wouldn't have any ADRs or you could have very, very high where the probability would be much higher. So that would be a misrepresentation of the actual information that you have by making it too simple. So I think there's the trick that you have to find the right balance.

DR. LESKO: We're in the world of safety here, and it would seem like in the clinical trial design, there's going to be prespecified endpoints of safety that the sponsor provides. It also seems to me that there's going to be a convention for a safety biomarker, let's call it, that will be continuous or categorical.

For example, if my concern with the drug is heart rate, I'm going to look at that maybe as a continuous variable because that might be the way it's measured and the way it's analyzed. On the other hand, if I was looking at a hematological toxicity like neutropenia, I might be concerned about grading the severity of that.

So I guess the question becomes, given there are certain prespecified endpoints in terms of severity and frequency, and given that there's a conventional way of presenting data as continuous or categorical, what would be the motivating factor to change continuous to categorical? What would the benefit of that be in terms of the pragmatic aspect of it, in terms of dose adjustment? Because you think about a drug that is going to be used therapeutically

and being monitored by these same endpoints and doses being adjusted by those same endpoints. So it's sort changing something from the usual to the unusual, and you'd have to say there's a reason to do that. And I guess it wasn't clear to me what the reasons for me to do that would be.

DR. SUN: I will give one example. Liver toxicity, one situation we saw in one NDA. If you see the frequency of liver toxicity based on concentration above a cutoff point, below a cutoff point, you don't see a difference. So it looks like all concentrations show -- if you only define liver toxicity as a yes or no parameter, you don't see a correlation between these two. But if you use actually the measurement of blood chemistry values, which is used as an indicator for liver toxicity, you're able to correlate concentration with the blood chemistry variable. So clinically it may be easy to see, well, 9 subjects or 10 subjects have liver toxicity, but we don't know correlation with concentration.

But on the other side, you can see actually when concentration or AUC increases, the chemistry value have some trend increase. Then you define a cutoff at that point that says on the curve where is the cutoff, then where is the concentration's cutoff? So there are two ways to present this data. The same data set we can do two ways.

My own experience in the drug labeling is that it looks like the first way is easier, the second way somehow, like you mentioned, has to be communicated to others in different professions as well.

DR. FLOCKHART: I would just emphasize that point. I think obviously you lose power with a categorical variable, but if you have a change that is picked up by the categorical variable, as well as by the continuous one, in general it communicates better. And we have a humongous problem with communication. That's an understatement. So I think when it's possible to state something in stark, clear yes or no terms to somebody who's practicing medicine in the area of drug interactions or recommendations within a label -- we're all aware of the vast majority of the label is just dust anyway. So when you can simplify it, that helps.

DR. VENITZ: Can I just make a general comment? Looking at some of the, I guess, labeling language or some of the statements that you reviewed with us, He, the only time they're going to be useful for a practitioner is if they actually draw blood levels because otherwise the fact if I'm above some certain level or some certain area, something happens or doesn't happen, it won't help me as a practitioner. The only thing that I want to know is can I change the dose and how do I change the dose.

So when you translate the information that you get from what you just reviewed for us and translate that into labeling language, you really have to target doses not concentrations, unless part of the therapeutic management with this particular agent requires dose titration based on plasma levels. Otherwise, I don't see how that is useful information for the practitioner to translate into practice.

DR. SUN: You prefer dose-response relationship rather than concentration-response.

DR. VENITZ: The only thing that the clinician can change is the dose unless part of the management requires taking blood levels, and then depending on the blood level, I can make certain adjustments.

DR. LEE: Can I say something here? I think what we propose here is we're not going to use a different approach either as population analysis or regular PK/PD analysis. So what we're presenting here is can we use population analysis to get a PK/PD relationship. But once we get a PK/PD relationship, we're going to follow the standardized approach to estimate the probability of an adverse event for special populations. So that's what we're trying to propose here.

DR. VENITZ: The statements that he reviewed for us, at least most of them, have some statement about

levels or areas that are too high, too low. And I'm saying from a practitioner's point of view, unless part of the management with this particular agent requires drawing blood levels so I can actually measure levels, it's useless. I need to know what to do with my dose because that's the only thing I can change short of changing the drug itself.

DR. SUN: But on the other side, let's say, same dose level due to drug-drug interaction change the concentration. This information will give us an idea if the concentration changed by such a degree, what's the probability in the ADR will be. So in terms of decision on this side, we still have to rely on this rather than only rely on dose. Although at the end, we can see -- the end and the level and we see in terms of drug-drug interaction, you have to deduct the dose by 50 percent, but a 50 percent deduction was got from concentration dose-response relationship.

DR. VENITZ: I don't have a problem with your conclusions. I'm questioning the usefulness of incorporating the conclusions as they are in a label to convince a practitioner to change a dose. That's all.

DR. SHEINER: Getting back to this issue of dichotomous versus continuous. There's a big difference as to whether you translate -- and I think this is what David

was getting at -- from continuous to dichotomous, let's say, on the way in -- that is to say, you change the data and then analyze the data that are now dichotomous -- versus on the way out where you take this complex model that deals with all the variables in their full complexity and then draw a very simple conclusion based upon simulations of that that says half the time it's going to be bigger than this if you do that. There's no problem with the latter, and that's very important for communication, to make things simple, to talk about the key issues.

The problem with the former is that you're losing information. The places you see it worst are in these clinical scores where you have 15 or 20 questions that you ask and then you add up the number of yeses and that's a number. They're clearly not combined optimally. You don't know whether one or another question would have more information than other things.

What they do is they allow you to do your analysis simpler. The only time that in my mind it's justified to simplify the data before you analyze them is when the loss of noise exceeds the loss of signal and you can afford some loss of signal and you're losing more noise by dichotomizing, let's say.

And I think most of us actually are probably --

I don't know. It would be interesting to take a survey of how many bits of biological information people believe there is in something, let's say, that's three significant figures like a serum sodium. Is there really the 12 or so bits that it takes to represent a three significant digit number or is it really only three or four bits? The sodium really 140 to 143 is all the same, et cetera. I think there's probably a lot less information in these continuous variables than we think there is because of the high feedback systems that we're dealing with.

But the main point, it seems to me, is the three questions that I always say you've got to ask before you do anything. So this one is, what's the question? If you want to write a label that says that as you increase the dose by these increments, the probability of this toxicity will go up by those increments, then you have to have a model that somehow represents continuous probability versus continuous dose. If you don't want to write a label that says that but only says don't give the drug to people who have values beyond these, then you don't need that. So first figure out the question.

There I think maybe is where one could sometimes schedule a meeting to talk about that. What should recommendations for dosage changes in special populations look like? What kind of statements ought we be

trying to make? And that will determine the data we want to gather and how we want to analyze it. I'm not sure that's all settled, but I'd like to hear at some point what the agency does think about what constitutes a complete set of statements and what degree of precision and what kinds of words you use and what kinds of things you want to be able to say.

DR. VENITZ: Any comments, Larry or Peter?

DR. LESKO: No. We've sort of thought about that last statement that Dr. Sheiner brought up. I think it's a very valuable statement, but I think we need to think about it and put the story together and bring it before the committee. But I think it would be a very interesting discussion to have, what elements of a good label should there be in terms of a probability of risk and an intervention of some sort and how do you present that in a consistent way across special populations or something like that.

DR. VENITZ: A couple of more comments, I guess more general comments, not specific to your question. But how to translate information that you would gather from this kind of analysis and translate them into labeling language. How is the drug being administered? Is it titrated on some kind of effect? That would make a big difference in terms of how some of those things would

translate into a dosing recommendation because you may not have to make a dosing recommendation because you're going to pick it up as part of the normal way of therapeutic drug monitoring. And it might not be a drug level. It might be some surrogate marker, some biomarker. Obviously, what are the available ways of adjusting the dose? Do they have the dosage forms to accommodate that or can you not?

And then along the titration route, do we know anything about intra-individual variability that would allow us to assess for a given patient how likely they're going up and down, and is our dosing algorithm going to pick that up?

Those are questions that aren't really addressed in what you're talking about, but I think they're very relevant for translating this information into recommendations in the labeling.

DR. LESKO: Just to add to that in thinking about maybe a future discussion with the proposed labeling rule that the agency has out, there's going to be some revamping of the label such that certain information gets to the top of the label out of the individual sections. And a question could be raised as to what criteria in the context of what we're talking about would warrant raising information to a more prominent part in the label. If that is done for catching the attention of a prescriber, it gets

to that language and how you present the information in translating it into therapeutics. So I think the whole thing sort of flows as a future issue.

DR. SUN: Yes. We do see situations where a label can be -- let's say it's an IV formulation or all doses are available. You can put a table. It shows about every year with a different dose. But when the formulations only have 15 and 30 milligrams, you only can do is categorize them to two classes. That is really true.

Let me summarize. As will be -- if we translate from continuous to categorical, we lose information. But on the other hand, maybe categorical is easier to communicate. And it also depends on the outcome, what are you going to do before you handle the data. Before you even start an analysis, how do you use this information.

Thank you very much.

