thickness measurements is .3 millimeters, roughly in the range of joint-space width measurement, and this was for 450 micron resolution image, something that in fact one can improve upon relatively easily with conventional clinical hardware.

Other quantitative measures are volume quantification. This is something that's easy to do with a number of image processing software, and the correlation between the MR-derived volume and the true volume of articular cartilage has been found to be relatively high by us and by other groups, as well as the reproducibility error for these measurements has been found to range

somewhere between 2 percent and 4 percent.

There has been relatively few longitudinal studies thus far using these new techniques. They are relatively new, and longitudinal studies take awhile to complete. One study that we did didn't find reasonable, some might say, rates of loss in a cohort of osteoarthritic women, around 6 percent in the femur and tibia, half that in the patella, compared to relatively no change in a small group of control subjects. This is roughly consistent to what joint-space width measurements have revealed in some studies in the knee and the hip.

In addition to these morphological parameters, a very intriguing aspect of MRI which you saw Phil Lang's discussion address, is the possibility of looking at earlier steps in the pathophysiological process before actual cartilage loss. In this case, collagen matrix damage, and on a conventional MRI image that's available on most magnets, as I said, normally one can see a great deal of signal from three water molecules like this, but in the presence of collagen, those water molecules become immobilized, and the signal from them therefore decays very rapidly, and so as that collagen disappears from the articular cartilage, the tissue water becomes more fluid in its behavior, and it shows up as an increased signal intensity like this, and this has been looked at at the

biochemical level and the histological levels by several groups, quite thoroughly in the correlation or, let's say, the biochemical validity of this is very sound.

What isn't known still is the dynamic range, sensitivity change and some of those metrics of performance for a marker as it applies to clinical trials.

Here is an example along the lines of what Philipp was doing. This is a patient who's two months post-lateral partial meniscectomy. If this was a little brighter, you could see it better, that there's a high signal intensity focus in that articular cartilage over the operated meniscus. That is the exact site of a focal defect nine months later.

One can quantify these T2 abnormalities in the articular cartilage and potentially track those using conventional MRI pulse sequence.

Other questions that still have to be answered, as I mentioned, sensitive to change, dynamic range, and also the optimal technique in a practical sense for large-scale clinical trials, and then a number of other markers that I won't discuss because they're farther from actual application, but that each offers some opportunity to look at either collagen or proteoglycans in variable degrees with MRI, and this is something that really was not possible even a few years ago.

I do want to mention a word about whole-organ assessment. This is something that is uniquely possible with MRI today, and that many elements of which we've actually been doing clinically for many years, looking at menisci, cruciate ligaments, et cetera.

I've combined them in a few studies into a scoring method that looks at nine articular surface features in 15 sites -- I'm afraid this didn't translate well in the MacIntosh -- and seven other articular features, including the menisci, the cruciate ligaments, collateral ligaments, and the synovium.

Basically, without going into detail, the output of this analysis is illustrated in this table where each of the values in the table represent the degree of damage to that particular feature in that compartment, and on the right, you see totals for that feature, and here, by compartment, the totals. One can find a global total for the knee, and really it paints a much richer picture of what's going on in the knee than we've been able to deal with up until now.

These are baseline values from a study, 208, a database here, two-year follow-up has already been done. We're just analyzing that data, and the longitudinal performance will be interesting for that. So far, we have looked in another study at the inter-reader variability of

this method and found it to show a high reproducibility with trained readers, and it represents really a technique that's simple to perform and therefore applicable to multi-center studies, will extrapolate easily to clinical service. It shows a low inter-reader variability and provides a more thorough evaluation of the articular surface and information, of course, about the other joint structures.

The limitations to date still include some information lacking about its longitudinal performance, essentially to change the dynamic range, et cetera. Clinical correlations are being carried out but have not been completed yet, and so just to summarize, for joint-space width measurement with radiographs, we now have methods that have been improved and adapted to multi-center clinical trials.

There still remains some fundamental limitations with radiographic information, particularly only small region of the cartilage is evaluated, and there's incomplete information about other joint structures.

MRI shows a number of fundamental advantages, including full anatomical coverage, broad tissue contrast allowing whole-organ evaluation of the joint, and there are now techniques for measuring cartilage thickness and volume

at least accurately, precisely and with pulse sequences that are applicable to multi-center trials, and there is some work ongoing with whole-organ scoring.

Outstanding questions still remain with regard to the clinical correlation of joint-space narrowing using the thick selection radiographic acquisitions and to the MR measures of articular cartilage loss I mentioned earlier. Exactly how much slowing of cartilage is clinically relevant has been a question that has been raised, whether they're different in the knee or the hip, and a direct comparison of the performance of MRI markers with the radiographic joint space narrowing.

Using these particular techniques has also not been answered yet, and the importance of other structural features in the clinical significance.

Now, the study --

DR. ABRAMSON: Dr. Peterfy, one minute, please.

DR. PETERFY: Yes. This is my closing.

DR. ABRAMSON: Okay.

DR. PETERFY: Therapeutic trials that we have in progress that use these techniques this way that are one year or older include some 650 knees with MRI that are being scored for cartilage, volume quantification, T2 quantifications, synovial volumes and whole-organ scoring, and 1,700 radiographs using the thick selection technique

that I mentioned earlier, and automated joint space with measurement, 200 MRIs of the knee using the same MRI scoring and quantification techniques, and 1,700 radiographs using a joint-centered acquisition and potentially automated joint-space width measurements.

And we're looking at and we're seeing patient populations in 1,900 cases, using standardized acquisition, and epidemiological studies greater than a year, we've got 2,400 knees, 2,200 of which are utilizing a 20-minute MRI pulsed sequence, and there, we're looking at cartilage score, whole-organ score, and there's some 1,800 radiographic knees with the fixed flexion automated joint-space width measurement being done with these, and, of course, each of these studies have the usual clinical endpoints being measured and also in many cases, especially in selective biomarker correlations as well, and so these are all already over a year into the study, and we anticipate at least some interim results coming out of these quite soon, but it's along this time frame that I think we can expect some of the answers to the questions that we've been talking about to begin appearing.

Just let me close with an acknowledgement of some of the individuals who gave me some images that I showed you here today.

DR. ABRAMSON: Thank you very much.

Dr. Brandt?

DR. PETERFY: Sorry for rushing through that.

DR. BRANDT: I'd ask the same question that I put to Dr. Lang earlier with regard to specificity of any of these changes or very, very sensitive technique, that's very clear. What did you mean by that last slide that said epidemiological studies? We have a large number of knees. Does that address this issue of specificity? What's epidemiologic mean?

DR. PETERFY: Well, what I mean is not clinical trials, not therapeutic trials, not drug trials, but rather an NIH-funded study, a health agency study, where there isn't actually a drug that's being tested.

DR. BRANDT: No. What I really --

DR. PETERFY: In terms of the specificity, if you mean morphological specificity as to whether an abnormality on the MRI, actually what it represents histologically or biochemically, that's one thing.

DR. BRANDT: No. What I really would like to know is whether these changes that you're elucidating and presumably in patients, in people who have symptoms, that's why they come to you --

DR. PETERFY: Right.

DR. BRANDT: -- are different from those in an age and sex match population of older people who don't have

symptoms.

DR. PETERFY: Half of the knees that I showed in the epidemiological studies or a thousand knees MRI'd controlled subjects without pain. So that question is being addressed in that particular study.

DR. ABRAMSON: Dr. Dougados?

DR. DOUGADOS: Just to come back to the radiographic evaluation, this morning in the first part of the meeting, we have emphasized the clinical relevance of the evaluation of joint-space width at the narrowest point, and Charles is giving two possibilities after digitization looking at not the joint-space width but the joint-space area evaluated after digitization, and that at least to my knowledge, there is no reason that the evaluation of the area is better than the evaluation of the anterior bone distance, and it's quite easy to imagine that if you have a a localized narrowing of the joint space, there will be either reparation or an artifactual increase in joint-space width in the -- so I am not sure --

DR. PETERFY: No, I wasn't advocating area over minimum joint-space width. There's several parameters with a digital image and a computer-assisted method. One can acquire numerous different ones. In fact, the beauty of image data is that you can test new measurement algorithms as they occur to you because it does not discuss the data

the way biomarkers specimen uses up a bit of the specimen each time.

But for minimum joint-space width, what we use right now in our clinical trials, there is an issue of shift of the minimum joint-space width sometimes between the original and the follow-up exam, but back registration of the two images to the narrowest point might be a way of overcoming or improving on that.

DR. ABRAMSON: Thank you very much, Dr. Peterfy.

We'll move on to Dr. Lang.

DR. LANG: Dr. Abramson, ladies and gentlemen, Dr. Peterfy just gave a beautiful presentation of the current state of MRI and the various techniques that are routinely available.

I want to take the opportunity here to focus on some of the new techniques that are evolving in MRI and that lend themselves in the future, assuming appropriate testing, reproducibility, accuracy, et cetera, and we're doing some of these studies as we speak, that are techniques that lend themselves as new surrogate endpoints in clinical trials in OA.

The three major areas of development that I believe we will see in the next couple of years to evolve from MRI is, A, morphologic analysis. MRI will get a lot

better at visualizing the cartilage than it is right now. Biochemical composition can be assessed, and if a drug has an effect at the biochemical level, such as enhancing glycosaminoglycans in the cartilage, that effect can be measured with MRI, and, finally, biomechanical assessment.

Because of the short time that I have available, I want to show just a very few representative techniques that I believe will have high impact, and I do want to point out that these new techniques are currently being used in two Phase II studies at Stanford University.

They are using these new techniques in conjunction with the current established techniques that have been validated in multiple previous studies by Drs. Peterfy, Rect, Eisler, and a lot of international collaborators.

