DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

ARTHRITIS ADVISORY COMMITTEE

Tuesday, July 30, 2002

8:00 a.m.

Holiday Inn Bethesda
Versailles I and II
8120 Wisconsin Avenue
Bethesda, Maryland

PARTICIPANTS

Gary S. Firestein, M.D., Chairman
Kathleen Reedy, R.D.H., M.S., Executive Secretary

MEMBERS

Jennifer Anderson, Ph.D.
Kenneth D. Brandt, M.D.
Leigh F. Callahan, Ph.D.
John J. Cush, M.D.
Ildy M. Katona, M.D., CAPT, MC, USN
Susan M. Manzi, M.D.
Wendy W. McBrair, R.N., M.S., C.H.E.S.
Yvonne S. Sherrer, M.D.

GUESTS

ARTHRITIS ADVISORY COMMITTEE

Steven B. Abramson, M.D.
Raymond A. Dionne, D.D.S., Ph.D.
Janet D. Elashoff, Ph.D.
Clifford J. Woolf, M.D.

ANESTHETIC AND LIFE SUPPORT ADVISORY COMMITTEE

Michael Ashburn, M.D., M.P.H.
Nathaniel P. Katz, M.D.
Mitchell B. Max, M.D.

NONPRESCRIPTION DRUGS ADVISORY COMMITTEE

Alastair Wood, M.D.

INDUSTRY REPRESENTATIVE

Charles H. McLeskey, M.D.

GUESTS

David Borenstein, M.D.
John T. Farrar, M.D. MSCE
Vibeke Strand, M.D.

C O N T E N T S

P R O C E E D I N G S

Call to Order and Introductions

DR. FIRESTEIN: Good morning, and welcome to the second day of the Arthritis Advisory Committee meeting. I am Gary Firestein still. I think because there may be some people here today that were not here before we can just go around the room again quickly with introductions since this represents a separate meeting. Then, we can have the meeting statement from Kathleen Reedy. Again, I am Gary Firestein.

DR. SHERRER: I am Yvonne Sherrer, rheumatologist.

DR. CUSH: Jack Cush, rheumatologist, Presbyterian Hospital, Dallas.

DR. CALLAHAN: Leigh Callahan, epidemiologist, University of North Carolina, Chapel Hill.

DR. WOOD: Alastair Wood, Vanderbilt.

MS. MCBRAIR: Wendy McBrair, nurse and health educator, consumer representative, with Virtua Health in New Jersey.

DR. WOOLF: Clifford Woolf, Harvard Medical School and Massachusetts General Hospital.

DR. DIONNE: I must have said something

offensive yesterday because they took my mike away--

[Laughter]

--but I am Ray Dionne. I am a clinical pharmacologist, from NIDCR.

DR. WITTER: Jim Witter, from FDA.

DR. GOLDKIND: Larry Goldkind, FDA.

DR. SIMON: Lee Simon, Division Director 550, FDA.

DR. MCLESKEY: Charlie McLeskey, from Abbott Labs, and serving as the industry representative.

DR. STRAND: Vibeke Strand, rheumatologist. I teach at Stanford and I am a consultant.

DR. BORENSTEIN: David Borenstein, rheumatologist, clinical professor, George Washington University.

DR. FARRAR: John Farrar, neurologist, Instant Pain Management at the University of Pennsylvania.

DR. ELASHOFF: Janet Elashoff, biostatistics, Cedars-Sinai and UCLA.

DR. ASHBURN: Michael Ashburn, anesthesiologist, University of Utah, Pain

Management Center.

DR. ANDERSON: Jennifer Anderson, statistician, Boston University Medical Center.

DR. KATZ: Nathaniel Katz. I am a neurologist from Boston.

DR. MANZI: Susan Manzi, rheumatologist, University of Pittsburgh.

DR. ABRAMSON: Steve Abramson, rheumatologist, NYU Hospital for Joint Diseases.

DR. KATONA: Ildy Katona, pediatric rheumatologist from the Uniformed Services University.

DR. BRANDT: Ken Brandt, rheumatologist, Indiana University.

MS. REEDY: Kathleen Reedy, Food and Drug Administration.

Meeting Statement

And, this is the meeting statement for the Arthritis Advisory Committee meeting of July 29th and 30th, 2002. It is the same one; you can sing along if you like.

The following announcement addresses the issue of conflict of interest with respect to this meeting and is made a part of the record to preclude even the appearance of such at this

meeting.

The Food and Drug Administration has approved general matters waivers for the following special government employees which permits them to participate in today's discussions: Gary Firestein, Kenneth Brandt, Ildy Katona, Yvonne Sherrer, Susan Manzi, Jennifer Anderson, John Cush, Alastair Wood, Nathaniel Katz, Michael Ashburn, Janet Elashoff, Mitchell Max, Raymond Dionne, Steven Abramson.

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

In addition, Leigh Callahan, Frank Davidoff and Wendy McBrair do not have any current financial interests in pharmaceutical companies, therefore, they do not require a waiver to participate in today's discussions.

We would like to note for the record that Ms. McBrair's employer's interests in two drug companies are exempt under 2640.203(g).

The topics of today's meeting are issues of broad applicability. Unlike issues before a committee in which a particular product is

discussed, issues of broad applicability involve many industrial sponsors and academic institutions.

The committee participants have been screened for their financial interests as they may apply to the general topics at hand. Because general topics impact so many institutions, it is not prudent to recite all potential conflicts of interest as they apply to each member, consultant and guest.

FDA acknowledges that there may be potential conflicts of interest, but because of the general nature of the discussion before the committee these potential conflicts are mitigated.

We will also like to note that Dr. Charles McLeskey is participating in today's meeting as a non-voting industry representative. As such, he has not been screened for conflicts of interest.

In the event that the discussions involve any other products or firms not already on the agenda for which FDA participants have a financial interest, the participants' involvement and their exclusion will be noted for the record.

With respect to all other participants, we ask in the interest of fairness that they address any current or previous financial involvement with

any firm whose product they wish to comment upon.

DR. FIRESTEIN: Thank you very much. Now Lee Simon will welcome everybody again.

Welcome

DR. SIMON: I think that yesterday was an intriguing day for the committee members and I think certainly for us, over here at the agency. Again, I would like to thank you all for making the effort to come and participate even for the second day. I am even more impressed--everybody is still here and suffering through the heat wave we are having, although I am told it is not so much the heat wave; it is the expectation of Washington.

I would like to make mention of two things. One is that, again, this is a combination committee from 170, OTC and the Arthritis Committee. So, there are members from everywhere and I think it is very important for us to have a mixture of people commenting on these particular issues.

Secondly, as we had a meeting with the NIH and the FDA in March, we are proposing to have another meeting in some months on the issue of function, healthful quality of life and outcomes in pain, both acute and chronic. Ray Dionne and Jim

Witter are planning to apprise the wonderful experience we had previously, and I have been advised to inform everyone here in the audience of that. In fact, for the companies' benefits, the sponsors' benefits, this meeting will include your participation so that we can truly get opinions from all aspects of interest in this particular field. So, look forward to receiving invitations for this particular upcoming meeting sometime this winter.

Back to you, Gary.

DR. FIRESTEIN: Thank you. There will be some comments and discussion of our charge from Dr. Goldkind.

Comments, Charge

DR. GOLDKIND: Thank you. Again, I want to thank the committee members for taking time out of their schedules to spend two days with us.

Yesterday we dealt with a lot of conceptual issues primarily related to chronic pain. While there wasn't unanimity and closure on every point, the discussion we had was very helpful and, hopefully, enlightening for you as well. Today we will be shifting a little bit and talking primarily about acute pain, probably a little more

detailed in terms of study design and analysis, and we look forward to another fruitful and stimulating day.

DR. FIRESTEIN: Thank you. In addition, at some point during the day, probably during the 10:45 to 11:45 block, Lee has asked us to revisit some of the issues from yesterday strictly with regard to global pain indications, and we are going to end up going around the table and soliciting two-minute opinions. That goes for everybody, two-minute opinions on the two questions of how many indications might be required and how many domains do you think would be important. So, we will come back to that a little bit later on this morning.

The first speaker today is Jim Witter, who is going to talk about ABC metrics for acute pain.

ABC Metrics for Acute Pain

DR. WITTER: Good morning. Kathleen, I was looking for the bouncing ball before so I could follow you!

[Slide]

As Larry said, we are going to have a little bit of a shift today and we will start off talking about acute pain and, hopefully, go from

there. But we will be transitioning eventually back to chronic pain by the time the day is over.

[Slide]

In terms of acute pain, the argument I guess could go that what we need to do is to be as informative--again, we are discussing labels so we want to be as informative as possible about the information that goes into the label for something to treat acute pain. We had a discussion yesterday about acute pain versus treatment in an acute situation.

But what we have I think are really two scenarios. We have an outpatient setting and an inpatient setting where we might find ourselves in need of acute analgesics. For example, for outpatient settings, most of us have experienced I think minor injuries, such as a sports injury. Some of us have experienced dysmenorrhea. Hopefully, fewer of us have had a major injury such as a motor vehicle accident. Then, some of us actually volunteer to have surgery. Now, the analgesics that are applied in these situations are for the most part oral, not exclusively but mostly.

On the other hand, in an inpatient setting we again are looking at surgical settings and these

are of the non-elective and the elective type. What I have indicated here by the stars are some of the models or some of the clinical situations in which drugs have been studied and ultimately have been approved so this isn't that we are taking a major change of course here.

[Slide]

I would like to take a second and talk about the analgesic box. Some people would call it the analgesic black box. What I have tried to depict here is a pain relief curve. There is some event over here that causes one to have pain and you take a drug and, at some point in time then there is this concept known as onset of relief. The pain relief continues and goes to a certain amount. This has been described in the old 1992 guidance document and in the EMA document as the peak. We talked about it yesterday as the pain curve, the whole thing, and today I am now calling this the effect size. So, this pain relief goes up and lasts for a period of time and then it ends.

We should be able to, particularly in a single-dose experience, really define these parameters of onset-- what I am calling here effect size, and duration quite accurately if we do our

homework.

[Slide]

For acute pain the needs are to look at these concepts of the onset of meaningful pain relief, its duration, the effect size and we should establish these then in circumstances of acute inpatient and outpatient settings.