DR. VENITZ: Any other comments?

DR. DERENDORF: Maybe it's way out there, but it seems that one problem is not so much the information and the data analysis, but it is really the ability of the user to do something with it. I think maybe one of the reasons is that we're limited in the label to just a written document, and with modern technology, a lot of that information can be presented to someone in a palatable way

like in a little computer simulation where all the information is entered, and then there's a recommendation or an assessment that comes back. That may be something to consider in the future.

DR. RELLING: I'm new to this, but you talk about special populations and you've presented different examples of covariates. is there some sort of list that you have of what you consider to define special populations and what you consider covariates that should always at least be asked about? We heard about ketoconazole and something else, age or renal dysfunction. Are you letting individual studies drive these things? Are you going through some algorithms to define what you should look for a priori? How are you deciding what you'll even include or think about looking at? Because there are all kinds of examples where we're making dumb mistakes like looking at ketoconazole but not itraconazole. How are you going through this stuff?

DR. SUN: The first part you're asking how to define the special populations. In the legal terms, they define special population as disabled patient/subject, blind subject or some others. In our clinical pharmacology term here, we refer to subpopulation like a pregnant patients, pediatric subjects, or other ethnic groups rather than the legal term defined those special populations.

That's the first part.

The second part. When we do clinical trials in phase III trials, we do have some inclusion/exclusion criteria. A majority of the time it's how the drug in the future will be given to patients in the clinical setting. So most likely we will include other subjects who potentially will take this drug.

Did I address your question?

DR. RELLING: Not really.

DR. LESKO: Let me add to it. First of all, the answer to the question is somewhat drug dependent, but there is a standard range of assessments that are expected within the drug development program. And most of these are revolving around the changes in pharmacokinetics. So, for example, there are demographic factors. We would want to know age, gender, and race and the effects on pharmacokinetics and whether that's pertinent to dosing adjustments that might be necessary in those subpopulations or special populations.

We then move next to intrinsic factors, as we call them, and they're predominantly disease states that handle drug disposition. So renal disease and hepatic disease is a standard study that's in most NDAs.

And then finally the issue of extrinsic factors and predominant amongst those are co-administered drugs.

So there's a heavy emphasis, ideally mechanistically driven, to look at clinical drug-drug interactions that are likely to be important in the clinical setting.

So the categories of special populations are defined by these demographics, the intrinsic, and the extrinsic factors.

Now, depending on the drug used, there may be additional special populations that would be looked at, but I would say the ones I just mentioned are the standard covariates that you would be interested in.

DR. RELLING: Do you require those to be addressed a priori in the trial? And what makes the difference whether you decide -- when are you going to start looking at genetics? Where do you start drawing the line of saying you've got to look at something besides a creatinine and a bilirubin or whatever it is you are asking for?

DR. LESKO: "Require" is kind of a harsh word. I'd like to think of it as recommendations. By and large, everything I just said is contained in guidances to the industry that say what the expectations of the agency are, and if those types of analyses aren't available, the burden of proof then is on the company to say why it's not important in the case of that particular drug.

When you get into what I would call evolving

areas or evolving covariates -- and you mentioned genetics I think as one of them -- we have now appearing on sort of the drug development scene the ability to look at changes in DNA sequence that would describe what we typically have called phenotypes, poor metabolizers and extensive metabolizers. It's a logical extension, to me at least, to begin to ask the question, given the availability of a test to measure a genotype, at least answer the question somewhere in drug development, is that an important covariate. It's no different in my mind. The fact that it's genetic doesn't make much difference to me in a sense because it's a covariate that could be identified that may have a significant effect on exposure and then subsequently response. So it would seem to me you either need to have some information that would say I need to worry about it or I have information that would say I don't need to worry about it.

If one thinks about the size of special populations, certainly a genotyped group, i.e., a poor metabolizer for a 2D6 substrate, is much larger in size than would be, say, a patient population defined by their renal function. So I think we're moving in that direction as the science evolves and as our ability to identify these covariates becomes more commonplace and available.

DR. SADEE: It appears to me that these models

all still assume exposure in terms of how much drug is there. So the alternative is to look at special populations that have a clearly different dose-effect relationship. So what I'm struggling with, in terms of trying to understand how to simplify this and how to still get reality in there, is if you have two different populations with the outlier, the toxicity occurrence, is because somebody is genetically or environmentally predisposed in a way that cannot be predicted by the model, that it's just looking at exposure, how much drug is in the body. And if you try to merge the two, you're obviously making an error. On the other hand, it's clear that even for those patients, the more drug you have, the more likely it is that there would be an adverse effect.

So I think we're talking about two different models here. One assumes we have this relationship between how much drug is there and the effect, and the other one would be there's a completely different relationship. Those two have to be merged, I suppose, without then incurring too large of an error.

DR. LESKO: Yes. I think that sort of brings home a very interesting point, and we kind of talk about it but don't know exactly what to do with it. And that is the paradox of drug development, as I call it, that drug development revolves around population signals and

population data, whether it's the efficacy signal, the safety issue, or the pharmacokinetics. You're looking at a population average.

Yet, as Dr. Sadee has mentioned, when you're treating patients, you're worried about the individual situation and how you can best optimize a dose in the individual. When we talk about individualizing or optimizing dose, I'm sure we're talking about it in the context of a subpopulation or special population defined by something I could measure or observe as opposed to the individual that genetically may be predisposed.

I think it's easy to focus on pharmacokinetics and much harder to focus on the factors that influence receptor sensitivity or the things such as long QT and things of that sort because they're a little bit more complex, especially in the polygenic nature of being more complex.

DR. VENITZ: Any further comments?

(No response.)

DR. VENITZ: Then it looks like we're going to get an early break. We're going to break from, I guess, right now until 12:45.

We have nobody signed up for the public hearing. So we are starting at 12:45 with Dr. Karlsson's presentation, the example number 4, in using exposure

response to recommend dosing adjustments.

So enjoy your lunch and we'll be back at 12:45.

(Whereupon, at 11:40 a.m., the subcommittee was recessed, to reconvene at 12:45 p.m., this same day.)

handle drug disposition. So renal disease and hepatic disease is a standard study that's in most NDAs.

And then finally the issue of extrinsic factors and predominant amongst those are co-administered drugs. So there's a heavy emphasis, ideally mechanistically driven, to look at clinical drug-drug interactions that are likely to be important in the clinical setting.

So the categories of special populations are defined by these demographics, the intrinsic, and the extrinsic factors.

Now, depending on the drug used, there may be additional special populations that would be looked at, but I would say the ones I just mentioned are the standard covariates that you would be interested in.

DR. RELLING: Do you require those to be addressed a priori in the trial? And what makes the difference whether you decide -- when are you going to start looking at genetics? Where do you start drawing the line of saying you've got to look at something besides a creatinine and a bilirubin or whatever it is you are asking for?

DR. LESKO: "Require" is kind of a harsh word. I'd like to think of it as recommendations. By and large, everything I just said is contained in guidances to the industry that say what the expectations of the agency are,

and if those types of analyses aren't available, the burden of proof then is on the company to say why it's not important in the case of that particular drug.

When you get into what I would call evolving areas or evolving covariates -- and you mentioned genetics I think as one of them -- we have now appearing on sort of the drug development scene the ability to look at changes in DNA sequence that would describe what we typically have called phenotypes, poor metabolizers and extensive metabolizers. It's a logical extension, to me at least, to begin to ask the question, given the availability of a test to measure a genotype, at least answer the question somewhere in drug development, is that an important covariate. It's no different in my mind. The fact that it's genetic doesn't make much difference to me in a sense because it's a covariate that could be identified that may have a significant effect on exposure and then subsequently response. So it would seem to me you either need to have some information that would say I need to worry about it or I have information that would say I don't need to worry about it.

If one thinks about the size of special populations, certainly a genotyped group, i.e., a poor metabolizer for a 2D6 substrate, is much larger in size than would be, say, a patient population defined by their

renal function. So I think we're moving in that direction as the science evolves and as our ability to identify these covariates becomes more commonplace and available.

DR. SADEE: It appears to me that these models all still assume exposure in terms of how much drug is there. So the alternative is to look at special populations that have a clearly different dose-effect relationship. So what I'm struggling with, in terms of trying to understand how to simplify this and how to still get reality in there, is if you have two different populations with the outlier, the toxicity occurrence, is because somebody is genetically or environmentally predisposed in a way that cannot be predicted by the model, that it's just looking at exposure, how much drug is in the body. And if you try to merge the two, you're obviously making an error. On the other hand, it's clear that even for those patients, the more drug you have, the more likely it is that there would be an adverse effect.

So I think we're talking about two different models here. One assumes we have this relationship between how much drug is there and the effect, and the other one would be there's a completely different relationship. Those two have to be merged, I suppose, without then incurring too large of an error.

DR. LESKO: Yes. I think that sort of brings

home a very interesting point, and we kind of talk about it but don't know exactly what to do with it. And that is the paradox of drug development, as I call it, that drug development revolves around population signals and population data, whether it's the efficacy signal, the safety issue, or the pharmacokinetics. You're looking at a population average.