This is a technique called projection reconstruction spectroscopic imaging. What this technique offers is very high-resolution images of the articular cartilage, but, in addition to that, you can actually quantitate water content in the articular cartilage and get spectral information, and as you see going from the subchondral bone plate to the articular surface, the line height and line width in fact also changes, which is a reflection of change in water content across the articular cartilage, and again we're currently using this technique

in an OA trial.

This is a technique that was pioneered by Deb Burstein in Boston and her group, gadolinium-enhanced imaging, and again also a technique that lends itself and that is actually technically very easy, lends itself to using clinical trials, gadolinium-enhanced imaging. What happens in this case, glycosaminoglycans in the articular cartilage, in normal cartilage, which you see here imperially, carry a negative charge.

Gadolinium-TDPA as standard MRI agent also has a negative charge. So in normal cartilage, the negatively-charged glycosaminoglycans will repel the gadolinium, and it will not get into the cartilage after an IV injection or only small amount will get into the cartilage.

In cartilage that is glycosaminoglycan-depleted, which is here shown superiorally, the gadolinium does not get repelled anymore. It leaks into the cartilage over a period of 60 to 90 minutes, and the change in T1 relaxation time is a direct reflection of glycosaminoglycan content in the articular cartilage, and you see this applied here in an OA patient currently enrolled in a trial, and you can in fact see areas of healthy cartilage with the long T1 relaxation time on this T1 map, high signal intensity, and areas of glycosaminoglycan-depleted cartilage with lower signal intensity.

By the same token, a new technique, sodium MRI, that has been pioneered by Dr. Reddy at the University of Pennsylvania, Philadelphia, is another in vivo probe to get at proteoglycan or at glycosaminoglycan concentrations. The principle here is similar. The positively-charged extracellular sodium inside the articular cartilage tracks very closely with the glycosaminoglycans.

What the Penn group has done here -- this is in fact the patella. This is not a proton but a sodium image, and the intensity in the image is a reflection of sodium concentration. What they did is they applied a membrane in the region of the knee enriched and then exposed the lateral facette to a protease, and the cartilage became proteoglycan-degraded, and you see the decrease in intensity which is a reflection of the loss in proteoglycan induced by the protease, and again this can be done in vivo, and we're currently doing this in vivo.

Three-dimensional thickness map. Dr. Peterfy touched briefly on this. This is an example of a focal defect in the posterior femoral condyle, and as we go from anterior, we see normal thickness cartilage going posterior, normal thickness cartilage. Here's the full thickness defect, and again normal thickness cartilage.

Now, you can map the thickness along the condyle here and along the X axis. Y axis is the thickness

in millimeters, and you see this thickness graft with a reduction to zero thickness at this level.

You can do it in 2D. You can do it in 3D, and this was first done by Dr. Eckstein's group in Munich, and our own group using funding from the Whittaker Foundation continues to develop these techniques, and here you see a three-dimensional thickness map of the articular cartilage in the femoral condyle in the same subject. The thickness is color-encoded. You see the focal defect here in the posterior lateral femoral condyle, and, currently, this technique still requires manual interaction, but, ultimately, this will be fully computerized. In fact, we believe in the next three to six months, we hope to have this fully computerized.

This is a technique pioneered by Professor Thomas Andreaki, formerly in Chicago, who has been for one year now at Stanford University. This is gait analysis. Gait analysis is a technique that can be used to study abnormal gait patterns in human subjects.

The retro-reflective markers are applied to the skin. The patient is walking up and down a defined path. The force plate is applied in the ground, and you get an estimate of abnormal loading patterns, like Dr. Brandt pointed out earlier, abnormal quadriceps mechanism, et cetera.

The problem with gait analysis -- this is in fact normal stair-climbing. The problem with gait analysis is that it assumes a standard femur and standard tibia. This is the same subject doing stair-climbing. This is a healthy volunteer, and I want you to think of this next time you're walking up the stairs because this is how your femur is grinding on the tibia.

But again this is a standard femur and a standard tibia. With gait analysis currently, we have no means of looking inside the patient. Well, not any longer. In the context of an NIH submission, we developed new software merging MRI with biomechanics -- i.e., gait analysis -- and here you see a patient who was initially in the gait lab, had the same test done that I just showed you, and subsequent to that underwent MRI, and the markers used for gait analysis are now filled with gadolinium. So we can cross-reference them.

Based on the MRI, the 3D reconstruction of the patient's actual femur, the true femur in the subject, was generated. Femoral condyle cartilage, tibia cartilage, and the tibia, and this is the type of information that you can get from this test. This is the actual knee of this particular patient based on MRI, and subsequent to that, the gait pattern that this subject had in the gait lab applied to this patient, and this just shows you where I

think this field will go in the future.

Yes, Dr. Brandt is a 100 percent right. Currently, we cannot integrate biomechanics, but with techniques like this, and there will be other approaches to this, this integration will happen, and I think MRI will become an even more powerful test.

This was a normal volunteer. Well, in fact, the subject thought he was normal, but as was elicited by this test, on every single heel strike, we look closely, there's hyperextension, which is not quite normal, which is not what you ought to see.

So in summary, with the current technologies, MRI can provide a very detailed assessment of cartilage morphology, and, yes, we need to get more data, and we're currently performing studies in fact to collect this data, to get at issues, such as work reducibility.

We can estimate biochemical information, such as water content and proteoglycan content, and again validation studies are underway, and I hope that we will have some data on this by the end of this year.

We can obtain quantitative information, such as 3-D maps of cartilage thickness, and, ultimately, over the next two or three years, I would hope, biomechanical information, and I believe that prognostication of defects will be possible based on size, location, composition and

biomechanical stress derived from MRI, and these are also potentially very useful surrogate endpoints in the future for clinical trials in OA.

I thank you for your attention.

DR. ABRAMSON: Thank you very much.

Any comments from the panel? Dr. Harris?

DR. HARRIS: Just one comment. It's remarkable that as things get more complex, and as one can define things a little more clearly, they become more difficult to measure, if you see what I mean.

If we just had thickness and volume, and we could do so reproducibly, there's a single measurement. Once one starts talking of cysts and water content and all the other sorts of variables, then one has to come up with global assessments and so on, and it becomes awfully more difficult, and, you know, this is just an observation, that perhaps being simple might help a lot more than being complex, at least as far as studies go of the sort that we'd like to do here.

DR. LANG: I agree completely, and my personal opinion, if I may share that with you, is that visual analysis, I think, will remain quite powerful. The subtle lesions that I showed earlier will be very difficult to quantitate with the computer, and the other surrogate outcome, I think, that will be very powerful in the future

are these thickness maps which are in a sense directly related to the work that has been done by Buckland-Wright, by Dr. Brandt, and a lot of other people here in this room, looking at joint-space narrowing, except in this case, you don't look at a focal point like Dr. Peterfy pointed out but really along the entire articular cartilage, and I think this will be an extremely useful surrogate endpoint.

Some of the other parameters that I showed here, I think they will be scientifically very interesting to better understand these drugs, how they really work in vivo, what effects they have because now we have a non-invasive in vivo probe to look at biochemistry, and indeed in that same context, perhaps to address specific questions also relating perhaps even to safety.

DR. ABRAMSON: Okay. Dr. Brandt?

DR. BRANDT: That's beautiful stuff you showed. It's important to keep in mind, I think, as we talk about joint-space width and surrogates, that in the standing knee radiograph, what's being measured is in fact a sum of the thicknesses of the femoral and tibial cartilage, plus their mechanical properties with compression under weight-bearing.

The thickness of cartilage measured by MRI in an unloaded joint gives something entirely different and not necessarily related because there's no compression that

you have in a standing knee film, and those have to be reconciled at least at this point.

DR. LANG: I agree, and Felix Eckstein at Munich has done some very beautiful work on MRI, actually compressing the cartilage, and in cadavers on the regular unit, he was able to show compression in the situation that was simulating weight.

Something that I didn't show here because of time is we have a vertically-open MRI unit at Stanford University. In fact, you can stand in the magnet, and we have done a couple of volunteers now, and we're trying to get OA patients in. In fact, we have the patients standing like this in the magnet in upright position, weight-bearing. You scan the articular cartilage, and, yes, you do see reduction in height of articular cartilage on the order of 10 to 25 percent in one subject even, and that is something that has to be taken into account.

DR. ABRAMSON: Dr. Dieppe?

DR. DIEPPE: I think I'm going to be at risk of sounding like a Luddite, but it does seem to me that it's rather more important that we figure out whether the structure of the joint really matters very much before we get more sophisticated at measuring it, and at risk of upsetting my colleagues who have given beautiful presentations, I've been hearing presentations for the last

10 years that say next year, MRI will have cracked it. It's always next year.

DR. LANG: Well, my answer in this case would be help us to get the NIH funding and give us another year.

DR. ABRAMSON: Okay. Thank you very much.

Yes, Jim Witter?

DR. WITTER: Just another kind of regulatory perspective question. Let's assume that we're seeing part of the future here, and, by the way, I really enjoyed the presentation. Do we need to worry about corruption of the data? I mean, because this will all be digitalized, and one could imagine it can be easily manipulated. Could you give us some sense of that?

DR. LANG: I think every time you have a visual analysis of x-rays, there is chance for operator error and incorrect transfer to case report forms or from the case reporter form ultimately to the computer, and, yes, the same thing holds true for a computer analysis. In particular, if you have, let's call it, an analog step in -- I think the weakest link is the human subject, where the human subject has to interact, and the biggest problem right now with these 3D volumes and 3D thickness maps is that you actually have to segment the cartilage from the MRI image, and that's the main thrust of our research right now to automate this, and we're making progress with it,

but we haven't achieved that yet, and, yes, that's the point where the data could be corrupted.