[Slide]

That leads us then to what we have termed the ABC's of acute pain metrics, that, in fact, you may not be able to accomplish all of these tasks in one trial and you may need to break this up. So, that is what we have done.

The A component is really getting at the concept of onset of meaningful pain relief. What we need to do is, to the best of our ability, establish this time very accurately. This onset should occur more frequently in drug versus placebo patients. It should be established in a variety of outpatient and inpatient settings, as I have described. And, this is really the single-dose experience.

[Slide]

I have depicted here a pain by time curve, a little bit different than the other presentation,

the other slide. We have pain intensity which is decreasing in a general sense. I have depicted two patients here, patient 1 and patient 2 and at some point along this curve these patients will let us know that they have established the onset of meaningful pain relief.

This is something that is not necessarily the same for everybody. So, I think what we need to do is make sure that while we are measuring pain intensity we also, particularly in the beginning, are measuring pain relief so we know how these two correlated because this is really a patient-derived outcome.

[Slide]

If we take an individual responder approach to this situation and this would seem to make sense--process analytical technology for an analgesic and for pain because pain is such an individual experience. So, the individual responder approach then focuses on a single person, not the group. It allows efficacy assessment to be very individualized, which we will be talking about later as well. It has the potential of eliminating imputation. We talked yesterday about forward filing of diaries. Michael Hufford talked to us

about that, and we all thought that was almost comical. We heard from Dr. Lu about last observation carried forward and other metrics to complete data. But if we can eliminate this, I think we all agree it would be better.

[Slide]

An individual response then for acute pain in terms of onset and duration for a single dose--I think the argument could be made that pain intensity should be measured throughout the entire trial. This includes not only the beginning but also at the end, when a patient either rescues or is censored, so we understand what is going on throughout the trial. Pain relief probably should be measured at least early to establish meaningful pain relief. If we do this properly, we should be able to really capture 100 percent of information on the patient's response to the analgesic, particularly during the single-dose experience.

[Slide]

What I have tried to do here is give us some idea of what I guess might be meant by the effect size. I have drawn some theoretical lines here. This says threshold; this says complete response. What I have depicted is the placebo drug

which crosses this threshold and goes to a certain point. We then have a drug which crosses the same threshold but goes beyond where the placebo response was and ends here. This is the concept of complete pain relief which is not happening, obviously, in this case.

But can we say then that the difference between the two blue lines here is really what we mean by the effect size? In fact, the difference between this line and this line is what we mean by that concept of a minimally clinically important difference. This is what we are searching for really because that is the difference from placebo. Can we, in fact, then really quantitate this response in a meaningful way?

[Slide]

The B of the ABC really refers to duration. What it is attempting to do in these series of studies would be to define the dosing interval, again, based on clinical data once more from outpatient and inpatient settings. So, here we are talking about the day 1 experience but it is the multiple-dose experience on day 1 if that is applicable for this particular drug.

We would then need to factor into these

metrics the concept of rescue in an outpatient setting or the use of concomitant medication such as opioids in an inpatient setting.

[Slide]

The C component is really meant to give us an idea of the minimally effective dose, and that is important because, you recall, yesterday one of the things that we discussed was our concern about carrying forward with analgesics, particularly analgesics studied in an acute setting, where the doses may be different than the doses that are carried forward in a more chronic setting.

If we have compounds which are not always going to be applicable and utilized, for example something like NSAIDs which we know are going to be used for the most part for something like OA, but, if we have medicines that have a very narrow therapeutic window but are really intended for an acute setting, we want to be sure that if they are used in what really would be off-label use that we have the lowest effective dose to be used in that situation. So, that is what the C portion of the studies are really intended to do.

Again, this is not intended to really inform chronic use. If there is a reason that

these compounds can be used in a chronic setting, we would encourage sponsors to do those studies and go for the indication. Again, establishing this in two settings, outpatient and inpatient, and this is a multiple dose over several days and the metrics may need to change in the sense of what we are interested in, as Dr. Lu had talked about yesterday, the area under the curve versus the onset peak duration mentality.

Once again, we are trying to establish the safety and efficacy here and we begin on day 2. So, day 1 in this particular series of studies is a time frame where we wouldn't have to be looking at any components of efficacy. These patients could take basically anything that they wanted. The randomization would then begin on day 2. So, what we are most interested in is from day 2, day 3 and on.

[Slide]

Acute pain has special issues with it, some of which we talked about yesterday. Pain is not equal in intensity or duration in various settings. For example, the pain after a dental extraction is not necessarily the same as after having bypass surgery, although maybe Leigh might

disagree. Pain does tend to improve with time. That is something we discussed yesterday. We will hear more today from Dr. Bashaw, but PK estimates in clinical results may really describe different aspects of pain relief in that PK may be more informative about early onset, for example, and may also then inform us about safety later. What I mean here is that if we have a compound that supposedly has a short half-life but in fact hangs around, for whatever reason, for days, and days, and days and has a very narrow therapeutic window, if the pain scores suggest that needs to be dosed more than once day we may have an issue of toxicity. In fact, we are faced with such issues.

[Slide]

So, the label in an acute pain setting should be as informative as possible and should contain information regarding onset, duration and minimally effective dose from two clinical settings, outpatient and inpatient.

[Slide]

If we cast in stone, so to speak, these concepts of acute and chronic, and if this were a river of pain I guess we are concerned about the bridging that needs to be done here because it may

be a time, as has been argued, that there is a transition from acute to chronic, wind up, plasticity, those types of issues. So, we should be paying attention to this interval between here and not lose sight of it.

If studies are conducted properly we may be able to support meaningful labeling claims for safety. We may, in fact, be able to get something for chronic pain if the studies would be supportive to push in that area, and we would encourage that if it makes sense. Or, this may also be informative for mechanistic claims that we talked about yesterday. So, it may be that this is the perfect time to be studying for some of these mechanistic claims, this time interval.

[Slide]

Why all the concern? Why don't we just leave things the way they are? Things have been working okay. Here is a drug that is in the PDR. I have given it the designation of X just to anonymize it a bit. This is the clinical study section. This is the entire section. It says: "In single-dose studies of post surgical pain (abdominal, gynecological, orthopedic) 940 patients were studied at doses of one or two tablets. Drug

X produced greater efficacy than placebo" and I have left out a few words here just to try to maintain the blind, "no advantage was demonstrated for the two-tablet dose." So, this looks like one tablet is pretty effective.

[Slide]

Elsewhere in the label, under dosing and administration, it says that this is "indicated for the short-term (generally less than 10 days) management of acute pain."

[Slide]

"The recommended dose of drug X is one tablet every 4 to 6 hours, as necessary. Dosage should not exceed 5 tablets in a 24-hour period."

The question is how did the clinical trials inform this dosage and administration scheme? There seems to be a gap here. This is an approved compound.

[Slide]

Could we make the case then that some of the ideal characteristics for a pain metric in this situation should be that it should be easy and understandable to patients and clinicians in the labeling and in clinical trials. It should be applicable across studies to facilitate IND

development and eventual NDA approval. It should define a clinically meaningful result so that it is a useful addition to our pain armamentarium. It should be valid in a variety of pain conditions, and it should be achievable with current meds, but also we need to think about having some kind of a tiered structure, which we have been talking about, so that we can really define and acknowledge important differences in drugs as they are developed.

[Slide]

Taking a responder analysis plan into a pain setting, it has the potential to characterize pain, as I have said, at an individual level in both acute and chronic situations, and Dr. Strand will be talking about the chronic situation later.

This may then be useful to allow a comparison of relative efficacy. This is against placebo or standard of care, not between drugs, in clinical trials in acute pain and in chronic pain.

[Slide]

If the hypothesis is correct, if it is properly constructed and validated, a responder analysis could be a major advance in clinical analgesia because it is currently not used. Later

we will be having more discussion about the concept of outcomes and domains, but I will discuss them here too. I think what we can say at this point is that if we can come to an agreement on outcomes or domains, we can do that even if we don't necessarily have the instruments because we can develop the instruments later. But if we can agree on the domains, that is definitely a step forward.

[Slide]

Just to step back for a second and look at the responder analysis that we do have in the division, the ACR, American College of Rheumatology, 20 responder analysis, and this is for rheumatoid arthritis, and this is really in a lot of ways a symptomatic responder analysis. What you have then to be approved for the indication of the signs and symptoms of rheumatoid arthritis, if you are successful with this metric you can then be approved, assuming you are safe. So, what you have to do is have a 20 percent improvement in swollen and tender joint counts. Those are required endpoints for this particular analysis. Then you can have three of the five following, a patient or physician global, a pain score, a modified health assessment questionnaire or some kind of an acute

phase reactant.

As Lee had mentioned and we talked about yesterday, we had the NIH-FDA workshop back in March. At that meeting we had a discussion of the responder approaches and certain domains were discussed. These included pain, rescue medication, patient global, health-related quality of life, physical function/disease specific measures, economic organ damage concerns, the issue of suffering which you heard about from Dr. Verburg yesterday, and adverse events. These were discussed as possible domains to be in some kind of an analgesic responder approach.

[Slide]

For the discussion this morning, I have whittled these down to the following that we should maybe be considering if we want to take this tactic, pain, concomitant medications, rescue medications, patient global, health-related quality of life, physical function, adverse events. Those are the ones that maybe make the most sense in this particular situation.

[Slide]

Were we to take this approach, could we begin to think about fashioning a responder

analysis by looking at our studies, our A, B and C type studies, and thinking through what needs to be applied or characterized in those settings? For example, for pain intensity the argument would be that that should be in all these studies. Pain relief, maybe more so in the onset and dosing interval. It may not be as important in the multiple-day use settings. Patient global might apply in all the settings, and continuing along. So, we may be able to already begin to get a sense of what a responder analysis might look like in an acute pain setting.

[Slide]

Let's just take a hypothetical example. It might be a bit hard to see. It is an AR20/12. So, AR then would imply that analgesic relief has been established. With an NSAID type compound that has generally been within an hour, but that time frame isn't necessarily applicable, for example, were we to develop a compound that would treat neuropathic pain, something that occurs sporadically like trigeminal neuralgia. That might not be the right kind of a time frame but, in any event, AR20 would refer to percent pain relief over the standard of care/placebo, and 12 would refer to

the hours of relief.