Yet, as Dr. Sadee has mentioned, when you're treating patients, you're worried about the individual situation and how you can best optimize a dose in the individual. When we talk about individualizing or optimizing dose, I'm sure we're talking about it in the context of a subpopulation or special population defined by something I could measure or observe as opposed to the individual that genetically may be predisposed.

I think it's easy to focus on pharmacokinetics and much harder to focus on the factors that influence receptor sensitivity or the things such as long QT and things of that sort because they're a little bit more complex, especially in the polygenic nature of being more complex.

DR. VENITZ: Any further comments?

(No response.)

DR. VENITZ: Then it looks like we're going to get an early break. We're going to break from, I guess,

right now until 12:45.

We have nobody signed up for the public hearing. So we are starting at 12:45 with Dr. Karlsson's presentation, the example number 4, in using exposure response to recommend dosing adjustments.

So enjoy your lunch and we'll be back at 12:45.

(Whereupon, at 11:40 a.m., the subcommittee was recessed, to reconvene at 12:45 p.m., this same day.)

AFTERNOON SESSION

(12:50 p.m.)

DR. VENITZ: Can we reconvene the meeting, please?

Our next presentation is Dr. Mats Karlsson, the faculty of Uppsala University of Sweden. He is going to talk about optimizing dosing strategies for defined therapeutic targets. Dr. Karlsson.

DR. KARLSSON: Good afternoon. I'm really grateful for this opportunity to get some insight into the American regulatory process.

So I'm going to talk about optimizing dosing strategies for defined therapeutic targets. I'm going to talk about target definition in relation to dose finding.

We have done some work, mainly simulation work, but then also applications to a few real drugs under development, and I'm going to focus more on those as examples of what we've been doing.

The dosing strategy alternatives that there are are, first of all, the single, one dose fits all always, which of course is very convenient for patients, clinicians, and producers, if it's appropriate, but often variability and pharmacokinetics and pharmacodynamics will lead us down the individualization route.

There are two types of individualization.

First, we can individualize based on patient characteristics, observable, or feedback individualization based on some measurement or observation of the patient, or we can have a combination of these two.

So the next thing is the defined therapeutic target. Anybody who is involved in decision making on dosing strategies have to have some implicit target concept. This might not always be quantitative. It might not always be stated, but there has to be some target. And if we were to quantitate it and actually spell it out, we could base it on, most reasonably, the weighted balance between beneficial effects and side effects. You could also consider other endpoints like only one side of the coin or drug concentrations or biomarkers, especially when it comes to individualization in subpopulations, as we have been discussing today. And, of course, the target might differ between different patient subpopulations. We want to treat more severe disease maybe more aggressively.

We need not only to know the target, but also seriousness of deviation from the target, and we can use an all or none criteria which we recognize from pharmacokinetics as the simplest concept of therapeutic window where all concentrations within the therapeutic window are equally desirable and all concentrations outside equally undesirable. And the pharmacodynamic correspondent

is the responder/nonresponder concept here.

But we usually think that biology works in a more graded manner and we might want to actually value the deviation from target in a more graded manner. Of course, what's important there is the clinical picture, but we might approximate that with various statistical distributions. And also the seriousness of target deviation may vary between patients, although I haven't seen any examples of such applications.

When I'm talking about the target concept and the penalty function, of course together these form the utility function, and I understand that that was the topic of talks at the last committee meeting.

So one scenario for selecting a dosing strategy would be that it's based on some implicit criteria on the target and penalty function, and then that more or less stops there.

Another scenario would be the same thing, but then take it a step further. If one has a population model for the dose, two-target variable, one could actually estimate based on the decided dosing strategy and the model, what target and penalty function does that correspond to and then assess whether it seems to be a reasonable target, and if it is, stop there. Otherwise, revise your dosing strategy.

A simple example that we came across when collaborating with a drug company was a drug under development for a disease which had two harmful events. The frequency of one was decreased by the drug and the frequency of the other was increased by the drug. Of course, what exposure, what dose you would choose would depend on how you evaluate these two against each other. If this effect is deemed more harmful than this, you would choose a low dose. Otherwise, the opposite, the high dose.

So we did some calculations. So the black line here corresponds to what I was just saying. If we weight the adverse event high compared to the event that is beneficially affected by the drug, then we would choose a very low dose, and if they're equally weighted, we would choose a much higher dose.

In this case, the project team had already selected a dose of 1, which corresponded to a weight of 3 to 1 for the two events. And when we presented the project team with this, they said, well, this seems to be a reasonable weight.

However, there were two sub-diagnoses that had actually different PK/PD relationships, but they had decided to go with the same dose in both subpopulations. So, of course, that meant that the weighting was different for the two subpopulations. So in one it was 4 to 1 and in

the other 2.5 to 1, and again the project team said, well, that's reasonable because in the sub-diagnosis, when the harmful event occurs, it's more serious than in the other.

Then finally for renally impaired patients, a dose of a quarter of that for the main population was selected and it corresponds to the same weight between adverse event and beneficial event.

So this is a way to rationalize a selected dosing strategy after the effect.

Of course, one might have more use of defining the target and penalty function beforehand, and I'm certainly no expert in this area, but what seems to be wise is to ask a clinician because they're really the ones who are sitting on the information here although maybe not used to formulating these type of functions in quantitative terms.

If there is a drug first in class, consult preclinical phase I data to assess what tolerability issues there might be.

Consult literature and marketing which might have done surveys in patients and clinicians of what are deemed important features of a drug therapy for a certain disease.

And then develop a few alternative targets and penalty functions, and apply it to historical data, maybe

make up a bank of hypothetical patients to be ranked.

And then again ask a few clinicians about the developed utility functions. Most likely they won't agree, which might be a source for revising the utility function, but even so, you might not always get full agreement between different clinicians, and just as Lewis said earlier today that we need to incorporate uncertainty in all our aspects, maybe we need to include uncertainty or variability in the utility function as well.

So if we actually have a defined utility function, then we can proceed in a more rational manner. If we have the utility function, we have a population model for the dose-to-target variable, we can estimate the best dosing strategies given different constraints such as we're going to give everybody the same dose. We're going to individualize, but only with two doses and based on a covariate, or we might individualize based on feedback, et cetera, and then select the dosing strategy based on target fulfillment and practical considerations.

So what we would do in more detail for the first step, if we want to optimize the one-size-fits-all dose, would be to, based on the utility function and population PK/PD model, we would actually maybe not need the full PK/PD models that we usually use if we're only considering steady state concentrations' relation to

effect. We might only need the model for clearance. Covariate models are essentially superfluous because they're not going to affect our dosing, and we would, based on these models, simulate a large number of hypothetical patients. We would obtain a prediction of each individual's deviation from the target for a certain dose, and then obtain the optimal dose by minimizing the overall loss. In this case, we can do this simply by just repeated simulations, trying different doses, but we could recast the problem as an estimation problem and estimate the dose instead.

This actually is just a pictorial slide showing the same thing.

So if we want to actually do individualization, the questions become more and it's more problematic of how to do it best. I'm only going to focus on this first question for the case when we want to do feedback individualization. What dose strength should be made available?

We came into such a problem when collaborating with a company that had developed a drug. It was planned to go into phase III as a fixed dose size, everybody getting the same dose, but in light of high variability in PK and PD, partly because of polymorphic 2D6, they were contemplating maybe doing individualization. And the

question was, what would we gain by that? Would it be sufficient? And they wanted the gain to be measured on a responder scale and the overall responder rate.

We went ahead and did the estimations based on one-dose-fits-all or a feedback individualization with two dose strategies where we estimated the lower dose size, the higher dose size, and the fraction of the patients that would be preferentially treated with the lower and the higher dose. I won't go into technical details here, but we used the $MIXTURE function in NONMEM.

DR. SHEINER: (Inaudible.)

DR. KARLSSON: No. This is feedback individualization without --

DR. SHEINER: This was feedback on the response?

DR. KARLSSON: Yes.

DR. SHEINER: So actually somebody was observed to be a responder or not a responder, and then the dose was changed.

DR. KARLSSON: Yes.

We had built population PK and PD models for both the satisfactory effect and for the side effect previously. These were more elaborate models with continuous and ordered categorical type data, but they could be easily reduced to the dichotomous question of were

they responders or not responders.

And when we estimated the single dose size to be given, it was very close to what the project team actually had come up with, and it resulted in 47 percent overall responders.

The best two-dose strategy was two doses, one lower and one higher, a 4 and 18, with about 60 percent gravitating towards the higher dose. This was predicted to increase the overall responder rate to 63 percent. This maybe was one piece of the puzzle that made the company actually in the phase III to go with both fixed dose and individualization in parallel.

We did simulation studies on similar problems, and just to relate two observations there, one was that although a particular dosing strategy may not be the most optimal for one utility criteria, it may be near optimal across all relevant utility criteria, and therefore may be superior to other dosing strategies.