I think with some of the new software algorithms that have been developed by a number of different groups, in particular also the Munich group is very active there, this will look a lot better. In fact, I think even this year, we will see some data that will be very encouraging, but these are issues that have to be addressed, I agree.

Once it is fully computerized, I think it will be more reliable, and we will be less subject to these types of errors than in fact any type of visual analysis.

DR. ABRAMSON: Thank you.

At this point, are there any members of the audience who did not register to speak at the open public hearing who would like to make a comment?

(No response.)

DR. ABRAMSON: Thank you.

What we'll do is we'll go back then to the questions. When we broke before the break, I guess Dr. Elashoff had made a final comment regarding the ability to define worsening.

I guess I would ask Drs. Johnson and Witter if they want to continue that piece of the discussion or move on to the Question Number 3?

DR. WITTER: Can I just maybe try and summarize what I think I heard?

DR. ABRAMSON: Sure.

DR. WITTER: In the sense that one could take away that generally, we're more concerned about in defining worsening or no worsening in a sense of measures, such as VAS pain scores, rather than some of the stuff that I tried to get at earlier in terms of serious adverse events, deaths, things like that.

I mean, is that kind of the sense of your main concern? You're assuming that those other parameters that can factor into worsening, meaning serious adverse events and deaths, are going to not be a factor in here?

DR. ABRAMSON: Well, I guess what I was thinking is that with respect to the efficacy of the drug and the joint-directed adverse effects, that would come through by continuing the parameters that one was measuring in terms of pain and function.

Whether there were unanticipated adverse events outside joint structure and function, obviously one would continue to monitor for those kinds of -- you know, whether it's shoulder stiffness and fibrosis outside the signal joint, one obviously would need to follow those kinds of things as well.

But unless other people feel differently, I

didn't hear a suggestion that one should change the kinds of parameters that one was following, and the question became more of a statistical one. Could you show equivalency to placebo or worsening in that Phase IV context?

DR. WITTER: Then, I guess, as a related question, could we have some more discussion on the psychosocial type of outcomes that have been mentioned? I mean, are we talking about SF36s or modified HAQs or are we talking about something entirely different?

DR. ABRAMSON: Perhaps people who are currently engaged in such studies want to comment. Dr. Brandt?

DR. BRANDT: Before that, Jim, just one comment with regard to worsening, and I think we have to take a broad look at this. It's conceivable that a structure-modifying drug could be associated with some worsening of pain in a patient who is in fact doing three times more than he was doing before, and so I think they both need to be taken into account and put into perspective.

DR. JOHNSON: But not worsened compared to control, though, because that's always the caveat here.

DR. BRANDT: Yes, he may be worsened. The pain may be worse than in the control, but in relation to a doubling of the activities that he's able to perform.

DR. JOHNSON: Yes, that gets around to sort of

integrating this efficacy measure which we also want to try to get some feedback on, but, yes, you could have a drug that differentially works wonders on one and not the other.

It sounds like Marc's suggestion about using prior trials to figure out an effect size might give us a handle as to what fraction of the effect size we could deem ignorable in an equivalence test, and, you know, if you make it a really tiny difference that you're going to ignore, then your sample sizes go through the wall, but if you make it relatively liberal, it might not mandate gigantic trials.

I, too, as Jim mentioned, would be interested in Paul and David Yocum, if he's still here, about what they would consider useful psychosocially.

DR. ABRAMSON: Paul Dieppe?

DR. DIEPPE: I think the data that's available suggests that the three key psychosocial variables that have been associated with pain and disability in OA are anxiety, depression and social isolation.

Now, I actually think personally that the SF36 is a nonsensical instrument. We all use it because we're all using it, but it doesn't seem to me it's got any sense to it.

Having said that, the mental subscale of the SF36 is probably one of the better subscales in my view in

this context and correlates reasonably well with other more specific measurements of anxiety and depression.

There are plenty of other standardized measurements of anxiety and depression that are well-validated that can be used, and I guess it's another meeting to discuss which of those should be factored in to this as a standard measurement, but I certainly think one or more should.

Isolation and other issues of that sort in terms of lifestyle change are much, much more difficult.

DR. JOHNSON: Does the epidemiology suggest that if you're talking about a two-year trial, for instance, these are not as important? At what point does this kick in? You had mentioned the short-term and the long-term dichotomy before.

DR. DIEPPE: Well, they kick in all the time because, you know, I talked also about the variability of pain and disability, and that's dependent on this. I guess the issue of when do those issues start to overwhelm other issues, I've no idea, in time frame. That's -- I don't know.

DR. ABRAMSON: Yes?

DR. LIN: I also think that this is equivalence problem, perhaps not, you know, from a physical point of view, perhaps not two-sided.

DR. ABRAMSON: Is your microphone on? I'm sorry.

DR. LIN: But it can be a one-side non-inferiority consideration. What?

DR. ABRAMSON: I'm sorry. Please introduce yourself.

DR. LIN: Wait, wait, wait. In thinking about this, I think that it would be important to specify -- because we're talking about joint-space narrowing as the primary endpoint in this trial, I think it would be important to specify at what point you're going to make assessment that pain and function were not worsening. I mean, you have to do that because somewhat after the Phase III trial, you will want to know if there was improvement in the pain measurement. So you want to measure not worsening during the trial, but then some time later, you want to see if it is correlated with pain improvement.

Okay. So therefore, you need to carefully specify at what point you want to assess the equivalence between the two, okay, and you can do that with a one-point measurement.

An alternative would be that you can make pain measurements repeatedly over the Phase III trial and perhaps take a repeated measures approach and see if the slope between the groups were different. I mean, it seems

to me that may be another approach or even whether the slope remained constant over time, and that would be another approach.

DR. JOHNSON: That was Stan Lin. He's one of our statisticians.

This scene comes up recurrently, too, whether you're just doing an endpoint analysis or a multiple-point analysis. We put a little verbiage in the RA document that unless there's a strong reason for not recognizing all the points, you should recognize all the points.

On the other hand, this will be a one-point in time analysis. It will be presumably at approval for structure, that we do some kind of non-inferiority analysis or equivalent analysis of the symptoms at that point.

DR. ABRAMSON: So what would worsening of pain that was observed during the course of the study, just in terms of analyzing the data in terms of the data monitoring?

One can imagine perhaps something that measured at one year might improve structure but could exacerbate pain in the course of the prior months, perhaps even as an unrelated adverse event if arthralgias, for example, are part of the profile of some of these kinds of agents.

So if there's worsening of pain and function before the one-year analysis, how would one think about

dealing with those data?

DR. WITTER: I just want to broaden it out a little bit to signal versus non-signal joint, that whole paradigm, too. I mean, we had a discussion on that.

DR. JOHNSON: The quick answer is we'd bring it to the committee. But presumably, if you have worsening of both pain and function, you're probably in trouble -- this is again compared to placebo -- unless it's just a very small worsening, and if it occurs at all points but not at the point of approval for structure, then you have to wonder what's going on, you know, if all the other points are positive but that one point is not.

DR. ABRAMSON: Okay.

DR. DOUGADOS: Excuse me, but I don't understand the discussion. I can't understand the situation where such as we've been discussing, we have a structural deterioration together with a symptomatic improvement, but in the real world, I don't know a situation where we will have an improvement in the structure within one year with the deterioration of -- I can imagine by chance, that you will see a study signal significant difference in favor of the placebo. That is only what you want to avoid because if you want to anticipate that, can you imagine the drug, whatever the mechanism of action, which will result in the deterioration

of the symptoms not by chance but due to the drug?

I can't imagine side effects which have just been emphasized, side effects, or statistical significance, but not to real deterioration in symptoms due to the drug.

DR. JOHNSON: Well, it is hard to imagine. If the hypothesis is wrong, however, that -- and in fact, joint-space narrowing is not a valid surrogate, then if you also didn't have a drug that was active, then -- or if somehow the mechanism of the drug engenders worsening of symptoms.

I mean, I think it's clearly a logical possibility, and it's the one that we need to guard against analytically at approval time. It's just a question of how to do it analytically.

DR. WITTER: And, Maxime, I think a lot of what we do in our deliberations when we discuss these kinds of issues is trying to imagine also what we might be seeing because part of the guidance document is not only building on some kind of foundation but also projecting what we might see in terms of therapies.

So my presentation, one of the things it was trying to do was to answer, I think, how do we make a distinction between a safety adverse event in, let's say, a joint that's not the signal joint versus something related to clinical worsening, and are those the same kind of

thing? I mean, do we think about those events in the same way?

DR. ABRAMSON: Dr. Brandt, then Dr. Hochberg.

DR. BRANDT: OA is a disease of an organ. I'll give you an illustration that would fit that situation that you asked about.

Twenty years ago, orthopedists, some orthopedists in this country, shaved osteophytes in patients with osteoarthritis, and they thought they were doing something good.

In fact, what happened as they increased mobility is the joint went to hell in a hand basket, and patients became symptomatically worse, and it ended up being a first-stage procedure for a total joint arthroplasty. So instead of having a Bard-Parker like osteophytes, consider a drug that might do that.

DR. DOUGADOS: Yes, but the question is not related to the symptomatic deterioration. The question arises, can we consider the change in the joint-space width as a relevant surrogate marker of the condition? So we are coming back at the beginning.

If no, you will never resist a drug with a structure effect without any evidence on clinical symptoms. If yes, you will accept the registration.

DR. JOHNSON: Yes. We're not advocating that

you should require, you know, clear standard clinical efficacy, but we're saying, well, should there be any test at all? I mean, would you register this drug if there was major clinical deterioration compared to placebo? You probably wouldn't.