[Slide]

So, let's take a hypothetical drug that has two forms. This comes in a 100 mg and 300 mg variety. This is what a future potential trial session might look like and it would describe in there then the A, B and C, how the onset dosing interval and lowest effective dose were actually established in outpatient and inpatient settings.

[Slide]

So, this drug at the 300 mg strength in the indication section may look something like this: Drug X is indicated for acute pain. It is described as AR90/24 so it is a pretty potent medicine; it lasts for 24 hours. See the details in "clinical trials" and daily use should not exceed five days. Again, what we are trying to establish here is that in acute setting with some of these medicines, they may not be able to safely be used in a more chronic setting.

[Slide]

With the 100 mg strength of this particular compound, it may look as follows. It is also indicated for acute pain. Here, it is described as an AR20/24, and it would say daily use

should end when the pain has resolved or can be managed in another way, getting at this idea that acute pain for the most part resolves.

[Slide]

Without further delay, I would like to introduce Dr. Goldkind, who will be talking to us more, along with Dr. Bashaw, about the uses of dose and dosing interval. Dr. Villalba will be talking about some of our experience with certain compounds in the division. Dr. Strand will be giving us some more thoughts about the responder analysis, particularly in a chronic pain setting. Then, our own Dr. Simon will wrap everything up for us later.

Estimates of Dosing Intervals

DR. GOLDKIND: Thank you, Jim. I want to highlight the extent to which our discussions and our talks today are really aimed at labeling information. A lot of Jim's talk and, hopefully, mine will really focus not only on minimum requirements for approval but actually what kind of data we should be collecting to inform the label.

[Slide]

I will be playing tag with Dr. Bashaw, who is the team leader that is affiliated with our division. He is in the Division of Pharmaceutical

Evaluation.

[Slide]

An ideal analgesic is one that would be once a day, 100 percent effective in 100 percent of patients without adverse effects. Unfortunately, most drugs available today don't meet those criteria. Most of the time we have multiple doses per day that are needed in the acute setting, suboptimal pain relief and dose-limiting toxicities.

[Slide]

Therefore, the majority of patients and I imagine everybody in this room as a patient, if not as a prescribing physician, has been faced with patients or oneself has had several critical questions to ask when their pain recurs or doesn't respond in the first place. "What do I do till the next dose? Do I change medications? Do I call and get a new prescription? Do I simply redose early? Do I take another drug concomitantly with unknown synergy or safety concerns?"

The reality is that there is no ideal dose interval in our current world, but the goal is to optimally characterize, particularly as I will be speaking of duration of drug effect, and have that

in labeling and be sure that that is not associated with toxicity that is unacceptable.

[Slide]

So, the question is how, in the real world, do we generate dose interval instructions? I will be using dose interval and dose duration somewhat interchangeably. The first step in drug development is pharmacokinetics and I will turn this over to Dr. Bashaw.

DR. BASHAW: I would like to thank the previous speakers, both Dr. Goldkind for the introduction and Dr. Witter, for their fine presentations, and also the fact that most of what I am going to speak of today, the groundwork has been laid yesterday in our discussions about chronic pain and pain metrics.

For the most part, as has been talked about already, PK/PD and analgesic response has been primarily geared towards onset. The dental pain model is certainly very good for that. As you go from no pain to almost instantaneous pain very quickly it is very reproducible for all those factors we have talked about. But there are some problems with its duration because eventually pain does resolve in that model in a very short period

of time relative to most chronic pain.

During my presentation I am going to briefly go over some data from a dental pain trial as it relates to onset and dose optimization, and contrast it to where we are going with chronic pain and also with duration metrics. However, because it is still early in the morning, or relatively early in the morning, I promise I will not take you through any model derivations or any model simulations because that is way beyond the scope of the time of the talk this morning.

[Slide]

As I said, we basically have very good single-dose metrics looking at blood level onset and pain relief. One can pretty much look at a successful development of many OTC analgesics and even prescription analgesics and see that we do have a very good handle on onset, and the next step is where do we go from there when we need a second dose, and how we get from it.

[Slide]

This is what one typically sees. In this particular case we have a dental pain trial where we are comparing three different doses of a nonsteroidal. Here we have what is calculated to

be a no effect dose; what was assessed to be a mid-range dose; and what was expected to be an antirheumatic dose but was put in the trial just to see what the performance would be for a new analgesic.

You can see this is where we would start off with pharmacokinetic data, concentration versus time. From this type of material one can get the standard pharmacokinetic analysis of varying the curve, Cmax, Tmax and those parameters which we normally work with.

In terms of making the next step, linking this to some kind of effect, analgesia being duration or whatever we are looking for, one has to make the next step as to how one combines this information with the dynamic response.

[Slide]

This is one representation I have. I tried to make it as simple as possible. Basically, what our theory is, is that we pretty much have optimized input rate. Input rate gets into the blood, gets into the plasma and then we have drug migrating into some effect site concentration that then exercises the effect.

The dynamic compartment is a theoretical

compartment. We tend to draw it as a separate box but in reality the effect site is subsumed within the central compartment within the blood and within the plasma. But for modeling purposes it is much easier to have this over here because it explains some of the things we see with the drug onset in terms of lag time, in terms of dose response issues.

Primarily what one needs to just remember from this slide is that effect site concentrations is what we are really trying to look at. However, we can't measure them directly. We can measure plasma blood levels, but we cannot measure the concentrations at the effect site. These are all theoretical and based on our simulations. However, we do know that the rate constant, if you model it this way, the Keo value, is equilibration between these two compartments. It is what is going to drive duration. It is what is going to drive the redosing issue because it is going to control time to accumulation at the effect site; time to onset; and also time for levels to go back in the plasma. So, that is really what we are trying to look at in terms of driving this situation.

[Slide]

Here is what we normally see. Again, we are taking our dental pain example. We have taken our concentrations and now plotted them against a dynamic effect. In this particular situation this PID score and placebo are corrected. Here is our no effect dose, some effect but not very much. Here is our mid-range dose which is getting a PID at maximum of about 1. Here is our antirheumatic dose which is getting up there but there is some lag time here.

This pretty much shows one of the problems you have when you try to direct correlations between concentrations and effect. You can see, for example right here with the mid-range dose, that we have concentrations of approximately 5 ng/ml and you get a PID change of only 0.2. Yet, up here at 6 hours you have the same drug, same dose and the same concentration but it has a PID change of 1.

What is going on there? How can you have the same concentration giving two different responses? Part of that is due to the fact, again, of the model. It is 6 hours in the dental pain model and pain is starting to resolve. So, even though your concentrations have dropped you are

seeing resolution of their pain relief because of other factors, which again shows the limitations of this model.

[Slide]

One of the things we do with this data in trying to develop a relationship is we try to collapse these responses. We call these hysteresis loops because or their curvolinear nature. This particular nonsteroidal is very typical of what you see, counter-clockwise hysteresis, as one sees here. This is basically due to one of three reasons: There is a significant time lag between drug entering the central compartment and going out to the theoretical effect site. Possibly also it would act on the metabolite if you were just following the parent and the activity is due to the metabolite. That is also going to give you a disconnect which is going to result in counter-clockwise hysteresis.

And, important for a situation with nonsteroidals, it is due to the fact that we are not having a direct effect here but a secondary effect due to the effects of arachidonic acid. Nonsteroidals, unlike opiates which work on mu opioid receptors, kappa receptors, etc. and have a

direct pain activity, nonsteroidals, of course, have to work through the arachidonic acid cascade and that is going to cause a lag time because it takes time first to use up those factors that have already been formed, and then when the drug wears off it takes time for the cascade to reestablish itself. This also results in that disconnect between concentration and effect, which is one of the problems we have in modeling this data.

[Slide]

But if one continues on with the same dental pain trial and you collapse the loops, this is what you can derive. You can derive a relationship, shown in this particular case using an Emax model, and you can make a response between dose and effect. You do see noise out here and this, again, is due to the duration issues. But one can see in this particular case that we do have effect of concentration. There is an Emax of about 1.2 PID units, which is about what you are going to see for maximum effect.

From a response like this, one could then go back and look at your doses, look at your dosage form and pick a dosage that would give you the efficacy you want, depending on how you define it.

Once you have a PK/PD relationship, you can look back and say you want to have a certain duration, a time above a certain EC50 or EC75. If you want what Dr. Witter was talking about, a 90 percent change or 75 percent change depending on what metric you are using, if you are using a quality of life metric or if you are using PID scores, or whatever, it is very analogous to how you go back and do this and look at time above for duration.

These are analogous to what is done in the surgical area where you use neuromuscular blockade and you have a train of 4 measurements, where you are looking at a pharmacological response in terms of muscle blockade and you must calculate your duration based on how long you want to have neuromuscular blockade, and a train of 4 is a way of doing it. It is very analogous to trying to look at duration of action issues with analgesia, except that we don't have as well defined a metric or observation.

[Slide]

As I said before, one of the primary reasons you have counter-clockwise hysteresis is, of course, the fact that one has this cascade of pro-inflammatory precursors and pro-pain precursors

that have to be used up in the modeling. The time it takes up for these precursors, both to ramp up in the case of the drug wearing off and to be consumed and onset, is what affects our hysteresis loops. It really is the modeling problem for duration.

For onset we have very good metrics. We have shown that and pretty much we have optimized drug delivery to deal with the onset. But what about duration? How can we deal with that in the drugs that don't have direct response?

[Slide]

We can model duration of action using indirect PK/PD models that allow for downstream activities. However, it requires, as I think has been reiterated before, an understanding of the underlying physiology; an understanding of the dynamics of the response; patient factors; and does require a large number of PK and PD observations across a number of doses.

With this kind of information together, understanding exactly whether or not it is, as Jim pointed out this morning, moderate or severe pain from a dental pain trial or from coronary-artery bypass graft pain, you have to understand the

underlying physiology of the pain. You have to understand the dynamics of response of the patient factors and how the patients are going to perceive their pain; how they are going to relate it back to you in terms of its intensity or their degree of pain relief. Then, from a calculational standpoint, you do have to have a large number of observations, both PK and PD, so that you can make predictions across a number of doses.