Also, another observation that all-or-none type responder definitions, like in this example, seems to favor individualization to a higher degree than gradual utility functions. Although this was obtained from a single example, it seems reasonable that if you have these very harsh, steep benefits of changing maybe somebody just a little bit on a continuous scale into having a utility of 0

to have one of 1, that that is more sensitive to individualization.

Moving on from feedback individualization to individualization based on covariates identifiable up front when the patient is to be started on a therapy, we will face questions such like what is the best covariate to base dosing on? What should the number of dose sizes be, and what covariate intervals should each dose size be applied to? Just illustrating here, the covariate here might be organ function, body size, age, et cetera.

If we want to go with two dose groups, the parameters we need to identify are what is the optimal cutoff value, what is the dose in the higher group and the dose in the lower group? So that's three parameters to actually estimate.

If we wanted to have three dose groups, that's five parameters, et cetera. The problem becomes more complex.

We would proceed to estimate those parameters in a very similar fashion as before, but with some differences. First of all, in this case it's, of course, very important to have covariate models for the covariates that we are intending to be using for dosing decisions, and we also need to have distributions of these covariates in the target patient population. We can obtain those from

simulations, but more relevant maybe from empirical databases from previous studies. What we're estimating are then the dose sizes and the cutoff values for them, where to change dose size.

We also had an example of this in relation to drug development. I can actually name this. This is NXY-059, a drug under development by AstraZeneca. We had a publication coming out in Clinical Pharmacology and Therapeutics in January this year.

This is a drug to be used for stroke. It's acute dosing, a 72-hour infusion with a 1-hour loading infusion.

The project team was worried about too high variability if one were to give everybody the same dose in particular since this is entirely -- not entirely, but mainly renally cleared compound, and they were worried about the end of loading infusion and the maintenance infusion concentrations.

So we tried to see what individualization could do in terms of bringing down variability to a reasonable level. The target that was set was a free concentration of 100 micromolar. The penalty function used was quadratic loss in log domain which means that half the target concentration is as bad as twice the target concentration.

We had a pop PK model developed from the first

patient study which showed clearance being highly dependent on creatinine clearance and volume on body weight.

We used empirical covariate distributions from previous phase III studies of stroke patients.

And for the loading infusion, we considered one dose, the same to all, or two-dose groups either based on creatinine clearance or on weight.

For the maintenance infusion, it was clear that we could not give everybody the same dose, but two to four dose groups were explored and dosing were to be based on creatinine clearance.

An additional constraint was made, which said that the therapy has to be fulfilling the following criteria, namely that 90 percent of the patients have to be above 70 micromolars and less than 5 percent above 150 micromolars.

As it turned out, actually these criteria could be met by just giving one loading dose of 2,400 units to everyone, but for the maintenance infusion, it was necessary to give three different infusion levels, depending on the creatinine clearance, with the cutoffs at 80 and 50 mls per minute. And you can see the dose units there.

This dosing design was implemented in a phase II study, and the target fulfillment was acceptable, with

more than 92 percent above the lower limit but more than 7 percent above the upper, which is slightly more than what was desired.

So actually before that, we had done some simulations to look at a priori dosing based on covariates, and we took dosing on creatinine clearance as an example. The standard approach often used when individualizing doses based on renal function is to use predetermined cutoff values for the renal function and quite large dose decrements, often a factor of 2 or higher, when going down to lower renal functions.

We wanted to explore what would be optimal approaches to renal based dosing and we wanted to see what drug characteristics and what other factors influenced what would be the optimal approach. So we did simulations where we changed the drug characteristics of hypothetical drugs, where we changed the creatinine clearance distribution in the target population, and also the utility function shape.

And two factors came out as by far the most important. For the selection of what would be the optimal cutoff value in the patient population, the median creatinine clearance in the patient population was most important. And if only two dose groups are to be used, the cutoff should be ideally positioned close to that median value regardless of other drug characteristics.

For the dose ratio, the ratio of the high to the low dose, the one factor that again by far was the most important was the strength of the covariate relationship, and for renally cleared compounds that can be expressed as the fraction excreted unchanged. So when the fraction excreted unchanged is 1, the higher to the lower dose ratio would be around 1.7. And for all other situations, the ratio would be lower than that ideally.

These two pictures are quite in contrast to what the practice is today which is using cutoff values below the median value and often dose ratios higher than 1.7. This might have practical and other reasons, but it's also to be maybe recognized that this has an impact on the target fulfillment.

This is just a picture showing some of the gains that could be made from doing individualization based on fraction excreted unchanged and the number of dose levels.

This is actually a picture very similar to the one that got us involved in this area. It was when again we collaborated with a drug company who had already decided upon a dose individualization scheme where they had actually selected a low cutoff value and a large dose decrement. So what in effect they were doing was they changed around the doses, but they didn't manage to lower

the variability in exposure. It was different types of patients that were in the tails of the distribution, but the overall variability in exposure was not reduced by the individualization. So doing this might actually not be the most simple task.

So to summarize with respect to target definition or utility function definition, this is something that certainly can aid data collection. If one knows what parameters are the most important and therefore go into the utility function, that will aid both data collection and modeling efforts, being able to maybe take both quantitatively and qualitatively better data for those variables and also do modeling more focused on those.

To improve communication within project teams or maybe between project teams and those outside. My experience is that many of the important factors for the utility is something that resides with only a few persons within the project team and many of the others that are to contribute to the dosing decisions are not particularly well informed about the weighted balance between effects and side effects or between different types of effects.

And of course, it's important to appropriately value the drug compared to other drugs.

And the last point. This was a slide prepared for a meeting in Europe last December that had the title

"Getting the Dose Right." If you ever want to say that you got the dose right, obviously you need to know what you're aiming for.

Separate from defining the utility function, which in its own right has a lot of benefits, dosing strategy estimation might have additional benefits to motivate the choice of dose or dosing strategies and to obtain conditions for optimal individualization and thereby assess the maximal potential value of individualization to justify doing it or, oppositely, maybe to justify not doing it before because the benefit of doing it isn't large enough. And if you do know the optimal individualization strategy, then you can directly offset any practical consideration that simplifies dosing against a decrease in target fulfillment.

I know that some people don't believe me when I say it's easy, but compared to doing population PK/PD modeling, compared to defining utility functions, which are the two more difficult tasks, I think this is very easy.

Thank you.

DR. VENITZ: Thank you very much, Dr. Karlsson.

Any questions or comments by the committee? Let me ask you, in your simulations, you used a symmetric loss function. Right?

DR. KARLSSON: We used both symmetric and non-

symmetric loss functions. Obviously, if you use non-symmetric so you penalize the high concentrations or adverse events more, then you're going to gravitate towards lower doses. If you are non-symmetric in the other direction, you're going to penalize the low doses, so you're going to get a higher overall dose.

DR. VENITZ: So your conclusions with regard to the effect of mean creatinine clearance, that was based on a symmetric loss function. Right?

DR. KARLSSON: We actually did explore various loss functions also there, and across a range of reasonable loss functions that we thought are reasonable, it was pretty stable towards that. But obviously if you have a very asymmetric loss function, you will tend to get other values.

DR. VENITZ: And my guess would be that's the reason why your recommendation is different than what's currently done because what people build in is a loss function where they're very worried about overdosing, less worried about under-dosing. So the easiest way to account for that is by adjusting the dose in people that have renal failure.

DR. KARLSSON: Yes. I think that's true, that they are usually dosed to an average AUC that's actually below what's seen in the main patient population. So if

that is what is decided, then certainly that's the case. For the example we had here, where actually we're operating at one point with different loss functions for people with lower renal functions because if there was a concern about renal function, then we wanted them to have a lower target as well. So there is certainly the possibility of incorporating all these types of considerations.

I think the reason why we see more lower cutoffs and large dose decrements is probably the way drug development pursues with more healthy patients to start with and then inclusions of larger and larger. So the initial dose levels are based on those with higher renal function and then the other ones are added as a tail more towards the end.

DR. SHEINER: I just wanted to sort of raise a point in the questions. You were romping through your slides. So you looked at individualization based on that drug that you told us the name of, that slide you just showed a moment ago where it had the 90 percent and the 5 percent?

DR. KARLSSON: This one?

DR. SHEINER: No. At the bottom of the slide, it had selection of dosing strategy.

DR. KARLSSON: Okay, yes.

DR. SHEINER: Yes, there.

Correct me if I'm wrong. I think one important thing to realize is that those numbers down at the bottom, of course, might not be attainable by any strategy. But to discover whether they are or are not attainable by some strategy, you fix on a model for the process. Then those percentages there have got to do with variability among patients. So this doesn't contain that uncertainty that I was talking about.

The model that says how frequently you'll get toxic at a given level, that very model is itself uncertain because it's based on assumptions and it's based on data. It would be very likely with not a great deal of data -- it would be impossible to attain a 90 percent probability of being in the right range or whatever it is if you incorporate that kind of uncertainty because that uncertainty says, I don't know how the world works. So how can you possibly get 90 percent certainty on anything?

And there's nothing wrong with this. This is the right way to proceed. Then you want to look -- presumably you did -- at the sensitivity of those two various assumptions that went into your model.