All we're just saying is that we need some sort of no worsening criteria, I think, in order to use accelerated approval.

DR. ABRAMSON: Dr. Hochberg?

DR. HOCHBERG: Well, I don't want to drag out this discussion because this will take us into lunch, and we'll never move forward, but, you know, the other issue that the agency will have to deal with is what are you going to do about co-therapy?

I can't conceive of a patient with symptomatic OA going into a trial of a structure-modifying agent and not taking a symptomatic drug, unless the trial is designed to look at an agent which is going to affect structure and symptoms.

So if you want to look at an agent which is just going to affect structure, there's going to be some background co-therapy, and it's unlikely that that background co-therapy is going to stay the same if you want to keep the patient in the trial.

I don't know what Dr. Brandt's experience is in

his doxycycline study, but, you know, most patients are going to continue to take medication. The amount of medication they take will probably vary day-by-day. They may change their medication. So this is again something that the agency's going to have to make the sponsor of the study collect data on and consider in this whole analysis, right?

DR. JOHNSON: Yes. I'll let Jim respond, too, but the co-therapy is fundamentally, you know, a risk factor and may be critically important as a covariate, and the no deterioration equivalence test or whatever it is at approval for symptoms have to be robust to all these objections or else it won't hold water.

DR. WITTER: Would your advice be that using less co-therapies is a clinical benefit for a structure-modifying agent?

DR. HOCHBERG: I'll defer the initial response to Professor Dieppe, who's anxious to talk into the microphone.

DR. DIEPPE: Thanks, Marc.

I think the answer to that has to be yes, but I wanted to say that don't restrict the concern with co-therapies just to drugs. You know, all my patients with osteoarthritis are using dozens of other things for their osteoarthritis in addition to drugs. Physical therapies,

behavioral interventions, walking aids, goodness knows what else.

Some people in this room know that I think they're much more important than drugs myself, and they've got to be factored in to any of this. So it isn't just drugs.

My view is that, yes, that the utilization of those other therapies is an important issue, and I think it has to be potentially looked at economically as well as in relation to an actual symptomatic change, and it sort of comes back to one of Marc Hochberg's early comments. What are we trying to do here?

It seems to me one of the things we're trying to do is reduce the utilization of expensive interventions, and physical therapy, for example, is a hugely expensive intervention.

DR. ABRAMSON: Dr. Brandt?

DR. BRANDT: Using co-therapy as an outcome variable puts us on ice as thin as that on which we skate when we use joint replacement. There are studies that show in older people taking non-steroidals, presumably for osteoarthritis, half of them don't need the non-steroidal. They can be stopped, and the next six months, they go on very, very happily without any requirement for anything. So to simply measure what they're taking does not

necessarily tell us they're needing it. They're taking it because they're having pain. It may be ritualistic and something that happens at the level of the thalamus.

DR. HOCHBERG: If I respond now after these two comments, my thinking about it for five minutes is I wouldn't be keen on using that as an important variable to consider registration of a structure-modifying agent, whether it reduces the need for co-therapy which is symptomatic or not.

DR. ABRAMSON: Dr. Dougados?

DR. DOUGADOS: One comment concerning the evaluation of the symptomatic effect in a trial in which the primary objective is structure, and it's difficult for two reasons.

The first one is I do agree with Ken Brandt that the baseline characteristics of the patients are usually different in a structure trial than a symptomatic trial with lower symptoms.

The second thing, that during the treatment, because of the concomitant therapy and the inter-relation between the concomitant therapy intake and our symptomatic outcome variance, there is an interference, and it becomes more and more difficult to pick the symptomatic beneficial effect because of the two things, the duration of the treatment, the baseline characteristic of the patient, and

the concomitant therapy.

So that is quite difficult to demonstrate, and I can't anticipate that with a drug which will be able only to have an action on the structure within one year, it will be difficult to demonstrate symptomatic effect even within three or four or five years. That will be very difficult.

DR. JOHNSON: Yes. We haven't even touched on the possibility that your structure trial -- I think this is what you were mentioning, actually. Your structure trial may a priori need to be different or that it would be desirable for it to be different than your symptom trial, and it's even conceivable that the dose optimization is different structure versus symptoms, but in any case, you know, if we're going to use the accelerated approval paradigm, we need to make some assessment as to what's happened symptomatically.

I mean, we would feel an obligation to do that, to at least ensure that there wasn't any major deterioration compared to control.

DR. ABRAMSON: Thank you.

Let's go on to Question 3. "Must Phase IV symptom demonstration be done only via continuation of Phase III trials in which joint-space narrowing was demonstrated?"

I guess I would ask for a clarification of

that. By definition, the demonstration of joint-space narrowing would lead to the registration of the drug, and so the question that's being asked is whether or not the same criteria should be carried forward for pain and function or whether new criteria should be established for assessment of symptoms.

Can you give us some clarity of what you're going after here?

DR. JOHNSON: I think part of what we're going after here is sort of the linkage issue that Bill Schwieterman brought up this morning.

I mean, to what degree do you think it reasonable to have something ongoing? It may not necessarily be the same trial but at least a trial ongoing, because the problem is going to be, you know, as was mentioned earlier, if it proves to be quite successful structure-wise, there's going to be a great temptation to sort of, you know, corrupt the control as was mentioned in one of these other questions.

I'm not sure. I don't think the accelerated approval statutes require that these things be done in the same trials. It may require that something be ongoing. Maybe Bill knows the answer to that.

DR. SCHWIETERMAN: Actually, I'm not 100 percent certain on the language. We certainly have

approved agents on the accelerated approval where the study was not already ongoing. However, we view that more as an exception than as a rule.

I've spoken at length on this issue with Dave Feigel, for example, in the Anti-Viral Division in the Center for Drugs, who has the bulk of the experience with accelerated approvals for the treatment of HIV therapies, and it was that particular center's policy that they like to see not only the trial structured and submitted but actually ongoing and continuing at the time and actually beyond that even spelled out what kind of time frame they expected to see the clinical data coming in.

So while I don't think it's part of the regulations themselves, I think it's fair to say that most people within the agency would want the trials accruing/ enrolling patients and an expected time frame for the clinical outcomes to be submitted to the agency.

DR. ABRAMSON: Any comments from the committee?

(No response.)

DR. ABRAMSON: I guess there are two issues, aren't there? One is the development of later symptoms if the one-year time point is the time point for approval, and the other is again as a surrogate. What is the long-term outcome of using that as a surrogate, the x-rays as a surrogate, for example, in terms of clinically-significant

outcomes in osteoarthritis?

So the length of time of this follow-up Phase IV is also at question, right, as to is it a year, is it three years, and how rigorous it is? Is it the full study?

DR. JOHNSON: Yes. The length would depend on what you think your drug is going to do. It could be six months, it could be three months, it could be three years.

DR. ABRAMSON: Yes, Dr. Anderson?

DR. ANDERSON: Yes. I'm just wondering about this expression accelerated approval. I mean, does it really mean in practice that it's a conditional approval, and the approval would be withdrawn if, in this Phase IV study, the results were negative or contrary or to what was thought to happen?

DR. JOHNSON: Yes, that's what the statute provides for.

DR. SCHWIETERMAN: One thing I just want to add, in addition. This is an evolving field. The accelerated approval, which is probably better termed "conditional approval," as I think Dr. Witter or Dr. Johnson mentioned, was done in 1992 almost exclusively in response to the AIDS epidemic and the problems associated with getting promising but unproven therapies available to that population. The agency has since taken the statutes and recognized that they can be applied elsewhere.

The problem becomes, however, with these trials, and this is, I guess, the underlying assumption here, is that if you approve a product under accelerated approval, how do you keep a patient on the placebo arm for any length of time without either making it an unethical trial because by definition you've approved the product with some reason, that is equipoise may be disruptive, and then just simply pragmatically patients don't enroll in these studies, even if there is a fair amount of equipoise because they can get the product simply through their doctors, if they pay for it.

So I think this has been something that has yet to be resolved with many chronic therapies where sponsors have approached the Center for Biologics at least, and I know that Dr. Johnson's mentioned this before, but it's going to be especially a looming problem in this field if we're going to require three-four-five-year outcome measures down the line. It may require this committee to go for enrichment studies where we go for shorter-term outcomes in the patient population, more likely to progress sooner rather than later.

The reason I think it's fair to bring it up now is because it's been an issue for several years anyway, and now that therapeutics are getting into the chronic, especially the biological therapies, getting into the

chronic phases of things, how to develop these studies is not obvious to us.

DR. ABRAMSON: Can I ask a related question? I think we've been talking around this issue all morning, but the reason that we're discussing this as an accelerated approval rather than just a conventional approval is because of the uncertainty about the radiographic or imaging endpoints that are out there as surrogates for clinically-significant, because otherwise one could argue why is this different from rheumatoid arthritis which is also a 20-year disease, where we have so-called remittive agents approved based on six months or a year of therapy?

DR. SCHWIETERMAN: Yes. I'll let Dr. Johnson speak to that, but that is the crux of the issue. I think this particular set of questions recognizes the possibility that in OA, we may come up with products that diminish the amount of joint-space narrowing without appreciably diminishing the amount of clinical symptoms, and it's a quandary for the agency to judge whether or not the risks are worth the benefits of having the public at large be exposed to these agents, have the benefits of these agents, whether you see the glass is half full or half empty when you may never know the answer given that you've put it on the market now and can't test it, and so obviously there's a tension there between enough evidence and putting

something out on to the market soon enough so that you don't have to spend 10 years waiting for clinical outcomes and randomized studies, and I think that that -- maybe Kent can amplify on that.