[Slide]

What one can get from an analysis such as this--this is some simulated data we worked on for an intravenous analgesic and what basically one can do when one has enough data. This is the probability of obtaining a certain PID score over time for a certain dose of the drug. You can do this for many different doses. What we see here is that if you are looking for a PID change of 1, we have a very good onset and we have maintenance of that PID score for at least an hour and a half. Right there is the last observation in this trial. For this trial, here, the probability of a PID score of 2 is about 0.5 and then it starts dropping off when you start getting out to 40, 45 minutes. PID score 3 is really not going to happen here.

But using simulations, using PK/PD and understanding the models one can, using indirect modeling, develop probabilities using a Monte Carlo simulation that can then be related back to duration of effect and the maintenance of effect over time. If one has enough data-- this is obviously for one particular dose level--one can take multiple doses, plot together and actually do three-dimensional response surface mapping and look at the effect of various factors, concentration, effect, time, duration, etc. and decide what is an optimal dose that can then be tested in clinical trials.

[Slide]

Before I hand it back to Dr. Goldkind, from a pharmacokinetic standpoint looking at exposure response analysis, you know, with opiates, because of their mechanism of action where they have direct binding to receptors, we have good assessments of onset and we can do pretty good work with duration because it is a direct receptor interaction situation. With nonsteroidals, the mechanism of action being indirect and they don't actually have pain relief themselves but work through other mediators, through a cascade effect,

we certainly can do onset. We have actually done a lot of work over the last couple of years optimizing drug delivery for onset.

Duration is more problematical, as we have said this morning. It is model dependent. It requires an understanding of the physiology. It requires an understanding and identification of relevant patient factors. Also, it requires certainly a good amount of data to work with because if you don't have the data your simulations and your work just won't have the power you want to have to make proper dosing selections.

With that, I will turn it back over to Dr. Goldkind.

DR. GOLDKIND: Thank you.

[Slide]

We now know that PK can take us so far in assessing dose duration, but only so far and the question is how do we add to that with clinical data? I will be talking about the endpoints that are used in adding value to PK data in assessing a dosing interval.

First I would like to go through the guidance that we have, both from the FDA as well as from EMEA. The 1992 guidance, in the section that

does deal with metrics for assessing the duration of analgesia, and I quote directly: Similar onset of analgesia, there are various approaches to defining the duration of analgesia. Examples include from the onset of study drug or the onset of analgesia until either intensity of pain returns to baseline; the patient indicates that analgesic effect is vanishing, which are similar; patient requests rescue, and the time to rescue is sometimes designated as TTR, can either be measured in the mean or the median; and the percent of patients who do not rescue during the specific interval. You can look at the converse, the number that do and the specific interval can be over a longer period than you anticipate a dose interval, or the dose interval you anticipate and end the study at that point.

[Slide]

The European Medicines Evaluation Agency's draft guidelines from 2001 state that a real effort should be made to obtain data on the best dose and interval regimen, time to onset of peak effect and duration of effect. The endpoints that are referenced a little bit further on in that document refer to duration of analgesia, which isn't a

metric per se but just reiterates that that issue needs to be dealt with, and time to rescue is mentioned as a metric.

[Slide]

I would like to go through the different metrics now and discuss them. The return to baseline pain metric, I believe, is a flawed one.

[Slide]

This graph, which is taken from real data but the specific drugs are not relevant, is a good example and reflective of what we see in I would say most curves for analgesics. The top two lines are both active drugs and the lower curve is placebo. As we all know, there is a substantial placebo effect. There is an onset for placebo as well as the active drugs. But what you see as you go out is that pain relief is pretty much steady going all the way out to 12 hours. Interestingly, the placebo response drops a little bit but nothing comes down to baseline. That is not uncommon in the studies that we see.

[Slide]

As Dr. Bashaw mentioned, acute pain resolves and that is just part of the model. So, you really rarely get a true return to baseline in

these studies. Therefore, this metric would give you a bias, extending the apparent dosing interval, if we were to use a return to baseline. In addition, during acute pain studies you typically have repeated questioning every 15 minutes, half hour, for the first short interval, and then every hour after going out variable periods of time. So, it is actually quite an artificial setting to collect data to begin with. I would imagine that as you ask patients what pain relief they have now compared to one hour ago, compared to two, three and four hours ago you really introduce a lot of bias and there is a lot of suggestibility. So, a return to baseline pain inherently is problematic. In fact, I think most pharmaceutical companies realize this, and this metric is rarely used in drug development, although it is mentioned in the guidance.

[Slide]

So, how do we generate dose interval instructions in clinical trials? Well, the first thing I will say is that true dose interval ranging studies, meaning to test out what you would get at fixed intervals, fixed doses rather than waiting for a sense of rescue or "I can't wait any longer"

are actually not done. Metrics primarily come from single-dose studies. There is some qualitative data that I will mention briefly later that does come from multiple dose studies but this is limited in amount and applicability.

[Slide]

Getting back to the other possible metrics from single-dose studies and, again, I want to reiterate that what these metrics describe are rescue, not optimal. Percent of patients who rescue during a study period is largely affected by the study design and the study execution.

What I mean by that in study design is quite fundamental. If you have a study that is explained to an investigator and a patient as a 12-hour study, let's say, and you tell them that if they need rescue to let you know, as they approach that 12-hour period they may well see the 12-hour mark as a threshold, as a success point, and simply hold out to ask for remedication. If it is a 24-hour period, that will affect how it is perceived. Likewise, a short study interval--if somebody knows that the study is going to be over in four hours, they may wait to that point.

Actually, the last hourly acute pain

measurement is kind of a flip side of the study duration. In most studies you will have hourly pain measurements up to a period of, let's say, 12 hours and then there will be a final pain measurement session at 24 hours if the study is designed that way, if the thought is that possibly it is a 24-hour drug. If it is a much shorter acting drug the last measurement may be at 12 hours, with a gap of these hourly measurements.

There are expectations that are transmitted to the patients through the very trial design that affects their behavior. We have actually seen this in studies, particularly the shorter intervals. A study that has hourly metrics, going out to four hours, with a follow-up later on, has a tremendous rescue rate right after that fourth hourly measurement. It is very profound when you see how the study design affects the patient responses.

In terms of the execution, simply the monitor behavior and how encouraging or discouraging the monitors are of rescue, whether it is called remedication or rescue, the very presence of a monitor--does the monitor walk around if there is more than one patient in the center? Do they

leave the room? Is the medication left on the table to take truly ad lib or do you have to come up and ask the monitor that may look like Nurse Ratchet or may look like an inviting personality?

[Slide]

The time to rescue varies also depending on the setting. Major surgery is different than minor surgery; is different than dysmenorrhea. I will actually show some case examples of this in a little bit. Whether you are measuring the time from the dose or the time from the onset of relief obviously changes the metric.

The statistic you use, whether you use the median or the mean--the median is obviously less susceptible to outliers and the mean will shift responses towards the shorter interval based on patients who simply don't respond to the analgesic to begin with.

[Slide]

I will be talking about this population for analysis a little bit more. Let me define things better so I don't confuse what I mean by responder and responder analysis that will be discussed later.

If you use the all-treated population to

analyze a dosing interval, then you are including patients who either rescued within an hour and who didn't rescue at all. This usually shifts the dosing interval towards the shorter time period, particularly in models of severe pain where there is a high rescue rate. So, we could call that either the all-treated group which does, as I say, include people who had no response; we could call it the ITT population.

The responders that I am referring to are those subjects who register some form of pain relief early on in the study, and there is variability, in fact, at that point as well. You can be defining a responder as somebody who had analgesia and, therefore, they are a valid subject to capture information on how long that analgesia that they obtained lasted. You could do it by time to onset of relief, and that can be broken down into either perceptible, meaningful, adequate or some prespecified either VAS or categorical improvement. So, you may want to say a patient doesn't really enter the analysis of duration of their drug effect if that drug effect didn't at least meet some minimal level. It could either be subjective or you can try and objectify it with,

let's say, a pain relief score of at least 1 or 1.5 on a scale of 4.

[Slide]

As I mentioned earlier, there is variability based on the clinical setting. What we have seen is not surprising. The percent of patients who rescue is highest in general surgery settings, whether it is orthopedic or gynecologic. Dental rescue rates tend to fall below surgery. Dysmenorrhea rates are very frequently very low, regardless of whether you are looking at 12 or 24 hours and almost regardless of the drug or placebo, and we will see that. The median time to rescue medication which in a sense is derived from the same database as the percent who rescue, obviously, then has the converse. Dysmenorrhea studies have the longest dosing interval based on time to remedication; dental, a little shorter; and surgery, shorter yet.

[Slide]

In summary, there is a lot of variability in the metrics that we use. At this point in time they are not well standardized. So, we see different analyses presented by different sponsors. The study design, the study conduct, which

statistic is used, what population is analyzed, the definition of relief, the setting and, actually I didn't discuss this earlier but I put it in the summary, even from trial to trial in the same model, roughly same study design has variability, as you would expect in nature.

[Slide]

Now I am going to go through some case studies. The first ones will deal with this issue of the population that is included for analysis. The stopwatch technique is very frequently used. What that means is that it can be either a single or a double stopwatch. The patient is given a stopwatch and when they feel that they have gotten perceptible, meaningful, adequate relief, they click that stopwatch. A two stopwatch technique attempts to differentiate perceptible from meaningful. So, the first stopwatch click is "I feel something is happening" but it may not be very meaningful for them. The second one is when "gee, this is significant for me."

[Slide]

In this dental pain study, median time to remedication and, again, the drug isn't really relevant but the half-life is worth noting because

we will talk later about how much there is or there is not correlation between PK and clinical results. Placebo I will call zero half-life. We could debate that. This is the all-comers or the ITT analysis. You can see that placebo has almost a 2.5- hour median time to remedication. A 2-hour drug has a 6-hour median time to remedication; and a 17-18-hour drug has a 9.5-hour median time.

When you only look at those who responded, based on the perceptible definition of response, you see that this stretches out. If you were to base a dosing interval instruction for a label on these data, you would have to ask yourself do I go with just onset, those who had onset? Just the ITT? Some kind of a gestalt approach between the two?