But this whole business of being clear with people about what uncertainty you're bringing in when, when it's appropriate and when it's inappropriate -- it would be here inappropriate to bring in model uncertainty when

you're trying to find the strategy because it's got to condition on some state of nature.

So I just wondered what your experience was with dealing with all of those, I think, somewhat subtle issues with people who are not used to speaking this language.

DR. KARLSSON: It is true that we actually didn't -- at first, when we saw these, we thought they were too stringent criteria. With normal variability, this would be very hard to achieve, and as you say, it might not actually be possible.

When doing these calculations, we did not take the structural model uncertainty into consideration, but we did do simulations looking at the uncertainty in the population PK parameters, what the impact of that would be. And it wasn't actually as large because they were relatively well defined. But this is, as you might guess, done with the point estimates.

In general, I do find it often difficult to discuss these matters with the project team I think maybe because it's not usually done, talking about quantitative models in this sense. The type of utility function that does seem to be used is the responder/nonresponder criteria. So it's easier if it's simpler, but again, the discussions around the responder criteria was where to put

the cutoffs and people differed in their opinion of where to put the cutoff, which is the same thing.

DR. LESKO: Mats, I want to bring you back to a little more maybe pragmatic question. But clearly this can be done in the context of what we talked about this morning in having a data set in front of us and then looking at an approach like this to adjust -- or make a decision based on adjust dose.

However, as I listened to you speak about the project team where there's a need to define a target and the penalty function, this method requires a weighted balance between the effectiveness and the safety. I guess in some cases that's very clear, depending on the nature of what the effectiveness and safety is.

If you think of risk as sort of an overriding issue in drug development as opposed to efficacy -- in other words, the risk of an adverse reaction, limiting approval, limiting dosing, limiting the label -- does the notion that a safety consideration would drive the relative agreement that you would have in a project team on the utility function -- in other words, I'd give more weight to a safety side of the drug's effect as opposed to the efficacy side. As an example, if I had a QTc issue, that would seem to weigh heavily in terms of my utility function consideration even against the most promising efficacy that

I might be speaking about. So it would be the driver, if you will, in trying to get an agreement on what weights to assign to the utility functions.

Can you speak to that a little bit? How does that work? And we sort of talked about this last time. You said ask a clinician, but we have actually asked clinicians after our October meeting and we do get quite different views of how a utility function would serve the purpose of what we're talking about. So I wondered if you can sort of pursue that a little bit.

DR. KARLSSON: I wouldn't have any hopes or belief that utility functions within the very near future would drive the decisions so that you would just define a utility function and then forget what went in there. I think a more reasonable way is to actually come up with some initial utility function and then use that in parallel maybe to illustrate different consequences of decisions. Then that would maybe show up where utility functions would actually fail because it wouldn't take something into consideration that is of importance, which would point to how they need to be refined, what needs to be considered in them.

When it comes to the safety, I think in some situations where there are tolerability issues of maybe not so severe nature as QTc prolongation, it's easier to

incorporate them. Obviously, severe side effects are -- by its nature, you're not going to see very many of them, and you're going to not want to expose patients or volunteers up to the range where you get a very good handle on what the function is at the dose levels. So you're going to have a large uncertainty in the models at that range, and you need to take that into account I think. So you're going to have a much larger uncertainty on the upper end than on the lower end probably.

DR. SHEINER: I know you know my response to this, Larry, but I just wanted to say it keeps on reminding me of the old data analysis argument about the Bayesians and frequentists and the issue of I'm trying to fit a model that's too big for my data, and so I'm going to fix some parameters. The Bayesian, of course, hearing that, says well, okay, you may not think you'll be acting like a Bayesian but fixing parameters is the same as saying they have a prior distribution that's got point mass at the value that you said. That can't ever be as good as giving it a little bit of wiggle room.

Well, it's the same kind of thing here with utility. Of course, if you have trouble convincing people that they ought to sit down and do a utility function, then there's some implicit utility function that's dominating usually by the most aggressive person or the person highest

up in the organization who happens to be at the table. And you don't know what it is. If you believe in decision theory, it's got to be there somewhere, and it's what's overriding everything else.

Just listening to the talk and saying, wouldn't I like to be in a room where the discussion was about how much do I believe that these data really do support this notion of what's going to happen, how much weight do I put on versus how much you put on, the various side effects. It seems like such a rational and sensible discussion to have rather than, well, I think we ought to go with 25 milligrams. What do you think, Joe?

DR. VENITZ: Any further questions for Dr. Karlsson?

(No response.)

DR. VENITZ: Thank you again for your presentation.

Peter, do you want to pose the questions for the committee?

DR. LEE: I think the question is very simple. Can this methodology be generalized to other scenarios? That's pretty much the question that we have.

DR. VENITZ: The question for the committee is, can this utility approach be generalized to other therapeutic areas?

DR. SHEINER: Well, why isn't it just about as general as it gets? This is one of those things that really is -- you know, it applies to drugs and automobiles and airplanes. How can you make any decision if you don't know what you're deciding about and what your values are and what the likely state of nature is? That's all we're really talking about here. Then after that, it's computation.

DR. VENITZ: I obviously second that, but I'd also like to maybe point out an approach of how we can convince other stakeholders that this is something useful. And it goes towards identifying utility values generically for certain kinds of adverse events, just like you would assign generically speaking for certain efficacy, life-threatening disease versus quality of life changes, and get agreement on that regardless of what drugs you're looking at rather than what you presented are specific to that drug where a decision was made this is how I define my utility function relative to my target, which makes it vary case by case.

But I think if you want to get consensus, let's start on the safety side. We agree certain adverse events are going to get certain utility values associated with it regardless of what causes it. By the same token, on the efficacy side, certain disease interventions get certain

utility values assigned regardless of whether it's a drug treatment, device treatment, or whatever it is. Then at least there's some common acceptance I guess on what the ranges are.

Otherwise, you're really getting into this swamp of, well, everything is a case to case. Show me the drug and that's the only way I can come up with a utility function, which defeats the purpose in my mind at least of using utility functions.

DR. LESKO: It just strikes me what you said would lead to agreement particularly on the safety side. I think you can identify certain safety signals that people would agree are bad and are most serious, and then weigh those, in turn, against a range of efficacy or benefits that one might get from a drug and against the loss of efficacy if you were to somehow misappropriately adjust the dose.

I think you can define boundary conditions in a sense with regard to certain safety signals, with regard to certain categories of drugs for efficacy. For example, hepatic toxicity, QTc. I think you'd get general agreement these are bad. We focus on those extensively irrespective of what the efficacy side is offering. By the same token, a disease state where efficacy is of extreme importance, there may be a different view of the safety signals. But

you need some anchor points, it would seem, and they might be those boundary conditions, and then the gray areas fall in between somehow.

DR. VENITZ: But if you look on the adverse event side, in oncology there is already a consensus of how to categorize and how to rank order certain adverse events according to organ function. It's unanimously accepted by the whole oncology community. Now, they don't necessarily use it in the context of utility. They use it to define dose limit of toxicity.

But why can't we do that as a general approach to adverse events? And you would do that regardless of the underlying cause of that. Whether it's a specific drug or other drugs, it's irrelevant. To get out of this discussion where everything is case by case, the moment you do that I don't think the utility approach is going to work. It might work on specific drugs, but it wouldn't work across the board because you've got nothing to compare to.

DR. KARLSSON: Although I agree and I like the idea. Maybe one complicating factor is those types of grading 1 to 4, which I assume you're talking about, are based on outcomes, aren't they? Whereas, if you have like a biomarker for a safety event, that's something a bit different and maybe more difficult to value.

DR. VENITZ: Maybe.

DR. FLOCKHART: I'm just thinking about something I know more about which is the QT interval. So the question would be, is a given QT prolongation, a prolongation of 10 milliseconds, by one drug the same as a QT prolongation of 10 milliseconds by another drug? Unfortunately, the answer is no because drugs do more than one thing. They have more than one mechanism of effect on the QT interval. For example, many antipsychotic drugs that affect the QT interval are anticholinergic as well, so they have effects on the heart rate, and that might be a little bit protective, make them a little bit tachycardiac. In situations where they got more drug, they would not only get more IKr blockade, but they would get more mu 1, M1, receptor blockage.

So because of that, I think hepatotoxicity gets even more complicated. If one drug causes X amount of change in the LFT's, is that the same as another drug causing the same? It's a very complicated thing.

Nevertheless, I think the effort is probably noble. It's worth venturing along that path at least to find out how different things are. I'm intuiting that QT interval drugs would be different. I don't really know.

DR. VENITZ: But I think the current --

DR. FLOCKHART: You need an outcome, though.

You need an outcome, and that's a very important point.

DR. VENITZ: As far as QTc is concerned, the current assumption is that QTc prolongation is bad regardless of what causes it. The relationship then really is what you're talking about, how predictive is a biomarker of a clinical outcome? Because the biomarker itself is not bad. It's the clinical outcome. You're worried about the fatal arrhythmia. So in addition to your utility, now you have to look at what are your uncertainties involved in linking that biomarker that you're measuring to the final outcome. In other words, for a given change in, I don't know, 40 milliseconds in QTc, how many people are likely to develop fatal arrhythmias.