DR. JOHNSON: Steve, there are really two separable concepts here, I think. The one you're addressing is the one that probably bothers all of us and always has and may always will, and that is, you know, what does really happen over 20 years? These are 20- or 30- or 40-year diseases. We're not going to get products if we require 20-year trials. So that's sort of aside. We have to think about it. It would be nice to have Phase IV. It would be nice to have good Phase IV, and that may or may not happen some day.

But the whole other issue is given that we've made some arbitrary decisions about durations of trials, and I think they're incredibly arbitrary, you know, six weeks to three months for non-steroidals, six months for a new agent for rheumatoid, you know, to a patient, that probably strikes them as pretty ridiculous, but given that that's done, you know, the accelerated approval dimension is simply a way to speed up something that looks promising, get it out there before you even show the clinical benefit, and when the epidemiology and the interventional trials have been done and knock you over like in blood pressure,

let's say, or cholesterol, then it's a moot point, and you get approved for the surrogate with no Phase IV validation, but that's not true at this point.

And in fact, if you look at the rheumatoid arthritis guidance document that came out, we engendered the exact same paradigm. We couldn't figure out how to describe what would be a major retardation in the rheumatoid x-ray, and if you go back and read it, you'll see all we said was, you know, one arm, you start with normal x-rays, normal hand x-rays, one arm, you get a lot of erosions, and in the treatment arm, you don't get any erosions, but we didn't specify beyond that.

But we're working through that. We went through the same logic for rheumatoid arthritis. No company has pursued that yet, and a lot of companies are interested in the joint-space narrowing utility right now.

DR. ABRAMSON: Dr. Moreland?

DR. MORELAND: I would just like to add that I think you will not get meaningful data in any Phase IV studies once you approve a drug. It's completely unethical to withhold these drugs, and the patients will ask for them, and they will be given them once they're on the market, and we will not be able to enroll patients in any Phase IV studies for placebo-control trials.

So I think we're talking about some issues here

that are not doable in the real world, and the question that I come back to is then are you really needing to make this an accelerated review? Why not do the right studies? If this is a big issue, and you're concerned about letting something out that's not going to be good enough, then don't let it out. Let's do these studies a little more carefully and maybe have a two-year placebo-control trial and not have the accelerated process.

DR. ABRAMSON: Dr. Witter?

DR. WITTER: But I think it still gets at the question of if we do a two-year trial or a three-year or a five-year or whatever, and if we don't get what we all feel comfortable with as a clinical benefit for that, I think the skeptics still would say, you know, why have we done this? Why have we bothered to do this?

If we've altered the structure but haven't shown any kind of benefit, then accelerated approval or not, why should we do it?

DR. ABRAMSON: Dr. Elashoff?

DR. ELASHOFF: I don't know if you can make quite such clear distinction in this area, but certainly in ulcer disease, there were studies for two separate indications. One was for acute healing and the other was maintenance.

I don't see why in this instance, you approve

the drug for anything when you approve it, why there can't be some distinction between whether it's known to work in the short term versus whether it's known to work in a longer term in terms of the kinds of indications that you approve it for.

DR. ABRAMSON: Dr. Dougados?

DR. DOUGADOS: I think I can't understand the question concerning the Phase IV trial. With regard to the previous discussion, we have in the past within the OMERACT or within the Osteoarthritis Research Society related to the lack of knowledge of the predictive validity of the changes in joint-space narrowing within the short-term period of time, and at this time, in '94-95, our strong recommendation was that if we are conducting clinical trial of one- or two- or three-years duration, placebo-control trial, please continue to follow these patients in order to evaluate the predictive validity, and as an example, the data I have presented.

So that was the first recommendation. The recommendation was only forecast on the evaluation of the predictive validity. Nothing to do with any registration, only to improve our knowledge. So if we consider this point, I can invest under Phase IV as you conduct a Phase III trial of one-, two-, three-years duration. You stop the clinical trial, but you continue to -- that is, you

have the answer in terms of structure, and whatever the subsequent treatment, you are evaluating the patient after five, 10, 15 years, and with regard to the treatment they have received during the first three years. That's the possibility of a design for your Phase IV trial. That will be the continuation of the Phase III. The continuation but without the treatment, only the follow-up.

The second possibility is to continue the Phase III trial that is the placebo-control trial and then to evaluate the symptoms of the requirement for total articular replacement, and the third possibility is to conduct an independent Phase IV trial. Those are our three possibilities.

DR. ABRAMSON: Dr. Witter?

DR. WITTER: Bill can correct me. Under accelerated approval, it's my understanding that the distinction between what Dr. Moreland's referring to in Phase IV, I think we're well aware, for example, that Phase IV commitments of an approved product do not necessarily always get completed, but under "accelerated approval," that is in fact a more or less requirement, and that if those studies aren't done in a timely fashion as Dr. Schwieterman has elucidated, the agency will have -- I mean, we're expecting to see those kinds of studies. They will be completed.

DR. JOHNSON: Let me try to address Larry's question because it's really quite fundamental.

I think in the end, it's sort of a philosophical question or a social question as to what the degree of evidence you want before you approve a drug. I mean, we mediate our perception of what Congress and society want, I suppose, and maybe that changed in the last 10 years, and maybe the AIDS epidemic is what made it change, I don't know, but it has succeeded in the past.

There have been drugs registered for surrogates that had ongoing clinical trials that didn't lose all their -- I mean, they continued in a blinded fashion and were completed and validated the supposition successfully.

Now, obviously if 10 years from now, joint-space narrowing is as bona fide as blood pressure, then it's totally unethical to not address it, but I guess it's a question of timing, and if historically we believe that, you know, we've got a window now, and we could do this ethically, then I think this is the strategy that we're trying to facilitate or allow anyway.

DR. MORELAND: My comment, I guess, is that if I'm a patient on that study, and at the end of the study, I want to know whether I improved or not, and if you tell me I improved, then I want that drug. So if it's that important, that we think these are going to be so good, and

they're going to be so accelerated then, that I think we owe it then to the patients to be more up-front, and they're going to want this, and so I don't think you can get into Phase IV and manipulating and following patients, saying nice knowing you for a year, you're out of here now, we'll tell you in 10 years whether this year of study really helped or not.

I think we're going to have to make a leap of faith. Do we really think that one particular agent that can inhibit one enzyme is going to be that important in a bunch of stuff in those joints that is not quite so simple?

DR. JOHNSON: So at approval time, you go through another informed consent with the patients, and maybe Bill knows the answer to this, whether or not that was actually formerly done with other accelerated approval scenarios?

You could fancy that you could do that, and patients would remain blinded to their treatments, which I don't think is unethical, but you could say here's your x-ray, here's what it may mean, here's what it may not mean, and you may get some dropouts as a consequence, but at least I think that would be preferable than only, for instance, dropping out the bad x-rays in the placebo arm because you may undermine the ability to ever get a conclusion, you know, an interpretable conclusion if you

did that.

DR. SCHWIETERMAN: Well, I'd just simply add that my experience has ranged from there has been very little difficulty getting the Phase IV study going because it's a relatively short-term study. There's been adequate sponsor funding for the study, and the endpoints are relatively near to the time that you depended on the surrogate.

Thinking about, for example, betaserine for the treatment of multiple sclerosis, the endpoints were one year or so away, and when accelerated approval was granted, it was with data that had been submitted six months -- I wasn't exactly primary on the team. So don't quote me on these numbers. But those patients had been enrolled during the time of the FDA review process. So there was little difficulty keeping those patients for several more months on the study to get to a particular endpoint.

And Flexamide for the treatment of rheumatoid arthritis was also given accelerated -- again, conditional -- approval. I mean, for the treatment of Crohn's disease. It's currently under review for rheumatoid arthritis.

The single-dose use for Crohn's disease in patients who had severe disease was so compelling that the agency felt that for that small subpopulation, you could approve this under accelerated approval, but there were

safety concerns and efficacy concerns about the broader patient population, and again there were short-term endpoints that could be used to follow how those patients were doing.

With products for the treatment of osteoarthritis, I think Dr. Moreland raises a very good point. I think it's going to be difficult in some of these patient populations to get anything approaching long-term follow-up for clinical outcomes for patients that believe that there's a product out there that works.

I guess, though, you have to consider the reverse alternative. Do we want to hold on to therapies in the agency for a large percent of the population, almost universally which get afflicted with osteoarthritis in their later ages, if we have something that is reducing joint-space narrowing by 50 percent, say, at a year or two years?

Do we want to wait till five years out and determine that in fact these things cause adverse clinical outcomes when there's some reason to believe that they might? And there, of course, it depends upon the equipoise and the amount of data you have on how severe the progression is on the ancillary/corollary benefits, but I would wager that there would be some cases where we would see perhaps startling degrees of changes in joint-space

narrowing where we would not want to deny the public access to these particular products, which again is sort of begging the question that Larry raised.

If we can't deny these things, how can we possibly do the trial?

DR. ABRAMSON: You may just be in a situation where you can't do the absolutely perfect study that you'd like to do, and that you might ask the question again, if we all agreed upon imaging techniques that showed retardation of structural change, you may begin to have to apply and think about this same criteria that we do for remittive agents for rheumatoid arthritis, where we have no better data to suggest that a retardation of x-ray changes in RA improves what happened to that patient 10 years down the line.

We make a reasonable presumption that it does, and I think one could argue at least that one could make a similarly reasonable presumption in osteoarthritis, if you had good imaging techniques that people felt were valid, and so the question is if you can't do the study that you would really love to do for five years, and you may not be able to do a good Phase IV study, why apply different criteria for these class of agents of structural modifiers where approval comes when the endpoint has been met, and then surveillance is applied the way surveillance is

applied for other kinds of medications that are released? Just to play devil's advocate on that side.