[Slide]

I am just going to show a slide demonstrating variability from study to study even in the same model. There is a second dental pain study added to this slide. Within study 1 and study 2, which really were conducted identically, there is some difference that you see in the two studies. Is that tremendous? Is it surprising? No, that is variability that you see, but if you

were interested in drug Y, you wouldn't really know whether to push this to Q8 hours. Should this go to Q8 hours? Should it go to Q12 hours? Then, if you are guided by the analysis of only those with onset, do you go to 12 to come up with some kind of a combo here, or do you go to the Q24-hour interval? I think that we would all agree that it is kind of difficult to know from these data what is the ideal dosing interval. For drug X, it is a 2-hour half-life. Is it a Q4-, 6- or 8-hour? For drug Y, is it Q8, Q12, Q24?

[Slide]

In summary, for dental pain studies we see that there is an effect of the population you are using for analysis. There is a limited relationship between PK and clinical data. The time to rescue and the percent who rescue within an interval are informative but not definitive. Then the question that, in a sense, we are asking ourselves, asking the committee for input, is would there be benefit in studying a multi-dose in the sense of at least a minimum of a second dose where you actually look at a fixed dosing interval to get an idea of whether, beyond the placebo effect, there actually is a pharmacodynamic effect of an

earlier dose compared to a longer dose that may be chosen based on convenience and perception of safety?

[Slide]

We will look briefly at dysmenorrhea. As I mentioned earlier, these are two studies. This is a 12-hour drug Z and a 17 to 18-hour drug Y. As you can see, the median time to remedication is very long even in placebo. The percent who rescue, and this was within 12 hours, you can see is quite low. Obviously, the greater than 24-hour median tells you that at 24 hours it remains very low.

What this slide tells us is that dysmenorrhea is not generalizable to other settings. I don't think we would want to apply these data to the label in a generic way. And, it tells us that dosing interval for dysmenorrhea is not going to be well guided by this.

[Slide]

Just looking briefly at postoperative models, and this is an orthopedic study begun day after surgery or when the patients came off patient controlled analgesia and when they reached a certain VAS of pain, I believe it was the threshold when patients where entered into the

study.

We have placebo, drug Z 12-hour half-life, drug Y 17 to 18-hour half-life. I only have the ITT population analysis for this study but you can see it is very short. It doesn't even resemble the other two models. The percent who rescue in 12 hours is extremely high in all groups. Again, if you were going to use this model to generalize to dysmenorrhea and dental, it would be problematic. We do see this across studies and across other major surgery models. Do we need a totally separate dosing structure for postop pain? Is drug Z a Q4 hour drug? Is it a Q6 hour drug? Is Y a Q4 or Q6?

[Slide]

As I mentioned, the surgical setting is quite different than the dental and dysmenorrhea. The question is how do we establish dose interval for postoperative pain and, again, if drugs Y or Z can't be safely given during that shorter interval, what do we do? Do we contraindicate it? Do we indicate it for postop pain but in conjunction with a rescue medication that should be available because we know that the interval will be short?

[Slide]

Now I will just briefly talk about the qualitative data we get for multi-dose studies to add to the single-dose study metrics I have discussed. Use of supplemental or rescue medication over a period of time is frequently collected. Patient global evaluation over subsequent days is frequently collected, as is average pain intensity scores over a period. These endpoints generally are not really sensitive and informative enough to give us information on a dosing interval.

[Slide]

Let's not forget risk/benefit. We could say take the drug every hour but that will have its problems. We are reminded of this in this "B.C." cartoon, "What's the strongest over-the-counter pain killer you got?" And, the answer is a mallet over the head. Is it effective? Yes. Is there going to be remedication at all? Probably no. But is this the ideal analgesic? Obviously not.

[Slide]

We need to balance safety and efficacy, and that is an issue that we need to directly address in labeling. Obviously, you want convenience. You want adequate pain relief,

optimal pain relief, but you have to balance safety and metrics, whiich particularly in the acute pain setting, for safety are usually not very informative. If you have a drug that has a very high toxicity during a short-term period, you don't have a drug. So, it is hard before marketing to really know how that will play out. If you make a drug a BID instead of once a day, you are not going to see in that safety database, even if you collect it for a week, substantial differences that you may see in safety after it is marketed. Increasing the dose may well increase efficacy but it also increases adverse effects.

[Slide]

I am just going to discuss a case study of attempts in labeling to optimize that information on risk and benefit. It is the tramadol label. In the clinical trial section it states that Ultram has been given in single doses of 50 mg, 100 mg, 150 mg and 200 mg in patients with pain. In the dosage and administration section it states that for patients with moderate to moderately severe pain, not requiring rapid onset of analgesic effect, the tolerability of Ultram can be improved with the following titration schedule, and it goes

on describing a titration schedule that has been studied, and describing in some detail the extent to which it spared some toxicities.

[Slide]

A little bit later in the dosage and administration section it states that for the subset of patients for whom rapid onset of analgesic effect is required and for whom the benefits outweigh the risks of discontinuation due to adverse events associated with the higher initial doses, Ultram 50-100 mg can be administered as needed for pain relief every 4-6 hours. There is a statement that clearly says not to exceed 400 mg per day.

[Slide]

So, we have a label that really attempts to put in all the different metrics and information available, and it really is a juggling act for the prescribing physician. This is an example, frankly, of what you would need to try to cull from any label. You need to ask yourself what is the best starting dose for my patient? Shall I give them a loading dose that is high, or are they going to tolerated it better if I start lower? What timing interval should I give them? That, to an

extent, is left to patients. There is nothing wrong in saying take it every 4-6 hours depending on how you need it. But then you have to deal with the maximum dose over a 24-hour period. You have kind of taken from Peter to pay Paul. If you want a high dose in the beginning you are going to have to lower it later. Of course, there is titration of dose which is frequently an issue with opioids particularly.

[Slide]

In conclusion, the duration of analgesia is guided by PK data. The return to baseline pain metric is not an adequate endpoint to assess dose interval. The clinical setting affects the apparent duration of analgesia and remedication use.

[Slide]

The analysis of time to remedication is dependent on what population you are analyzing, those who have some onset versus those who are enrolled in the study and may well not have onset. The percent who rescue is informative, but it doesn't distinctly and clearly define any optimal dosing interval. The current metrics, as I have described them with the limitations, are not

standardized.

[Slide]

Additional information on dosing interval is needed. More formal study of dosing schedules may further characterize optimal dosing intervals, and different acute pain settings may need to be addressed in labeling.

I do want to say at this point that, with the second point on this slide, we are kind of venturing into a new area here. We don't really know what those studies will tell us if we ask for them, and that is one of the questions for the group this morning, to discuss how valuable such studies might be. Thank you.

DR. FIRESTEIN: Thank you. The next speaker is Lourdes Villalba, on safety databases for acute analgesics.

Safety Databases for Acute Analgesics

DR. VILLALBA: I am a medical officer in the Division of Anti-inflammatory Analgesics Drug Products.

[Slide]

Throughout our presentations at this meeting, we have tried to emphasize how important it is to collect adequate data to write a label

that is informative to patients and physicians.

[Slide]

I am going to talk about the kind of safety databases that we would like to see. I think my talk actually was titled safety in acute analgesia trials, but I need to spend some time talking about chronic requirements. Actually, instead of chronic, this should be long-term use.

[Slide]

We do have some guidelines. We have the ICH, International Conference Harmonization guidelines that were published in 1995 and refer to the use of products intended for long-term in known life-threatening conditions. Long-term is defined as continuous or intermittent use for six months or more.

The minimum requirements are 300-600 patients for 6 months, and 100 patients for a year, and a total exposure of 1500 patients including single-dose and short-term multiple dose studies. These numbers are given as a minimum guidance, and exposure should be available at clinically relevant doses or doses intended for clinical use.

However, the same guidance has said that a larger N or longer-term safety databases may be

needed. That is in the case when there are specific safety concerns. For example, if during drug development in preclinical studies or early Phase I for some reason we may identify some specific event, or we may think that some adverse event may be more frequent with time and that the hazard rate will increase with time, in that case we may need larger and longer safety databases. Or, when there is need to make risk/benefit decisions such as in the case when a new drug has a tiny effect size and, therefore, even if an adverse event is not very frequent we need to quantitate how often that happens in order to make those decisions.

[Slide]

As I mentioned, the guidance says that exposure should be in doses intended for clinical use. However, one of the safety concerns that we do have, which applies to all analgesics, is the dose creep phenomenon. Dose creep is the use of medications at doses above the recommended dose. That means doses above the demonstrated doses that are effective and safe in clinical trials.

We do have an example of the dose creep phenomenon from the Celebrex NDA. In the

randomized controlled trials part of the NDA, Celebrex showed efficacy in osteoarthritis at the 100 mg BID dose and efficacy in rheumatoid arthritis at the 200 BID dose. There was no obvious efficacy advantage of higher doses of 200 mg and 400 mg respectively. Of course, they were also efficacious but there was no major advantage of higher doses.

In the open-label part of the development program patients were allowed to increase the dose up to 200 mg BID in the osteoarthritis study and 400 mg BID for the rheumatoid arthritis patients. Actually, it was shown that most patients, 70 percent of the patients increased the dose and most of them moved to a dose twice as high as the initial dose even though there was no evidence of worsening efficacy right before they increased the dose and there was no evidence of improvement in efficacy after they increased the dose. So, this is just an example and the good news is that there were no major safety concerns observed with these increases in dose.

[Slide]

Therefore, out of a summary regarding exposure requirements for long-term use, more than

fulfilling a minimum number, what we want to see is an adequate safety database that will address specific issues that may arise during drug development. We want to see minimum ICH guidelines at the highest labeled dose. We also want to see special populations addressed, particularly the elderly and the pediatric populations.

[Slide]

Now I am going to talk about exposure requirements in acute or short-term use. The approach that we have had in the division for the last several years is to require as much as if it were intended for chronic use. The reason for this approach is that we know, I think everybody is aware, that drugs are used for longer than approved. There is no analgesic that is going to be used only once. Even if the label states that the recommendation is for short-term, they are used for longer term.

I have two examples here. One is from the Vioxx database and the other is Duract, bromfenac sodium.

[Slide]

This slide was presented at the advisory committee meeting in February of last year so it is

a little outdated but it makes the point. Vioxx 50 mg was approved for the treatment of acute pain. It was recommended in the label to be used for five days. This dose is twice the dose approved for chronic use, the highest dose approved for chronic use in osteoarthritis and twice the dose approved for rheumatoid arthritis.