DR. FLOCKHART: Yes. You can put a model on it, and you can also model in things like there are a significant number of drugs that prolong the QT interval for which torsade has never been reported, and we have them on the QT.org website. You could model that in as well. That's a chance that a drug in that class would not do it.

DR. VENITZ: But from my perspective, that's not a utility. That has something to do with how your biomarker relates to the clinical outcome.

DR. FLOCKHART: Right.

DR. VENITZ: Because really what you're assigning utility to is the bad outcome and fatal

arrhythmia in --

DR. FLOCKHART: Well, it is affecting the utility, though, because it alters that number.

DR. SHEINER: (Inaudible.)

DR. FLOCKHART: Dr. Sheiner was pointing out that it affects the expected utility, and he's right.

DR. VENITZ: Any other comments or more specific questions of either Dr. Karlsson or the FDA?

DR. KARLSSON: I was just going to add to that that maybe if you do start with a utility function that incorporates many different aspects of the drug therapy, you will find that actually in the end it's only really a few of them that are going to be important and you could reinspect those and the assumptions that go along with those particular events.

DR. SHEINER: Let me just ask. Some of these examples were actual examples from your interaction with the industry. I guess we can't draw too much of a conclusion from anecdotal experience, and also these folks asked as opposed to having it pushed at them by a regulatory agency.

Is it your feeling that this exercise was appreciated, informed people, and had some consequence in terms of the way in which the development plan went forward?

And finally, do you have any information on whether or not the regulatory agencies with which these companies dealt -- how they responded to the justifications offered through this means?

DR. KARLSSON: I don't think any of these drugs have gone to the regulatory authorities yet, although I might be wrong.

For the example that I presented, the AstraZeneca one, they were very helpful and really appreciated all our efforts.

In the other cases, I think it was more add-on and maybe it was not the core of the project team that was really wanting to have this information. It was more a side effect of having the possibility of doing it, and they thought maybe it was nice to know, but to what extent it influenced their decision I'm not sure.

DR. DERENDORF: A follow-up question. How many times do you think this is done a posteriori so that the decision is made first and then you need a justification for it?

DR. KARLSSON: I don't think that's done very often. I don't know, but in my experience outside the examples I've been involved in when I've actually done this, I haven't heard these kind of discussions going on with people in industry.

DR. LESKO: Just following up on the same question. Maybe this isn't a fair question, but I sense, at least in our regulatory agency, a stronger desire to understand dosing strategies than there has been in the past. It's partly related to a lot of things.

The question I sort of have is a follow-on to one that was asked, and that is, is there anything in your experience in working with the method, within the context of drug development, that would cause any concern on the part of a sponsor with regard to what a regulatory agency might ask of this kind of methodology?

It would seem like it would be received rather positively because it provides a fair amount more than we normally would see in terms of a dose justification or a rationale for dose selection. Therefore, my sense would be it would received positively.

But on the other side of that coin is anything different is different, and what are the issues that somebody might think about with regard to if I presented this, the regulatory agency might do X, Y, or Z?

DR. KARLSSON: Well, I've heard both arguments, both that it's good to do your homework and to be able to know your drug so you can argue well. I've also had the response that let's not do it because we're waking the sleeping bear, or whatever your expression is, that maybe

they wouldn't think about this if we haven't done it.

DR. DERENDORF: Well, but for internal decision making, I think it's always helpful to do it.

DR. KARLSSON: I think to some extent, though, it goes hand in hand. If you don't want others to know it, you don't want to know it yourself.

(Laughter.)

DR. VENITZ: Any final comments?

(No response.)

DR. VENITZ: Thank you again for your presentation.

Then we are going to start with our second topic and the last topic for today which is the pediatric database. I think Dr. Lesko is going to give us an introduction to that topic.

DR. LESKO: Thank you, Jurgen.

This is going to be a rather brief introduction. It's to lay the groundwork really for the next two presentations. The focus of this now, switching gears, is pediatrics, and we're going to be talking about two topics. The first is going to deal with a template or standardized approach, if you will, for pediatric studies that utilize sparse samples. The second is to discuss with the committee some ideas that we have for mining the pediatric database that we have as a result of the

pediatric rule.

Let me start with that. The pediatric rule is something many of you are familiar with who were actually here at our last meeting because we discussed it in some detail. We shared with you the types of studies that we've received under the pediatric rule. Basically the rule encompasses the idea that we want to bridge the adult data to the pediatric situation, usually not directly but through the initiation of some studies in the pediatric population, and depending on the nature of the adult data and the nature of the assumptions that we make in doing that, the types of studies that are conducted for pediatrics would be either efficacy, safety, or pharmacokinetics. We presented last time a pediatric decision tree that sort of guided the thought process on what studies to conduct.

Basically the idea with the pediatric rule is to avoid large-scale studies. It's considered inefficient to redesign the efficacy and safety trials when there's a preexisting database in adults. Although the challenge from a clinical pharmacology perspective and a clinical perspective is to decide what studies are, in fact, appropriate to conduct and will provide the most information.

The pediatric rule, as you know, is intended to

speed up access of drugs for children, to do that in a cost effective way without reinventing the entire efficacy-safety spectrum. I think overall I think most people feel this has been generally successful in meeting the goals that were intended for it.

But some questions always need assessment within the context of the pediatric decision tree and the pediatric rule. Is it reasonable to assume a similar PK/PD relationship as exists in adults? This sounds very familiar to the questions that we were discussing this morning in terms of different populations, and probably the underlying principles are very much the same.

We feel there's a need -- and we've begun to look at this -- a need to develop standard methods for answering this question for specific drugs and drug classes particularly where we've seen submissions in pediatric patients in drug classes over a number of drugs.

The other question is what is the appropriate dose. And to answer this question, we generally rely on pharmacokinetic studies. They can be of two types: the full exposure, sample-rich type study design, or the sparse-sample, population PK approach.

The overall goal of these studies is straightforward, to achieve a dosing that's intended to achieve exposure similar to adults. Generally PK studies

aren't the only studies conducted in pediatrics. Frequently a safety study or in some cases a small efficacy study is also required.

But it's probably safe to say that a sparse-sample strategy has been under-utilized in pediatric studies. We have some ideas why that might be. It might be that there's an insufficient understanding of this approach. It might be that there's some concern that the regulatory agencies won't accept such an approach. But it seems like there's room for opportunity here to do these types of studies that call for pharmacokinetic information in a sparse-sample strategy that is based on good principles.

So the first thing we wanted to talk about today was a discussion of a standardized -- and I use "standardized" very generally, but a standardized study design template that would be useful, for example, in communication between investigators or between companies and the agency to agree on a sort of starting point for designing both acceptable and informative PPK type studies. So that would be the first thing we're going to talk about, and Dr. Peter Lee will talk about that primarily.

The next topic is a follow-on to what we talked about in October. We had made the point and we shared with you the specific drugs in October for which we have a

database on pediatric studies. We raised the question about if this were your database, what would you want to learn from the database and what would you look for. What would be the questions you would have? We didn't spend a lot of time with that. We have more time today. But we've begun to look at this database and you'll hear from another individual from our office, Dr. Gene Williams, who is on detail to the immediate office of OCPB to specifically look at what we can learn from this database.

This is a work in progress. We began to assemble the database, in an organizational way of age groups, the PK data that we have. We've begun to look at individual drugs in the database, elimination pathways, the clinical endpoints that were studied.

I'd have to say this has not been easy. Unfortunately, this database is not in a form where you can just press a few buttons and pull it out. So there's a fair amount of up-front work that goes into assembling the database. I think in recognition of that, it's important that we understand where we want to go with the questions that will derive from this effort to organize the database once we start moving in that direction.

So Gene is going to do an overview of some of the research objectives that we have, with the goal of generating knowledge from this database. What we think

we'd like to do is look at underlying mechanisms where there are exposure differences between kids and adults, look at breakpoints perhaps on age, look at specific elimination pathways.

The reason we want to do all this is it's sort of common sense. We don't feel we should be asking the same questions now of pediatric studies that we asked three years ago before we had 50 or 60 studies in house. So the idea is to learn from the database and then use that information and knowledge to revise our pediatric context or pediatric decision tree for the future.

So the goal of this effort: to improve or revise our pediatric decision tree based on identifying better studies that we need to conduct in the future -- and by "better," I mean more informative studies -- or perhaps reduce the number of studies in the areas that the data would allow. So for that, Gene Williams will present an update.

And Peter is first.

DR. VENITZ: Lew.

DR. SHEINER: Can I just ask you a question, Larry? On your third slide where you have "some questions always need assessment," and you say, "is it reasonable to assume a similar PK/PD relationship as adults," I understand that that's not particularly the focus of today.

I'm sensitized to this because of some work recently that I've been doing with a topic that's related to this with Novartis.