DR. SCHWIETERMAN: Well, the counter-argument to that is then you are risking, if you make no attempt, short of surveillance data, which is very difficult to infer efficacy data on, say, from a registry and so forth, it's very difficult if you have something that meets its structural outcome measure, yet is not studied in a way, in a rigorous enough way with which to measure efficacy, you're left with the question of whether in fact you've done the thing or not, and whether in fact you're exposing patients to a potentially-dangerous product, whether short term or over the long term.

I don't think there are any easy answers here, and I think that this is, despite it seeming somewhat confusing and chaotic, it's still helpful for the agency to think of people's thinking about this because obviously the devil comes in the details.

DR. ABRAMSON: Other comments? Dr. Yocum?

DR. YOCUM: I appreciate Larry's concern, but I think that the problem here is we've put in placebo, and we're dealing with this now in long-term follow-up studies with RA comparing to the recent release of accelerated review of certain products over the last few months, and we've looked for a control population for those studies as

well long term, and I think to say that Phase IV would be a placebo group may be a misnomer.

There are going to be a lot of people out there early in the phase of development of these products that are going to walk in and listen to the spiel about being on the products and saying no, I don't want to do that, but I will be followed as a control for this population.

So I think given the size of the osteoarthritis population out there and having been involved in Ken's study with, what is it, nine out of 10 are excluded for various reasons, there is a large group of controls that aren't placebo that make the self-choice of not being involved in protocols. So I think that we fool ourselves here by saying, oh, this has to be a placebo group.

Later on, once there are a ton of these products and nobody wants to be without one, that will be a problem, but early on, I don't think we're going to have trouble developing a control, not a placebo control.

DR. JOHNSON: Yes. I wanted to comment on Maxime's suggestions, but prior to that, fundamentally, you could have no placebo, and if you could still show a dose-response, that's also a possibility, if your structure approval range was broad enough that you could get two different doses, and you could sanction doing that ethically.

Maxime was suggesting that you continue to follow up even if they're not on the drug or if they dropped off placebo. I mean, if you've lost all your placebo, if you've lost all of your control, you will be able to epidemiologically make some statements about the predictability of the joint-space narrowing effects that you've seen, but you won't know how to compare it, except to other epidemiologic studies.

I guess what I think David is saying is important in that we may have a window right now with osteoarthritis because they're not throwing a co-therapy around, and hence any proposal like you're talking about wouldn't be as confounded by lots of other agents being tossed on board because there aren't any out there, and if it's food supplements, presumably everybody's already on those.

DR. WITTER: Can I just maybe clarify something?

DR. ABRAMSON: Yes.

DR. WITTER: Accelerated approval -- and I make the distinction between priority review, things like that -- I don't think accelerated approval necessarily means that the review will be a priority review, that we try and get it done, for example, in six months.

I think Bill Schwieterman has said it much

better, but I would just like to kind of ask Dr. Moreland a question. Do I understand what you're saying is if you were in the trial of one of these agents and working by whatever mechanism that it's working, and that is, you were aware that your joint space was improved, do I understand what you're saying is that you would probably opt to take that medication?

DR. MORELAND: I think the issue is that when a patient finishes a study, we can't just look at that one patient and say you improved or you didn't improve. The patients want to know what happened to this study, and whether the drug worked or not, and I think we, as clinical researchers -- it's a growing issue.

We need to be able to report back to that patient, yes, this drug did work versus placebo. They want to know whether they're going to get this. They spend a year or so many months or weeks in a study given out of their life. They want to know the results, and they want their physician to make an educated decision on whether they should stay on that drug forever.

So they don't necessarily want to know what happened to them, but they want to know whether the drug worked, and can they get it because they went into that study knowing that they wanted to get better or wanted to prevent something, and that's the issue that I have to face

day-to-day with patients. Can they continue to get the drug because they spent the time participating in that study, and I think the answer they want to know is, yes, it worked or didn't work, and if it worked even a little bit, and it's going to make it on the market, they want to be able to get it then themselves.

DR. WITTER: Well, let me just continue. Let's say that it works, this particular therapy. Would you recommend to your patients then that they continue on with that, and why?

DR. MORELAND: Well, I think it depends on the overall picture of what else is available, what else -- you know, the other profile of that particular agent, and so I think it's not just a simple yes or no, and did it work a little bit, but I think that's the spirit of where we're talking about drugs now that will inhibit perhaps very important radiographic or issues with regard to OA.

It's not going to be an outcome you can measure over a couple weeks or a couple months, and patients are worried about that as they're living longer. Can I take something to prevent this? Can I take something to prevent my hands from looking like my mother's or grandmother's? And so that's the issue that we have to come to grips with here as to -- and there's not going to be one of these. There's going to be five or six of these potentially who

might pass your hurdle, and my point is then what do we do with those five or six that pass your hurdle?

DR. WITTER: My question, too, is what are those hurdles?

DR. MORELAND: I would set the bar a little higher, and again as Paul and others have articulated here, I don't know whether we know any data as to where to set that bar, especially in the patient populations, but we perhaps need to make that step and put the bar up there.

DR. ABRAMSON: Dr. Hochberg?

DR. HOCHBERG: Well, what Dr. Witter may have been alluding to is let's say that this committee makes recommendations to the agency, and the agency decides to sit down with companies, and you come up with some scheme of whereby somebody will do a study which is not NIH-supported but which is industry-supported, and they're going to do a study for a period of time to try and register an agent as a structure-modifying drug for osteoarthritis.

Now, let's say at the end of that study, which may be one year or two years or whatever, they've reached the hurdle that they reduced joint-space narrowing or decrease in inter-bone distance by whatever you decide to set, 30 percent or 50 percent or whatever, and there hasn't been a worsening, and then there's an approval.

So the other question is, well, how long does somebody need to be on this medication? And you could ask the company to do a study whereby they could offer patients who completed the trial the opportunity to enroll in a Phase IV study, whereby they could be randomly allocated to continue on the medication or to go on placebo and see whether in fact they need to continue on the medication or not to halt the changes in inter-bone distance, whether or not in fact a shorter term of therapy is just as good as long-term therapy, and you then look at the placebo patients, and you enter them into a registry and look at clinically-important outcomes for which this was presumably a surrogate.

Some of them will, you know, go on the medication when it's available, others will not. Of those that don't, some of them may go into other trials, but you can then maintain registry data on them.

DR. WITTER: Then clarification of clinically-important outcomes.

DR. HOCHBERG: Whatever the agency decides is the clinically-important outcome for which this is a surrogate.

DR. JOHNSON: That might be a feasible approach. I mean, that's a double-blind respond to withdrawal analysis essentially, but you're doing it with

only your surrogate having responded. We'd probably have to think through that, but on the face of it, what do you think, Bill? You guys have used some double-blind withdrawals.

DR. SCHWIETERMAN: Yes. Actually, we have. Dr. Siegel helped Immunex design that pediatric study with that particular endpoint. It begs a lot of questions, though, because if you simply take the responders and then randomly withdraw them, the doctor's still left with the question I have this patient in my office, what do I do about long- versus short-term therapy?

In other words, this is a subgroup of the patients that have responded, and you get some efficacy data, but it really doesn't help you with the dosing.

DR. HOCHBERG: I didn't mean just the responders. I mean, the patients who completed the study on active medication.

DR. SCHWIETERMAN: The only problem -- again, it's very difficult to discuss this in the abstract. To the extent that the endpoint that you're trying to prevent is serious and results in a morbid condition, such as debilitating osteoarthritis, you have to ask the question, is it likely that you can effect a long- versus short-term beneficial outcome just as easily?

In other words, is it likely that a short-term

outcome treatment of therapy is just as good as a long-term therapy to prevent that, because if it's not likely, if you have reasons to doubt that a short-term course of therapy's going to prevent that, equipoise is lost, and therefore the trial becomes undoable.

DR. HOCHBERG: Well, I think my colleagues might agree that we don't know, you know. We don't know whether osteoarthritis has a continuous course or whether it has an episodic course, whether the deterioration is related to episodes of inflammation or not, and, you know, you could probably poll a dozen people in this room, and you'd clearly get two different responses. You might not get 12 different responses.

DR. SCHWIETERMAN: I'm sort of playing devil's advocate to take Dr. Moreland's position here because I would agree you don't know, but to the extent that there's a risk to the patient going off of a drug that has been demonstrated over a short period of time to prevent what was thought to be a surrogate for a meaningful and serious clinical outcome measure, to the extent to whether you can do this particular study or not, if they believe they can stop this drug, and they've maxed out on the effect because of the arguments that you make.

Again, it becomes less of an issue if you have less serious signs and symptoms. For example, this very

design that you're talking about is one that we've talked about with issues with Crohn's disease, where you can treat patients short term/long term, and they get signs and symptoms. Signs and symptoms are very bad. You get diarrhea. You get belly pain and so forth. But they're not the same as debilitating arthritis and so forth, and you can abruptly institute therapy to reverse those and have your answer within several months.

I don't know. If we had a product that was remarkably good at preventing joint-space narrowing, whether we could convince patients that there was just as equal likelihood that they were going to maintain their goodness state for the long term and thereby prevent long-term disability as -- if they stopped the product, the same thing would happen.

Frankly, if I were a patient, I'd be reluctant to stop something that was so remarkable.

DR. ABRAMSON: Mrs. Malone?

MS. MALONE: I agree with what was just said. If I were a responder, I wouldn't take the chance that I might be a placebo in the next phase.

DR. LIN: I just want to say that the trial design that Dr. Hochberg proposed certainly is one approach, but I don't know that in this particular disease, that there's a difference. One design may not be better

than the other one, meaning the one that continued everybody, because you might take so long that for whatever effect you accumulated in the drug group to win out, that will take longer to see a difference. That's Number 1.