At that time, the total number of drug appearances was approximately 13 million. Of those, 5 percent were for the 50 mg strength. Of those, one-fifth were for more than 30 days. So, this is just to show you some numbers because with the next example, which is actually much more dramatic because of the public health issues that came with it, we do not have numbers or denominators.

We have also seen with Vioxx that there are some patients who used the 50 mg dose twice a day, that is, 100 mg a day. That actually is very unwise, I would say, because there are very limited data on the 100 mg dose in long-term exposure.

[Slide]

This is the next example. This is an unfortunate example but very enlightening for us, for the division and for the agency. Bromfenac was

a nonsteroidal anti-inflammatory drug approved in July, 1997. There was a voluntary withdrawal in June, 1998 due to reports of hepatic failure.

This is a very interesting example because the original development program was towards acute pain, dysmenorrhea and osteoarthritis and there were also some rheumatoid arthritis studies. The proposed dose in the original NDA was 25-50 mg every 6-8 hours up to 200 mg a day.

At filing, it was noted that there was insufficient exposure for the osteoarthritis indication. Therefore, the osteoarthritis indication was withdrawn but chronic safety data from the chronic studies was submitted.

[Slide]

I want to show you the size of the database which is actually a very good size if you look at total numbers. The total exposure was close to 2200, with 1000 patients exposed in the acute pain studies, close to 400 patients in the multiple dose, up to one week studies. There were also some dysmenorrhea studies of 250 patients and the chronic exposure was about 900 patients in osteoarthritis and rheumatoid arthritis. So, if you look at the total numbers these look very good.

[Slide]

However, if you break it out by dose and duration of exposure--this is the dose in milligrams a day and this is the duration in days of exposure, you see that the number of patients exposed to the 200 mg dose for a year or more were only 24. The bulk of the exposure was at doses below 150 mg.

At the safety update there were more patients, and when we get to the 900 patients exposed for more than three months--I do not have the breakout of these numbers but it was mentioned in the medical officer's review that there was sufficient exposure to support the 150 mg dose a day and, again, the dose was 25-50 mg up to three times a day.

[Slide]

I don't want to go into details but just to show you that this was a very good database in the sense that there were placebo control studies, active control studies up to one year with several comparators. They used fixed dose, as I said, 25-50 mg BID, TID and four times a day but in fixed dose, not in flexible dose. There was a good number of patients with OA and RA, and there also

was an open-label experience up to four years and that involved flexible dose, some of them up to 225 mg a day.

[Slide]

Therefore, the safety review showed that acute pain studies that were conducted at the 50 mg and 50 mg single doses, and also in short-term multiple dose studies conducted with the 25 mg and 50 mg a day dose, showed absolutely no safety concerns. There was some nausea, some vomiting, a little allergic reaction but there was not even mention of any liver effects.

[Slide]

However, the chronic studies showed a flag for hepatotoxicity. This is what the NDA review showed regarding liver function test elevations, 15 percent of patients had mild elevations, that is less than 3 times the upper limit of normal, and 2.8 percent had clinically significant elevations of LFTs, 3 times the upper limit of normal or higher. Of note, the NSAID template mentions that LFT elevations in clinical trials of NSAIDs are usually seen in 15 percent of patients. Therefore, the number of patients with mild elevations of LFTs was nothing outstanding. The clinically

significant elevation was higher than what the template says, which is 1 percent but, again, it was not something terribly dramatic here. This number is actually similar to what was observed in the diclofenac NDA.

The elevation of LFT particularly clinically significant events were dose related. They were observed at the 100 mg dose but most of the cases were at higher doses. Most of them were reversible after drug discontinuation. Some of them were reversible even without drug discontinuation. The majority occurred within the first 90 cays, but the important observation was that the earliest occurred around day 30.

[Slide]

Based on those observations, the drug was approved with warnings for risk of hepatic effects. Short-term use for pain should be less than 10 days and, because of the risk of hepatotoxicity, if longer therapy is needed, LFTs should be monitored after 4 weeks. So, we think it was pretty clear that there was some concern with liver toxicity here. In addition, the maximum daily dose would be limited to 150 mg a day, and there was removal of any reference to treatment of osteoarthritis,

chronic pain and dysmenorrhea.

[Slide]

Within months the agency started to receive postmarketing reports of liver toxicity, including hepatic failure, need for liver transplantation and death. Most of the reports were at doses within the labeled dose, but most of them were exposed for longer than 10 days. The majority was for 2-8 months, and some of them were exposed for a couple of years.

We have this unfortunate example, but I think that reflects something that everybody knows, which is that drugs are used for longer than initially intended. As was discussed yesterday, if a drug is approved for acute use but somebody thinks that it may work for chronic pain physicians are going to use it.

[Slide]

In summary, short-term safety studies are certainly insufficient to address safety concerns that may come up with some patients who will be using the drug for longer than intended.

Drug development for acute pain drugs should address the potential safety concerns of dose creep, use for longer than the intended, and

potential for abuse which is another whole issue.

We propose that for a short-term indication, unless there is a contraindication based on safety, formal efficacy studies should be done in a chronic setting. I think this is the new concept that we would like your opinion on. We are not saying that off-label use needs to be addressed for every indication because that is impossible, but for a drug that belongs to a class that is used for a chronic indication it is very reasonable to ask for some efficacy studies. If it doesn't work, if it is not efficacious in the chronic indication, then we can put that in the label, that this doesn't work for chronic pain; do not use it. So, we think that this would be a way to address the possibility of off-label use and also allow us to do a better risk/benefit assessment. That is the end.

DR. FIRESTEIN: Could I ask a quick question? Do you think that that final recommendation would essentially nullify yesterday's discussion about having separate acute and chronic indications? I mean, if for an acute indication you are going to require formal chronic safety and efficacy what is the value then of

having separate tracks?

DR. VILLALBA: Well, we are not going to require replication in three different models for the chronic indication. What we want is at least to have some efficacy studies. For example, for a new NSAID or a COX-2 inhibitor, if someone would come with only the acute pain indication, then we would ask for osteoarthritis studies to see if that worked in the chronic setting. That would provide also better safety data because safety data collected in an open-label way is not the same as safety data collected in a controlled way, with placebo control and active control studies. But we actually would like to hear your opinion. Thank you.

DR. FIRESTEIN: Are there any other comments or questions from the group?

DR. MAX: I have some questions regarding the dosing interval. I think a lot depends upon what you want to tell people about. My question is has the FDA studied what percentage of patients whom you are trying to inform who are taking acute analgesics take two doses total versus three doses or four doses? Because if you mostly want to tell people about the second dose, single-dose duration

is enough. If there is a large number of people who take three doses, the second dose is important, and so on.

DR. GOLDKIND: That question will really depend on what studies show the dosing interval should be. There may well be off-label usage TID for a BID drug, but if the best studies have identified a twice a day regimen, actually PK and some Phase II clinical studies should give you an idea of the ball park. I mean, we don't have examples of every two-day drugs or drugs that are taken very infrequently. I think, as you pointed out, you need to at least get data on doses going out beyond the first interval that you would be prescribing in terms of usage data on how many patients go beyond the frequency advised. We don't have that.

DR. MAX: Yes, my question is have you studied general use of analgesics for acute pain and how many people just have one day treatments or one dose treatments, or two day, three day treatments?

DR. GOLDKIND: We don't have that, no. In clinical studies it is hard to get a model that will get you out multiple days. So, I think that

answers the question to some extent. Most people only take acute analgesics in the postoperative setting or acute injury setting for several days on a regular basis.

DR. MAX: But do you understand what I am referring to?

DR. GOLDKIND: If I do understand, we don't have usage data to tell us how many days patients take acute analgesics for most indications. I don't know if that is available. I don't know if IMS data could give us that.

DR. FARRAR: As somebody who has focused primarily on chronic pain as an area of study, I would admit to this being the first time that I have sort of seen the full scope of the approval process for acute pain. I commend the FDA for reexamining the entirety of the approval process because I think there are a clearly a number of issues that can be addressed that aren't currently being addressed, some of which were being hinted at by Dr. Max.

One of the things that strikes me is that I have never, ever seen a drug that is used as a single dose, ever. It may be tested that way; it may be used that way perhaps in a hospital setting,

but if it is over-the-counter it just doesn't exist. Therefore, I think it is probably necessary to study certainly the effect of several doses over a period of time. I think that that would clearly generate a completely different set of data perhaps.

The second issue that I will just raise, and I am just raising all of these and I think they would need discussion at length in a different setting, but the second issue relates to the safety data. Dr. Goldkind showed very nicely sort of the need to look at risk/benefit ratios. It seems to me that it doesn't make obvious sense to look for safety day in use over six months and not look at least in some way at efficacy data over the same period in terms of just thinking about how a medicine is going to be used in terms of the general public.

What that raises is really the last point that I want to make, which is that we know that these drugs are going to be used in a variety of different ways by different patients and different physicians. And, I think it is imperative that we look at the way in which the drug is going to be used and use that information to guide us in terms

of both the safety and the efficacy data that we would want prior to or following approval.

There are two points that were made in the last presentation which I think really speak to this. With the Celebrex example, the fact that 70 percent of people increased their dose when allowed to do so tells you two things. It tells you, one, that that is the way it is going to be used. It also tells you that even though the study was not large enough to show that a larger dose provided better efficacy, or that there was some development of--I don't want to call it tolerance but getting used to the medicine, whatever you want to call that, that over time an increased dose was more beneficial. The patients were telling you that. The patients said when given the option I will take this medicine at a higher dose because it works, number one and, number two, doesn't cause acute side effects. That really is telling and indicates that there needs to be at least some approach to the concept of if given free access to the medication, if it was placed at the bedside so the patient can take it without asking the monitor, be that person nice or not nice, then they will use it in the way in which they would probably use it at

home and that would perhaps dictate the way in which a study could be conducted.

The very last thing that I would like to point out is that we need to keep in mind with all of the PK data, all of the time to effect data, all of the time to return to baseline although I think I agree that that is a lousy measure, time to remedication, those are all mean values. What a mean value indicates is that there are 50 percent of the people who did either better or worse. I don't think that 50 percent is the number we are actually targeting in terms of what a reasonable dosing schedule would be. I certainly would never treat my patients and allow 50 percent of them to suffer for an hour or two before I gave them a second dose.

I think that needs to be dictated very carefully by the risk/benefit or the minimum amount that they can take to be effective and the maximum amount they can take and still be safe.