It's an interesting question. What would constitute evidence that children are like adults with respect to PD now, not the PK part? I just wanted to put it on the list of many things that you have to think about, that maybe we ought to address that at some later point, or maybe you might want to have us address that at some later point. It's a vital question. Obviously, if what it takes to establish the similarity relationship is more work than it takes to clear a drug to start with, you're done.

DR. LESKO: Yes. I think it's a vital question. We have a few instances, a few drugs in hand where we actually have information both on the adult where that happened to be done in the NDA and it was also done in the kids. We're going to focus on a few of those examples to try to answer that question.

We've often turned the question around and asked how much of a difference would be important in the PK/PD relationship and would that warrant necessarily a different dosing strategy in kids. These are open questions, but we do not have a lot of opportunity to look at this issue based on, at least, what I've seen of the database to date. Maybe Gene will comment a little more on

that.

DR. SHEINER: Right. It does relate to that. But it's also a question of style in the following sense. If the data are sort of anecdotal, are they really data? So the issue becomes -- when you say I think that the relationship is the same, and then I say to you, okay, here's some evidence, and you say, well, that's not really evidence because that's physicians' opinions, let's say. So the question is what constitutes evidence that it's the same, given that you're inclined to believe it. That's the issue I was getting at.

DR. LESKO: Right. That's a good question whether it's in the statistical domain or whether it's in the modeling domain. I think we need to talk about that, but I think that's a good open question for another discussion.

The point of bringing that to the committee might be when we've done our analysis of the database as we have it and share with the committee what we've learned from the data that we have and maybe what we might want to know from data we don't have at the moment but might recommend somebody begin to look at, not necessarily in addition but maybe as an alternative to the studies that are being conducted today.

DR. VENITZ: Peter?

DR. LEE: What I'd like to present to the committee today is a proposal to develop a pediatric population PK study design template. I believe we have sent a copy of the full proposal to each committee member a few weeks ago, so hopefully you would have had a chance to look at the proposal already. But what I'd like to do in the next few minutes is to just go over this on the key points in the proposal.

The objective of the pediatric PK study design template is the following. First, to provide a consistent approach, like Larry mentioned earlier, to design and evaluate a pediatric PK study. We'd like to develop a computer-aided pediatric study design template which will take the user-supply study design and automatically estimate the study performance and study power. So it will be making it easier for the user to determine which type of study design is appropriate for their drug and for their design.

We also want to select case studies from the FDA database to test the template and refine the template.

Finally, through this template, we hope to promote a wider use of population design in pediatric studies.

I think Larry has mentioned this pediatric study decision tree. He also presented this decision tree

in the last meeting. Basically we use this decision tree or propose to use this decision tree to determine what type of study will be necessary to bridge the adult efficacy and safety data to pediatrics. Depending on the answer to many of these questions, we may end up doing or recommend doing a efficacy study or clinical study, a PK study or a PK/PD study, or all of the above, or safety studies.

But today I just wanted to focus on this particular box here which is we will recommend doing a PK study as a bridging study between adults and pediatrics.

So once that decision has been made, then we will have to use the PK study for dose selection in pediatric populations. So this goes back to the dose adjustment in special populations, the same decision tree we talked about in the morning today.

Based on this decision tree for dose selection, we have to answer several key questions. First, we have to ask whether there is a clinically significant difference in pharmacokinetics between adults and pediatrics. Secondly, we have to ask what is the pharmacokinetic parameter in pediatrics, and we had to use that pharmacokinetic parameter to adjust the dose in that population.

So based on this decision tree, we had to conduct the pediatric population PK study to get two information. First, we had to identify whether there's a

clinically significant difference -- not just any difference, but clinically significant difference -- between adults and pediatrics. Secondly, we had to accurately estimate the parameters in pediatric populations without any bias.

There are, of course, many factors that may influence study performance, population PK study performance. For example, the study design which may include a number of subjects, demographic information maybe, the number and timing of samples. Also, the study conduct, such as the compliance of the patients, the variability of dosing time and sampling time and in missing dose and in drop-off. Of course, the pharmacokinetics itself and the variability of the pharmacokinetic parameter also influence the study performance or how we design the studies.

In the proposal, we also bring up several important points to be considered during designing a population PK study. We believe that when we design a study, we have to take into consideration the objective of the study which I just mentioned to identify the clinically significant difference and to estimate the PK parameter in pediatrics.

We also believe that because the pharmacokinetic parameters are very different between

drugs, also there are many different varieties of population PK designs that will provide sufficient study performance to achieve the objective that we just mentioned. There is no one-size-fits-all design for population PK. Each design should be looked at on a case-by-case basis, but using a consistent approach which we believe a clinical trial simulation will be a good practice to estimate the study performance and study power.

In our proposal, we also bring up several other points to consider. For example, we had to look at a number of factors that may influence study performance, such as dosing time, sampling time. Compliance is an important factor. Of course, the number of subjects, the number of patients, and how the sampling time and dosing time is distributed with time. We also need to consider if there's an unbalanced design in terms of number of subjects as well as the number of samples between different populations.

This is a flow chart of the proposed study design template. It consists of a module where the user can input their study design protocol. So this is where the user can enter the number of subjects, the number of patients, when do you take the samples, and what is the variability of pharmacokinetics, and so on and so forth.

So this will be your input parameter where the

user can put into the module. And the template will take this information and automatically generate or translate the study design template into a simulation code. With that simulation code, the software will generate a study performance indicator. In this case, it will be the power to determine whether there's a clinically significant difference in pharmacokinetics between populations, also what is the accuracy and precision and bias of the parameter estimations.

So the input and output of the proposed study design template includes the following. The input will be a study design, the pharmacokinetics in adults, but also the variability of the demographic or patient population you will include in the studies, and also the study design variables. The other input parameter will be the criteria for evaluating the study performance.

But the output from the template will be the estimated study performance which is related to the two objectives of the population PK study. One is the power to identify -- I want to emphasize -- the clinically significant difference, and secondly, the precision and bias of parameter estimations.

The clinical trial simulation that we propose to be used in the study design template is pretty standard. It includes the following steps. First, it will generate

demographic variables and pharmacokinetic parameters based on the user input information. It will simulate a study design as well as the study conduct, and it will generate population PK data. Once the population PK data is generated or simulated, then we will conduct a population analysis based on the simulated data. And finally, we will repeat the process perhaps a few hundred times to a few thousand times to estimate the power of the study as well as the precision and accuracy of the parameter estimations.

Like I mentioned earlier, there are two main objectives to measure the study performance. The first one is to identify a clinically significant difference in pharmacokinetics. Based on the decision tree or dose selection that we had presented in the morning and early in my slides, the first option is to just look at a 90 percent confidence interval and see whether the 90 percent confidence interval of a PK parameter is within a default boundary.

But the second option which we might prefer is to determine whether a clinically significant difference in pharmacokinetics will exist. In order to determine the clinically significant difference in pharmacokinetics, of course, we had to know the PK/PD relationship and we have to assume that the PK/PD relationship that we have perhaps from the adult population is similar to those in the

pediatric population.

For example, in the simulation we can assume that there's X percentage which is considered a clinically significant difference in the body weight normalized clearance. With that difference, we want to ask the question whether the population PK study that we tried to design will be able to capture this clinically significant difference. So with that we can calculate the study power of the population PK studies.

Now, the second criteria for study performance we propose is the precision and bias of PK parameter estimations. Precision and bias can be presented in terms of a percentage prediction average. Precision will be the standard deviation of the prediction errors and bias will be the average of the prediction error. The prediction error is defined as the percentage difference between the true value and the predicted value because we're doing a clinical trial simulation, so we know exactly what the true value is.

So the output again are two information to measure the performance of the study. One is whether the study has a power to identify a clinically significant difference. The second is precision and accuracy of the PK parameters.

So this is basically a general description of

the proposal. Again, the detail is elaborated in the actual proposal itself.

I guess we'd like to ask two questions to the committee. The first question is, are the proposed objectives for pediatric PPK studies reasonable, considering the decision tree for the dose adjustment? So we have mentioned two objectives for the population PK studies. We also talk about criteria for study design performance.

The second question is, is the proposed pediatric PPK study design template reasonable? This is related to the clinical trial simulation approach that we propose, as well as the factor to be considered, a study design factor.

We also would like to ask what feature we should include in the pediatric study design template so that we can make it more user friendly and more useful for both the clinician and the reviewer who might use it to design pediatric population studies.

That's it.

DR. VENITZ: Thank you, Peter.

You have the questions for the committee. Any comments by the committee? Hartmut.

DR. DERENDORF: I have a question for clarification. On your decision tree, right on top you

said it's reasonable to assume similar response to intervention. Then you have yes, and reasonable to assume similar concentration response in pediatrics and adults. What's the difference between the two?

DR. LEE: I think one example was asthma. For example, it may be a different endpoint that can be measured in a pediatric population than an adult population. That will be the first block of questions. So if we can answer that question yes, then we go to the next block. But if the answer is no, then we have to go to the clinical studies. So if the answer is no to the first two questions in the top block, then basically either the clinical endpoints are different between the two populations or the disease is totally different or disease progression is totally different in the two populations.