Number 2. If you look at just the group of the drug group patients and randomize those, you're going to have a reduced number of patients, and that's going to make it a little harder to see a difference.

Also, I think what Jim and Dr. Moreland was discussing was that, you know, you prove that the joint-space narrowing worked at one year. Okay. Then the question is do you want the patient to get this thing? Well, you would give the patient the drug if you know that the drug has clinical benefit, but that's something we don't know, and that's why this extended study has to be done.

So it's almost -- it's necessary. I mean, I would call it Phase III/IV study, keeping it blind, and go as far as possible.

DR. ABRAMSON: We'll take one or two more comments before lunch. Dr. Dieppe?

DR. DIEPPE: Just I think the purest answer to this question, getting back to the question, is that you should go on forever without changing any of the parameters of the trial. That's obviously the purest answer.

What we're talking about is compromises and the level at which you're prepared to compromise. A word of caution, if I may, about some of the suggestions that are coming through, which is if they depend on patient willingness to participate in a continuing phase or reapproval from patients, then your generalizability is really going to go down.

We've got a generalizability problem any way because we're already dependent on those people who are willing to sign up to go into a trial anyway. What sort of people are those?

Now, if we then take it a phase further by people who are prepared to go on with placebo when they know the drug's been shown in others to do good and so on, we're getting down to generalizability problems of a major sort, I would think.

DR. ABRAMSON: One last comment. Dr. Brandt?

DR. BRANDT: I think the discussion, whether it's intended to or not, implies that this is a disease of a single joint right now that we're focusing on, a symptomatic joint.

Certainly for hip OA and knee OA, these idiopathic -- these tend to be bilateral diseases, not necessarily in temporal synchrony, but both joints are involved. It's increasingly clear that pathogenetic

mechanisms that are operative in the early stages, in the stage of initiation of OA, may be different from those that dominate and drive the process later on.

Also, there's a suggestion at least from animal models of OA that drugs that are effective in the stage of initiation may not be effective later on and vice versa.

Now, the doxy trial was structured to provide an opportunity to look at the effect of a drug on progression and on incidence, prophylaxis, in the same patient with very stringent recruitment requirements, but if we look at the universe of OA with more generalizability, it is not hard to find patients who are very lopsided, and the opportunity exists in clinical trials to look at both, and in a patient who has an index joint, a painful left knee, it may be possible to hold out the promise with a longer-term study, that this may have a protective effect against the development of disease, symptomatic or radiographic or both, in other joints, back to your signal joint issue.

DR. ABRAMSON: Thank you.

Any final comments from Dr. Johnson or Dr. Witter? No? Okay. We'll break now and reconvene at 1:30. Thank you.

(Whereupon, at 12:44 p.m., the meeting was recessed, to reconvene at 1:30 p.m.)

AFTERNOON SESSION (1:40 p.m.)

DR. ABRAMSON: Can folks begin to take their seats, please? I saw Dr. Johnson a second ago. Are Drs. Johnson and Witter here?

(No response.)

DR. ABRAMSON: Okay. I saw Kent.

Well, let's begin. We were on Question Number 3 and 4 under "Endpoints," and hopefully, now that we have a good sense, a better sense, of the goals of the agency, I think we can go to Question 4.

"Is a Phase IV design which specifies withdrawal of placebo patients who show severe joint-space narrowing -- that is, corruption of the negative control -- fatally flawed, and how can this be avoided?"

Any of our study design people want to begin the discussion of this?

DR. DOUGADOS: Could we have some clarification concerning the definition of "severe joint-space narrowing"?

DR. ABRAMSON: I guess Kent wrote this question. Let's assume we have a parameter of more than 50 percent.

DR. DOUGADOS: Because you have seen that I presented this morning, if you have degradation of more than 50 percent, the probability that the patient will

undergo surgery is 73 percent of the people.

In other words, you don't have to specify that the patient will withdraw from the study. He will withdraw from the study because of surgical intervention. But perhaps if severe referred to a .1 millimeter, that's completely different.

DR. WITTER: I think what is also being -- well, as you had mentioned, the idea that these are mandated withdrawals, and somehow we would -- the sponsor might, for example, actually look at data and have some kind of stopping rules based upon that, I think that's part of the heart of the question.

DR. ABRAMSON: So Dr. Johnson, the comment from Dr. Dougados was that people with severe joint-space narrowing, let's say more than 50 percent based on his data, would be forced to withdraw by the nature of the fact that a high percentage of them will need joint replacement.

DR. JOHNSON: Yes. Sorry. I missed the beginning. There's always going to be withdrawals for one reason or another. I think the question here is whether or not it's desirable to have the protocol prespecify a mandatory withdrawal for certain -- you know, arguing that you've affirmed some sort of hypothesis, and now you're going to actually believe it and act on it and change your protocol accordingly.

It goes back to the discussion we were having just before lunch, frankly. How do you continue a protocol with sort of partial information that you think, hope, wish, is going to translate clinically?

DR. ABRAMSON: Other panel members who want to address that question? Dr. Dieppe?

DR. DIEPPE: Well, I think the answer to the question is yes, it's fatally flawed.

DR. JOHNSON: What if the withdrawal were required to apply to all arms, not just the placebo arm, and so maintain the blind? Less flawed?

DR. DIEPPE: Well, less flawed, but dangerous as well, I think. Although we've seen Maxime's data, and that's fine, there nevertheless are some patients, particularly with knee and some with the hip as well, who can have virtually no joint space who do just fine.

So there are groups of patients with very severe joint damage who are doing fine in the long term. So I think it's kind of a dangerous policy to have a cut-off and say if they get there, we're going to take them out of the study.

DR. JOHNSON: But are you arguing that it's not ethically necessary to so design?

DR. DIEPPE: That's exactly what I'm arguing, yes. I don't think you need to do it. So it is fatally

flawed to do it, but I don't think you need to do it.

DR. DOUGADOS: And I do agree.

DR. ABRAMSON: Mrs. Malone?

MS. MALONE: But if your premise is that the joint-space narrowing is indicative of problems to come, where does that leave the patient? I mean, if you know that it is narrowing, you know?

DR. DIEPPE: Yes. I think the point is that we know statistically that you're more likely to get into problems if you've got a bad joint, but that's a statistical issue. There are still a number of individuals and quite a lot of individuals who do fine in spite of that narrowing.

So I would say that if you want to have a cut-off and say this is all too awful for people to go on, it's got to be on the basis of how the patient is, not what their x-ray looks like.

MS. MALONE: But wouldn't you have to inform the patient about that?

DR. DIEPPE: Sure.

MS. MALONE: And give him the choice?

DR. DIEPPE: Yes, sure. I mean, you could do that, but I still think that there's big dichotomy that we've not really sort of put out in the light of day, which is this dichotomy between, on the one point, talking about

structure and x-rays, and, on the other hand, how people are, and the poor correlation between those two things is at the heart of our difficulties.

My argument would be how people are is what we should be driven by.

DR. ABRAMSON: Dr. Brandt?

DR. BRANDT: There are data that say that Kellgren and Lawrence Grade 4 bone-on-bone in hip OA among males, only 50 percent are symptomatic.

DR. JOHNSON: I think one of the easier ways of seeing how it might in retrospect have been undesirable to drop out of the placebo arm or drop out of any arm if you have a terrible-looking x-ray was if the drug turns out to be seriously problematic, if the drug doesn't work or even more, if the drug is toxic, let's say, in the second year -- I mean we've got a premise, but we don't have it fully established is the problem, and what to do with this sort of partial knowledge.

MS. MALONE: Well, just as an aside, if you are with bone-on-bone, and it's not hurting, why? You know, why don't we go after that?

DR. DIEPPE: Hear, hear. That's exactly what we should be doing.

DR. BRANDT: That's exactly the real question. What's the difference between painful OA and painless OA

with identical radiography or levels of pathology?

DR. HOCHBERG: Less anxiety, less depression, better coping skills. It's a short answer. But let me get to another issue. Is this related to the Phase III or this is related into the Phase IV?

DR. JOHNSON: This is related to the transition into Phase IV.

DR. HOCHBERG: So let's say that this is a multi-year study. People are coming back annually to get radiographs done, and you notice a large decrease in somebody's joint space over time. This is akin to some of the osteoporosis studies, where there were people who were identified pre-hoc as rapid losers, who could be notified that they were rapid losers and could choose to either stay on coded medication or to go off coded medication and drop out of the study, not be unblinded, and then go on to whatever alternative therapy their physicians wanted to put them on, and for some protocols, you know, this was greater than three standard deviation change in one year, and that didn't fatally flaw the studies, at least the agency didn't think it fatally flawed the studies.

DR. JOHNSON: Well, I don't know that particular one, but that strikes me as a design that would be pretty problematic in trying to interpret the result.

It's a question of attribution, you know. If

you've lost the patients based on a partial affirmation of a hypothesis, you can always leave for bad symptoms, and in fact, you know, in certain more dangerous medical scenarios, those clauses are put specifically in the protocol, that you have to drop out for clinical deterioration, X, Y and Z, and we can put them in these protocols, too. That's not the issue.

The issue is what to do if you've got partial validation. I don't know when this study was done, but the osteoporosis thing is complicated because of the fluoride case where the surrogate failed.

DR. HOCHBERG: Well, I think the other thing about these kinds of compounds that we're talking about today is that everybody's going to be on co-therapy here. Everybody's going to be on symptomatic therapy, if they have symptoms of osteoarthritis.

So unless companies are going to recruit subjects from the general population and just look at radiographs and do a study of asymptomatic individuals with radiographic changes consistent with osteoarthritis, people are going to be recruited because they have painful or symptomatic osteoarthritis. They're all going to be on treatment.

So some people may, sure, drop out because their treatment isn't working, but more likely within the

protocol definition that's allowable, they're going to change therapies.