DR. FIRESTEIN: I think actually you are referring to median, not mean. Actually, the points that you raise bring us to the first point of discussion. I think based on what we have heard and our own clinical experience, it is reasonable

to expect not single-dose studies but at least multi-dose studies involving a variety of metrics. I would like to open this for discussion with regard to what sorts of metrics people might feel would be appropriate. Susan?

Discussion Points # 1, 2 and 3

DR. MANZI: I just wanted to make one other comment first. I agree that I think the purpose of clinical trials is to accurately simulate clinical practice. As I was listening to these talks, I said I can't even imagine where you would use single-dose analgesic even in the most acute situations. So, I would agree with multiple dosing.

The only other point, and I guess this is the epidemiologist's hat that I wear, is that when you are looking at how to figure out dosing, you really learn a lot from the outliers. It is the people who extend beyond the bell curve where you get the most information. My point would be that if you look at time to rescue, you shouldn't exclude the non-responders in that because in clinical practice we can't predict who those non-responders are going to be and when they are going to need some additional dosing. I think most

people don't take a drug and say "it didn't work; I'm not going to try it for another dose."

So, my point is that I would assume the most narrow time based on the outliers for time to redosing and test safety of that in that setting.

DR. FIRESTEIN: Clifford?

DR. WOOLF: To come back to the issue of onset and duration, Dr. Witter's presentation, the context of when even a single drug is given, whether it is given pre- or postoperatively may profoundly change both of those metrics.

DR. FIRESTEIN: Coming back to the question of what the appropriate metrics might be, a series of possibilities were raised, and I can't remember in which presentation it was but is the gold standard for an acute pain medication going to be quality of life, or is it simply pain? Somebody?

MS. MCBRAIR: I would go for pain relief. I don't think we are worried as much in the short term about the quality of life, especially for a post surgical patient. They are going to be, hopefully, in a hospital setting and well monitored, and they need pain relief and we would not hold it back from them.

DR. CUSH: I would also say that when looking at the metrics you should rely upon, obviously, pain is where we are going to go. Unlike other diseases where our metrics are maybe multivariate where we are going for a disease response, here we are looking for a symptom response across many different diseases, and having a multivariate definition of response might be very difficult to arrive at, as we discussed yesterday. But if we had an acceptable measure of pain relief that was universally agreed upon, we could go for the variables that Jim was looking for. For instance, if you defined an acceptable response of 50 percent, pain relief of 50 percent, you could then define the time of response and the percentage of patients actually receiving that response in a placebo population and in an active treatment population and then also maybe even define the duration of response with a PR 50, or something along those lines.

DR. FARRAR: I think the point about quality of life as a measure in an acute pain process brings up an important point, which is that the quality of life is defined differently in different circumstances. I would argue that

adequate pain relief postoperatively is, in fact, a very good measure of a postoperative six-hour period of quality of life.

But I think ultimately that pain is the primary outcome. What I would like to point out though is that it is not a single measure of pain that is paramount. Certainly, in treating postoperative patients, clinicians are aware that the onset of action is vital to the control of pain and you certainly would not give a medication to a postop patient who is writhing in pain a drug that would take two hours.

So, the onset of action is of extreme importance, as well as the duration of action only inasmuch as it dictates dosing. The duration by itself--you know, a long-acting medication may well be of benefit but if you have a short-acting medication, as we know, in terms of intravenous use of various short-acting opioids, they can be very effective and the short-actedness can be overcome with either an infusion or multiple dosing.

So, I would argue that there needs to be pain measurement as a primary outcome with at least two issues. One is the onset of action and then the duration of action as it dictates the use of

the drug.

DR. KATZ: Just to continue the discussion of appropriate metrics for onset, first of all, I wonder if somebody could explain to me what the relevance is of placebo response to measuring onset? That doesn't seem to make any sense to me at all. If you are lying there in bed, looking up at the nurse giving you the medication, you want to know how long it is going to take this thing to work. You don't want to know when is the pharmacodynamic of the response of this medication going to separate from placebo. That is a completely noon-intuitive and clinically irrelevant measure. I would propose that for onset we look at actually onset, when the medication starts to work as opposed to when it separates from placebo.

The second issue I have with onset is that it is not at all clear to me why we are only interested in drugs that have onset within one hour. There are other characteristics of onset, aside from time to onset, that are also relevant. For example, in an NSAID I don't know what the typical rate is of responders that you see, but if you see that, for example, 60 percent of your patients will respond within an hour, I also might

be interested in a drug where 95 percent of patients respond but it takes an hour and a half and there are other ways to bridge the gap. So, I am not sure why we have this rigid notion that you have to meet your onset criteria, whatever that is, within an hour.

DR. FIRESTEIN: Can you clarify your point about differentiating from placebo? You don't think it is important to differentiate from placebo during that first hour?

DR. KATZ: Let's say, for example, that you give your drug to a group of patients and the median time to onset of the drug itself is one hour. In other words, you have a clinical sense that it is going to take on average an hour for that medication to work. If it doesn't separate from placebo for an hour and a half, what is the difference?

DR. FIRESTEIN: Because then you could just treat with placebo.

DR. KATZ: No, no, no, that is not true at all. The confusion I think is between looking at measures of efficacy of the drug compared to placebo versus looking at onset compared to placebo. Obviously, you have to show that your

drug is better than placebo in some way--a SPID or one of your measures that has been shown to be effective for that. But in terms of giving clinically important information about when the drug works, the clinician wants to know when the drug works; he doesn't want to know when the placebo works. So, whether the drug separates from placebo within that hour or it takes an hour and a half or two hours, or what-have-you, is a completely separate question, and I don't think the separation from placebo is a clinically useful metric of onset. The drug works when it works. The effectiveness of a drug is a combination of its pharmacological effectiveness and whatever placebo or non-specific effect it brings to bear, but in the real world both of those issues are operative.

DR. FIRESTEIN: One would wonder if you can't distinguish it from placebo whether or not it is truly a pharmacologic effect.

DR. KATZ: No, no, no, that is not my point at all.

DR. FIRESTEIN: I understand. Dr. Ashburn and then Janet.

DR. ASHBURN: I hesitate to speak before the biostatistician speaks, but I just have a

couple of issues that I wanted to point out or bring to the table. First of all, I want to remind folks that pain measurement in the acute pain setting needs to be both at rest and with movement, particularly in patients who are undergoing major operations, because that has been predictive of good quality of outcome.

The other one is onset, and in an acute pain setting I would reinforce Dr. Katz's remark. There is not necessarily a limit of one hour with regard to meaningful analgesia in the acute pain setting. There are medications that can be given preoperatively that do have a longer duration of effect, which is no longer relevant if you are trying to use a long-lasting medication and prophylax, if you will, for analgesia at the end of the operation. So, a one-hour onset may not necessarily be important when looking at a medication still intended for acute pain use.

Duration of effect, depending on the route of administration, may be very important. A 24-hour duration of effect in a patient who is going to be NPO for the first hour after surgery may actually be a very meaningful, important aspect of a different medication.

The other one is that adverse side effects tend to be overlooked with regard to blending that in with safety. Adverse side effects can be very important in a postoperative period. If a medication has a very low incidence of nausea and vomiting, for instance, that will be perceived as a marked advantage over parenteral opioids which do have a fairly high incidence of nausea and vomiting.

Of course, safety is paramount in these areas because one would tend to not tolerate a medication that even has a fairly low incidence of a catastrophic event. A medication that is relatively safe, that doesn't have opioid-induced risk of respiratory depression may actually have marked advantage even if it is equally as good as an opioid analgesic.

DR. FIRESTEIN: Excellent points. Dr. Elashoff?

DR. ELASHOFF: I wanted to comment on the issue of what was being called separation from placebo, which I assume means statistically significant separation from placebo, which is a combination of whatever the true separation is and the sample size that you used to look at the issue.

So, the whole issue of when they get far enough apart is both the issue of a clinically meaningful separation and the issue of whether the study is actually big enough to address that question.

DR. FIRESTEIN: Thank you. I always enjoy being chastened by the biostatisticians! Yes?

DR. KATONA: Just looking at the world from the pediatric point of view, even in other situations we do not like to do placebo-controlled trials. I am just wondering, in the acute pain situations, especially the postop pain, in special circumstances like with the children and the elderly, is that something that we need to compare, the active drugs with placebo, or could we do some other designs? I personally even wonder about the general population, if we could design these studies as comparison studies or some other ways.

DR. WOOD: Gary, I wanted to return to the point that you were raising right at the beginning of this discussion, and that is how long do we need safety data for, and how will that duration of safety data affect the potential for indications.

It seems to me that we have excellent data, going back to the question Mitch was asking, to say that labeling changes are not very effective

and are generally not followed. I mean, if we think of the example of fen-fen, the example of truplidazone, or the example of even Accutane, which has extraordinarily rigid labeling, physicians and/or patients are still not following these. Certainly with truplidazone the liver function tests were ratcheted down week by week and with relatively little effect.

So, the lesson from all of these, it seems to me, is that even a drug that was approved exclusively for acute use, such as one that was limited eventually to ten days' use in the example that was shown, was used for much longer than that. So, common sense would dictate that we should have safety data that extends for a much longer period than just a single dose.

If that is the case, you have to then say, well, how are you going to get that safety data? You could give patients or volunteers an analgesic for a long time for no indication which would seem to me to be dubious ethics and you are probably unlikely to get lots of volunteers. So, it seems almost inevitable, therefore, that if you are going to look for safety data that goes longer than the acute setting, you are going to insist de facto

that you look at chronic pain relief even for a drug that you might initially be looking at for only the acute setting.

I don't see a way around that, and you sort of touched on that in your question but I think we need to return to that because that actually is pivotal to how we think about this whole issue of development, perhaps not labeling but certainly how you develop it. If you are unable to go forward without chronic studies, then that is important to think about in terms of how you pitch your development program.

DR. FIRESTEIN: Would you require efficacy?

DR. WOOD: I would.

DR. FIRESTEIN: For the acute indication? If you propose that you would look for efficacy endpoints simply as a safety study, would you require efficacy in the chronic study in order to have approval for an acute indication?