DR. JOHNSON: Well, does that make it analytically intractable? I don't think so. You just gather all that stuff and look at it, and that all comes under the co-therapy rubric. No?

DR. HOCHBERG: Yes. It doesn't get at Mrs. Malone's issue, but I would agree, it doesn't flaw the study.

DR. ABRAMSON: So Marc, you're saying it is not a flaw to remove the people with extensive joint-space narrowing?

DR. JOHNSON: I don't think either one of you two are saying that. In fact, Paul said the opposite. I don't know what Marc was saying.

DR. HOCHBERG: No, I'm not saying that it's not a flaw to remove them. Yes, but it should apply to both. If it's done, if people are going to be removed, then it should be applied while maintaining the blind.

DR. ELASHOFF: And you need to have a plan for how they're going to be dealt with in the analysis.

DR. HOCHBERG: Right.

DR. ABRAMSON: Okay. Any other comments on this question?

(No response.)

DR. ABRAMSON: We've addressed some of these later questions already, but should designs address other measures up front with face validity, the use of rescue medication?

I think there's been a sense of the committee that ordinary care, symptomatic care would proceed in these patients. The question's how standard that ordinary care can be.

Dr. Dougados?

DR. DOUGADOS: That's a real possibility to standardize that, and I'd like to look at the amount of, as an example, milligram of NSAIDs, and then you have to obtain an equivalent score because in the large multi-center trial, usually the people are taking different NSAIDs.

The other possibility, which there is a publication on that, to only to focus on the percentage of days during which a patient had to take any drug that is in its easiest way to analyze. Otherwise, it's very complicated. You need a diary. You need to calculate the amount of NSAIDs or the amount of analgesic, and if we are focusing on the concomitant drug therapy, it's complicated but physical, but if we want to add non-physical therapies, such as it has been emphasized this morning by Paul, it's much more complicated because you have a lot of

information. One concomitant therapy that's important is total articular replacement.

DR. ABRAMSON: Other comments about concomitant co-therapy?

DR. PUCINO: Yes. And I think as what Dr. Brandt did with his study, to control some things that may be affecting structure, glucosamine, things we don't know about, should be accounted for and controlled for.

DR. ABRAMSON: Dr. Brandt?

DR. BRANDT: There's another practical issue. In a large-scale trial, especially with patients recruited from the community, subjects were recruited from the community who may become patients during the interval of the study for pain who are cared for by local doctors. It becomes logistically unfeasible in many instances if you're recruiting and randomizing a 120 patients in this trial to provide their total care as a rheumatologist or a caregiver.

You rely on their local physicians. You communicate with them, but those are the people who regulate symptomatic therapy by and large, not the study coordinator or clinical PI.

DR. ABRAMSON: Ms. Malone?

MS. MALONE: Wouldn't the psychosocial that Marc was talking about enter into this, too? You know,

some people may not need any medications, and, you know, they may, for whatever reason, be able to cope better and have a higher threshold of pain.

DR. ABRAMSON: Dr. Dieppe?

DR. DIEPPE: Well, just responding to a couple of comments. I think it would be nice to take the approach that says we will try to avoid other things that are structure-modifying in a structure-modifying trial, but with respect, I don't think that's possible because everybody in the U.K. any way is using glucosamine and stuff because they can buy it over the counter, and many of the other things, including some of the physical measures, may well be structure-modifying, we just don't know.

So it's not practical to avoid all other things that might be structure-modifying, sadly.

DR. ABRAMSON: Okay. And Number 2, the use of patient global, including but not limited to non-signal joints. Ken, is that something in your study that's built into the study or is it primarily signal joint?

DR. BRANDT: We ask a global question overall how do you think your arthritis is doing.

DR. JOHNSON: Not your knee arthritis, your arthritis?

DR. BRANDT: Yes.

DR. JOHNSON: Yes. I think part of this is to

try to get a certain agreement that this question should be a global question. I think, you know, five or 10 or 20 years ago, it was as often phrased as how is your knee arthritis doing in a knee, is that right? Yes.

DR. BRANDT: Yes. The difference is that the broader question, the more open-ended question, allows for some input relative to side effects, and it's not only non-signal joints but overall side effects with regard to medications, too.

DR. ABRAMSON: But do you not do both? The patient's global assessment for the signal joint as well as their arthritis or is it limited to just --

DR. BRANDT: No.

DR. ABRAMSON: -- global?

DR. BRANDT: A true global. Yes. We're not asking the global question about the index joint.

DR. ABRAMSON: No? Okay.

DR. BRANDT: But we're quantifying with WOMACs and things of that sort for those joints.

DR. ABRAMSON: Dr. Dieppe?

DR. DIEPPE: I think a global should be included for reasons Ken states. But I would also point out that if one uses things like WOMAC, one is doing more than the index joint because, of course, generalized function depends on more things than one knee or one hip.

So one is capturing things over and above the index joints with those types of measures, and then if you do put it other so-called quality-of-life measures like SF36 we mentioned briefly, you're again capturing a much wider spectrum than the index joint.

If you don't include those things, I think it's even more crucial that you do include a global that's not an index joint global.

DR. ABRAMSON: Dr. Dougados?

DR. DOUGADOS: I think there is some misunderstanding concerning -- not misunderstanding but several interpretations of the global because it is one recommendation of a lot of international societies to evaluate the patient's global assessment, but in fact, there are three levels.

The first one, which is the global assessment concerning the health status of the patients. The second thing is the global assessment concerning the arthritis as a disease, and the third is the global assessment of the knee arthritis, and it's a pity that when you are looking at the publication, you never know what is beyond, and sometimes there are some trials which have been conducted with a global assessment related to the health status of the patient and sometimes related to the knee arthritis, and I assume we make the mistake at the OMERACT last

meeting not to go into the detail because I'm not sure that we are using the same overall assessment, but we will produce the same thing.

We will use the same line, patient's global assessment, yes, zero to 100, but beyond that, the wording was different because in the trial I am coordinating, it's the knee or the hip that is a signal joint. I have no experience of VAS of arthritis apart from the signal joint. That is, I have one investigator that is one of your physicians. He does a global assessment of arthritis different than the signal joint, but I have no experience in that.

DR. BRANDT: That's why the question should be explicitly stated in the publication.

DR. JOHNSON: And the protocol, too.

DR. DOUGADOS: Usually it's written in the protocol. Usually it's written in the protocol.

DR. JOHNSON: Yes. What was your first of the three -- I understood the focus on the knee and focus on arthritis. What was the first one?

DR. DOUGADOS: Health status.

DR. JOHNSON: Just global? Okay.

DR. DOUGADOS: Yes.

DR. ABRAMSON: The next question is "other assessments of structure, such as osteophytes and joint

instability." I mean, osteophytes are built in -- I'm sorry. Dr. Dougados?

DR. DOUGADOS: Two things. If you want to pick up the information concerning new affected joints, that is, to use x-ray to make a classification, yes or no, as to arthritis, it's seen that at least in the knee, based on the data provided by the inspector in U.K., it's in that the presence of osteophytes is better than the joint-space narrowing. That is, to classify in a population people with or without knee osteoarthritis.

If you want to monitor patients in order to see whether the disease is progressing, yes or no, we have some data that most of the potential articular variables -- that is, joint-space widths, osteophytes, subchondral sclerosis, subchondral cysts -- are able to change over time if you have a large sample size, but the most sensitive and probably the most relevant is joint-space width.

DR. ABRAMSON: Dr. Brandt?

DR. BRANDT: I wouldn't be so sure about osteophytes. You're right in what you said, for example, about Tim Spector, but there was not rigid positioning, and in some instances he had osteophytes on the first film that weren't there on the second film, which is not a reflection of resportion, I suspect, as much as differences in positioning between the two exams, and there wasn't a

really rigid measurement of joint-space narrowing like you do today in those studies.

So I don't think we can categorically say that joint space beats osteophytes. I think we should look at both.

DR. DOUGADOS: And there are also other possibilities. That is, the structure can be also related not only to the cartilage or the bone, but also to the synovial tissue, and with MRI or bone scan, as an example, if a drug is able to reduce the old scan seen within 3 months, is it a potential relevant drug?

Paul?

DR. DIEPPE: Well, unless the chairman insists, I don't know that I want to get into that particular thing, Maxime. But I do want to comment on the question.

I think that we certainly should be trying to get as much information about osteophytes as we can. I don't think there's any way we can use osteophytes to power a study as a primary endpoint, because its assessment is so very crude at the moment, the best we can do is sort of naught to 3 scoring on a visual assessment. But having said that, as we've already said today, osteophytes might be crucially important in pain and indeed in relation to stability of the joint. So I think not to try and get some assessment of osteophytes would be silly, and I think it

should be a requirement of all studies to at least get a naught to 3 scoring of osteophytes.

I think when it comes to joint stability, I would say don't bother with joint stability, because we can't measure it. It's quite clear that clinical assessments of joint stability are almost a complete waste of space. Our inter-observer studies suggest that we have absolutely no agreement between us about which joint is stable and which isn't. So I don't see much point in trying to do that. There are sophisticated ways of doing it with clever biomechanical devices, but I think that's a step too far, not justified at present.

DR. ABRAMSON: Yes, Dr. Dougados?

DR. DOUGADOS: I am wondering whether you will analyze a trial in which a drug is able to reduce the size of osteophytes when compared to placebo, is it good or bad for the patients.

DR. ABRAMSON: On that, let's go on to assistive devices. I think that the discussions there would be similar to the co-therapies, that they should be allowed, and usual care should be provided. I don't know if there's anyone who wants to add anything to that piece.

(No response.)

DR. ABRAMSON: A question about the duration of the study: "One year minimum, in principle, for