DR. WOOD: Well, let me rephrase the question, if I may. I don't think the question is would I require efficacy data in the chronic safety study necessarily. I think it is improbable that a company or that you would advise a company to not

do an efficacy study if they were collecting chronic data because, otherwise, you would be doing a study in which you are giving an analgesic to somebody chronically for no very obvious reason, and I think it would be tough to get volunteers for that, frankly. Therefore, for relatively little additional cost you could get the efficacy data. I think most people would do that.

If someone came to you and said we don't want to do that, you would almost wonder why. I mean, is the reason that they don't want to do that because they have data that suggests it doesn't work chronically or it is toxic chronically? As a regulator, it would make me very uncomfortable if someone was adamant that they didn't want to do an efficacy study chronically when you were telling them they had to collect safety data chronically.

DR. SHERRER: I think that goes back to one of the original questions for why we came, and that is should we really then be dividing into acute and chronic pain? Because if we say that we are going to give these drugs for acute and chronic pain, in a sense we are saying that they work for both. Maybe the dosing is different but, in fact, the drugs work for both acute and chronic pain. In

practice that is really what is happening. So, does that go back into the mechanistic differences again, and are we really back to saying well, pain is pain? You know, we treat one way for acute and a different way for chronic.

DR. WOOD: Well, I think my point is a little more than that. I think that even if we could divide it into acute and chronic pain, and even if we really thought that that would be a good division to make--and I am not arguing for or against that--de facto, we have come to recognize that physicians and their patients are relatively poor at following that advice. And, it is not just true of pain; it is true of lots of other drugs. You know, fen-fen was taken for much longer than it was supposed to be. Truplidazone was taken without the appropriate liver function tests being done. Dosage creeps occurred with other drugs.

That is not a criticism; that is the reality of the marketplace. That being the case, it seems to me foolhardy to say that we are going to ignore all that data and say if a drug comes in only for acute pain we are not going to require a safety database that goes beyond that, even if we could make recommendations about how it should be

used and hope that it would be used in that way.

DR. FIRESTEIN: Dr. Max and then Dr. Farrar.

DR. MAX: I would like to comment on the metrics in the multi-dose studies. I think now the standard metric in looking at doses past the first dose is the choice of the patients when to rescue. I see nothing wrong with that because you are really using that just to tell patients when to expect to do that. The problem is this, I have spent many horrible afternoons sitting with drug companies, trying to massage a bunch of repeated dose data into some meaningful information and you can't get anything out of it generally because there are PRN doses with one regimen. The beauty of dose response studies is that you make the dose regimen the independent variable, and when you have the dose also be the dependent variable you muck it up completely.

So, I heartily endorse what I hear in your talks. Should we use the dose response type regimen and take multiple different regimens, either doses or times, and try to stick to it and use some other drug for rescue and find out what is too high, what is too low, and what is just right

for Goldilocks? That is the way to go about it.

There is one other finer point, and I think you have to define whether your main orientation is towards exploring the clinical pharmacology or usage study. That gets to the issue of whether you include placebos. Say you want to compare a six-time a day regimen of the same drug with three-time a day, there are some studies I have seen where they give placebos intermittently and then people say, well, the placebos gave analgesia and you really can't count them. It may be that if you really want to mimic usage, you want to do it unblinded so you get the full impact of the placebo effect of taking extra pills. But I think you need to spell this out so sponsors won't go ahead and use placebos or not use placebos and have the study be voided.

DR. FARRAR: I would like to pick up on something that Mitchell just finished with and get back to something that was said before. There are designs that are possible and completely valid to look at the way in which patients use medications. Two of them that are specific, one of which our group has suggested to some drug companies in terms of ways to look at long-term use but have not been

adopted.

The first one is in terms of the onset of effect and the efficacy, and that has to do with whether a patient at the end of the pharmacologic time period where they should have their maximal effect, whether or not they decide they need something else to treat that pain. That is very clinically oriented and it is a valid measure of whether the drug is ever effective.

The second thing has to do with long-term use. I think it was suggested before that giving patients drugs for a long period of time with no indication is a problem. What I would like to suggest is that one possible mechanism for dealing with that is, in fact, to do a very tight and carefully controlled study for a period of 4, 6, 8, 10, 12 weeks, whatever seems to be appropriate for the drug. In the long-term study it is possible simply to continue to give patients the medication as long as they want to take it. That sounds a bit odd perhaps, but ultimately what we are asking is how are patients going to use that, and is the drug safe for the period of time that they use it? If you want to study it long term, as in a safety study, you would give them the medication; follow

them as long as they are willing to take it, meaning if it still helps them, they claim it helps them for whatever reason; and look at the safety data over that period of time.

There is actually a more elegant way to do that which would in fact, be to continue to give the patients the medication in a blinded fashion long term. One of the arguments against that has been how can you possibly give somebody a placebo over the long term? My argument is to reverse that and to say if the placebo is providing real relief for the patient, then why not give it long term?

One of the ways of knowing whether a drug, in fact, works better than the placebo long term would be simply to give it blinded for a long time and follow, as was suggested yesterday, the number of dropouts.

DR. WOOD: But how would that differ from a placebo-controlled, long-term study? I mean, giving a placebo and an active drug for long term in a blinded fashion sounds to me like a placebo-controlled, randomized, controlled trial, which is what I am saying we need to do.

DR. FARRAR: Right, it is. The difference is the following, which is that in most of our

placebo-controlled trials there is a monitor that calls you every day and says, "have you used the drug? Did you write in your diary? Did you use your electronic diary?" What I am suggesting is that over a brief period of time, 4, 6, 8, 12 weeks, whatever is decided, that is reasonable.

But what you want to then study is the actual use of the medicines. So, what you want to do is to give them the medicine for, let's say, two weeks or a month, a month's supply and have them come back to visit you, and nobody calling them in between and finding out whether they took it or not; whether they filled out their diary. The issue is you use simply the continued use of that medicine and metrics that you measure once a month to determine whether or not they actually used it.

There is very clear evidence, as I think was suggested earlier, that if the monitor is somebody who makes you feel like you want to do what is right, or scares you into doing "what's right" you may use the medicine in a way that is very different than the chronic, normal use of that medicine.

DR. FIRESTEIN: Dr. Strand?

DR. STRAND: I just want to comment that

that is a rather standard design in, say, rheumatoid arthritis trials, and that is that patients are allowed to continue if they have had a response, open-label treatment for continued safety analysis.

But another thing that we have also done with placebo-controlled trials is that the responders, not unblinded, are allowed to continue treatment and that treatment is maintained blinded. We have actually had patients take placebo for as long as three years who respond clinically.

DR. CUSH: The limitations of that are as far as recruitment. I mean, I tell patients up front that you may be on placebo for three years and that is somewhat of a deterrent.

DR. STRAND: I think we say not that but that on or after a certain period of time, if you are not responding, you are allowed to go to active treatment. Then, all responders can go on to continued treatment and that way we don't imply that they will be on placebo for a long period of time.

DR. FIRESTEIN: Dr. Woolf?

DR. WOOLF: I would like to come to the issue of dose creep and the relevance of that for

the primary outcome measure, which I think we have all agreed should be pain. But I think the fact that patients tend to take higher doses than have been demonstrated to be effective might be a reflection of the fact that our measurements of what is effective are insensitive, and that patients may be getting a greater benefit than we can actually detect.

So, while primary pain outcome measures clearly are appropriate, there may be other aspects of the treatment that are making the patient feel better in a way that we are not detecting.

DR. FIRESTEIN: Yes, Dr. Brandt?

DR. BRANDT: Fundamentally I agree with what is being said about long-term placebo studies. But, as Vibeke said, there are practical problems with IRBs that are very significant in being able to do this.

DR. STRAND: It is not that they were told that they had to be on placebo; it is that everyone was offered to drop out for documented lack of efficacy, and only those people who responded stayed in and, therefore, we selected for a small group of patients who were placebo responders.

I would say part of any of these designs

would be the same thing, and that is people could not continue treatment beyond, say, the blinded time of the trial unless they were responders. But you can maintain a blind and find out some interesting information.

DR. FIRESTEIN: And even open-label extensions with safety rather than efficacy as a primary endpoint would not raise the bar that much higher for an acute indication.

There were a couple of other issues that were raised that the agency has requested that we discuss. One has to do with the parameter used for assessing dose intervals for acute analgesic drugs. The other, item three, is the issue of how one measures clinically important differences. Actually, I think Dr. Katz yesterday used a quote that I think I am probably going to put on my slide, which is if a difference doesn't make a difference, then what is the difference? Or some variation of that.

What I would like to do is try to steer us towards addressing those two issues right now. One is if anybody has specific thoughts on what sort of dosing interval studies would be required, or whether that is appropriate. Dr. Elashoff?

DR. ELASHOFF: Specifically with respect to 2(b), which is median time to rescue, and to (a) as well, which is the T 1/2, part of what was remarked earlier is that just looking at the mean or just looking at the median is not bringing in variability from patient to patient. One kind of thing which could be helpful in that is looking at the 25th percentile or the 75th percentile, that sort of information as well to help characterize how typical, in some sense, the median is of people and to try and get into the variability from one patient to another issues.

DR. KATZ: I am happy to say I was actually going to say the same exact thing. We have been talking a lot about how to get a precise estimate of duration by whichever metric, whatever that will wind up being, 8 or 11 hours, but to have some sense of how variable that is I think is very important. If two-thirds of your patients are within an hour of that, that is different than if two-thirds of your patients are within 4 hours of that and informs clinical practice better I think.

DR. FARRAR: I agree with what has been said, and I think what was just being suggested is actually best described as a box plot. It is a

very simple mechanism for actually displaying in an understandable format the 25th, 50th and 75th percentiles.

I think what it brings to mind is a second issue which is that patients are really quite different. In trying to help physicians understand how to use the medication what we really need to tell them is what is the minimum time that a patient should wait before they take an additional dose. That really is dictated by safety data. The question really is if a patient only waits an hour to take a second dose, an hour to take a third, and an hour to take a fourth they are clearly going to take much more medicine than if somebody waits three or four hours.

The example that comes to mind is when we prescribe medications for a patient 2-4 mg every 3 hours. What our patients will do sometimes is to take 2 mg but then, because they have taken the 2 mg they decide they have to wait the full 3 hours before they take an additional 2 mg, even though the intention was for them to be able to take up to 4 mg in that period of time.

What I am suggesting really is that in the label what it probably ought to say is something