Now, the second question on the right-hand side of the block is that once we decide that through our clinical opinion the disease and disease progression and endpoints are similar, then we ask in our database do we know there is a PK/PD relationship and can we assume that a PK/PD relationship is similar between pediatrics and adults. For example, for a proton pump inhibitor, we know the PK/PD relationship of gastric acid versus drug concentration.

Now, can we assume that the relationship we

have seen in the adult population is the same as the relationship we will see in the pediatric? If the answer is yes, then we can just rely on PK information to determine the dose selection in the pediatric population.

DR. JUSKO: A comment and a couple of questions. This, in general, seems like an excellent idea. Of course, we always want to utilize information we know ahead of time to anticipate the study design and changes that will occur in a new group for study.

One small point that seems to be totally missing is dosage form and bioavailability. Young children are typically getting liquids and chewable tablets, and in order to anticipate what will happen in the kinetics and dynamics in young children, you may need a comparable database in adults with a similar dosage form or at least make an adjustment for perhaps faster dissolution or absorption. So at some point some reminder of that question may need to be added.

Then there's a certain vagueness when you talk about dose adjustment or dose selection because in pediatrics there's always dosage adjustments. Children are seldom given the same 500 milligram tablet that adults are. So something needs to be said about what do you mean by dose adjustment. Are they getting certain milligrams per kilo already or certain dosage sizes depending on weight

ranges? There's a great deal of flexibility already inherent in selecting dosages in children.

DR. LEE: I think the answer to the first question is I think we are talking about dose selection because we know that we're going to adjust the dose anyway. We're going to give a different dose. Normally the dose is given on a per body weight basis or sometimes we will look at the body surface area.

DR. LESKO: It depends on the age group. Sometimes the dose is adjusted based on average exposure to the drug, and depending on the age, it may be down to a milligram per kilogram basis, something like that.

But I was thinking of the first question on the response that Hartmut asked and I don't know if Peter clarified. But I was thinking basically you're asking the question is the response the same that you'd be interested in in kids as you would be in adults. That's sort of like a two-part question.

So the first question is, is the response, let's say, FEV-1 in an asthmatic patient the appropriate response? And that sort of gets to the heart of the mechanism of action of the drug and the progression of the disease similarity. Often data to support those assumptions or answers is not available, but if you do agree that, yes, I believe that's true, then the question

is, is there a concentration related to that response that you previously agreed is similar to the adult? Then that puts you further down in the decision tree.

DR. CAPPARELLI: I just had a question and a comment as well. In regards to the scope of this, is this looked at as part of a safety study as well, or is this really structuring a population PK study with the only endpoints being population PK? Because I think one of things that's missing in the objectives is getting at the question that we've been asking how do we assess these potential age-related exposure-response relationships. And the design that may be very robust for estimating the PK parameters of the population may not give you all of the estimations in the individuals that you may want to do some of that exploratory analysis. So I think that it needs to be very clear along those lines.

It was brought up before, and I spoke with Dr. Lee before a little bit about the concept that, really, kids are different. The question is whether or not we can predict those differences a priori. So saying are there differences, there are always differences. The question really is that based on our knowledge of modeling, our knowledge of pathophysiologic changes, developmental changes, can we predict those well and then go from there.

I also just wanted to amplify Dr. Jusko's

comments about the dosage formulations. A lot of the exposure issues are going to be based on what size doses you have. So when you say clinically significant differences in PK, it's really what sort of exposures we're going to get out of the available dosage forms. There may be a modest change in PK, but because of where you're left with your cut-points, you may end up having big changes in dosage exposures which again it needs to get back to, I think, what we're interested in and at least getting the exposures as comparable when we don't have the information in terms of differences in exposure-response relationships.

DR. RELLING: Is it implicit that it's only worthwhile to do these pharmacokinetic studies in children if there aren't good a posteriori methods of dose adjustment based on more readily available clinical measures like blood pressure, like immediate response to anesthesia, like pain relief? Is it any part of your interest in the pediatric rule that pharmacokinetic studies be performed for drugs for which there's a narrow therapeutic range or small therapeutic index and there's no other good way to adjust doses?

DR. LESKO: No. I think the pharmacokinetic studies are routine in these types of situations. I don't think dose adjustments -- I'm trying to think if I have any experience with dose adjustments being made without the

availability of PK information, for example, being based on observed responses in kids relative to adults. Is that sort of the question?

DR. RELLING: My point is that I feel that there are a lot of studies being done by pharmaceutical companies, which they claim are being done to fulfill FDA regulations or requirements or suggestions, which are done for medicines that don't need to have pharmacokinetic studies done. They're done for anesthetics that could be easily titrated based on the response of the patient to the oxygen saturation. They're done for narcotics for which the drug is going to be adjusted based on pain response. And I feel like a lot of resources are being expended on these studies for unclear reasons. So I'm trying to figure out is this a suggestion that's being made for all medications that would ever be used in children regardless of the therapeutic index and the ability to adjust doses based on other parameters besides PK.

DR. LESKO: Yes. I'm not sure "all" is the right word, but I'm thinking of the implications in labeling the drug product with a starting dose in pediatrics. You need to have some information, it would seem to me, to begin dosing, and that generally has been the pharmacokinetic studies to recommend some changes in that initial dosing strategy.

This isn't to say this is the only objective measure that one looks at in the pediatric area. There are always other studies, in particular safety studies and, in many cases, small efficacy studies as well.

DR. KEARNS: I want to thank Dr. Relling for her insightful question because everybody wants to know the same thing about children and that is how much do I give, do I need to give a different amount as the child gets older, and will it work like I want it to. That's really crux of all this. I won't belabor all my soapbox points about this issue, which are many, but I want to make two points about this recommended approach.

First is the issue that Dr. Sheiner said we should grapple with at a later date, but it's at the top of the decision tree, and that is, do we believe that whatever condition occurs in a child it is substantially similar? And that's the language in the law, "substantially similar" to what it is in an adult.

For the last four or five years, I have seen people at all levels of the agency grapple with this as though it were a very large, mean animal, and at the end of the day, people rather than slay it, seem to run from it and invent ways to try to avoid it. I think that's tragic because what that has produced is an incredible consumption of resources, not to mention the exposure of children to

clinical trials that is a needless exposure. That doesn't get talked about enough, and I'll stop talking about it now because that's probably another whole day.

Now, with respect to "substantially similar," what comes to my mind is something very similar. I'm starting to sound like one of our politicians who uses the same word over and over. There's an article in CP&T, Art Atkinson. It's near the front of the journal this last month on biomarkers. It talks about the goodness or badness of biomarkers on how far away they fall from the trunk of the tree of drug effect. The same thing could be looked at with respect to this issue in children. Drug action is obviously something we can, at times, determine whether it's similar. If we can't talk about action, we can talk about drug effect, a physiologic response in an association with a drug dose or a concentration or an amount. If we can't talk about that, we can then talk about disease response and lastly disease progression.

Sadly, I hear people put disease progression at the top of the discussion list as they look at that box because if we were to get a bunch of pediatricians in a room and ask them if they agreed that the progression of GERD was the same in adults and children, they would never agree. They would never agree about asthma. They would never agree about leukemias, other malignancies. Pick a

disease. No one will agree at the end of the day. Which means then you punt the ball, and the ball is punted in terms of time, dollars, delay, and for all the reasons that people line up on the opposite side of the pediatric argument to say it's bad things to do, it gives them fuel for their fire.

So I think it is incumbent upon the agency and those of us who you've elected to advise you to, at some point in time, grapple with what is substantially similar so that any well-designed pharmacokinetic approach can get on the right track and do what it's intended to do. So I applaud Dr. Sheiner for suggesting that and hope we can talk about it.

With regard to population pharmacokinetics which is, Peter, I think central to your presentation, my question is always the same, and that is, when we use a pop PK approach in a pediatric study, are we aiming to explore relationships with age or, alternatively, are we aiming to define them? I think we certainly can use pop PK, appropriately designed, to explore them.

But keep in mind that for those drugs where age is an important covariate with respect to metabolism or perhaps pharmacodynamics or response, an exploration and a definition can be very, very different with respect to impact because at the end of any pediatric program that's

conducted, we are trying to do things on the quick, on the cheap, and on the small. That means the generalization that we want for an entire population of patients that represents about 15 percent of the population of the United States is predicated on a fraction of the numbers of subjects and things that we would do in adults.

So what you're proposing is very important, has incredible potential if done correctly, but we've got to be mindful of knowing how that tree starts and making sure that at the end of the day the people at level of the review divisions and the Office of Pediatric Therapeutics understand the power of this tool and how it can help them as opposed to what's going on now in the area of PPIs -- I hate to harp on this, but it a plays a nice tune -- where all the things that people around this table know, if you go into the little, bitty room downstairs on the third floor and you listen to the recommendations to a sponsor, you would have thought you woke up in the stone age. The magic is still very much there, and this approach has to be used to make the process better.

DR. SHEINER: I'm glad we're moving off in a direction away from the techno-nerd thing.

Mary, at the risk of maybe setting you up as a straw man, let me just think about what you just said. I'm going to take away the word "pediatrics."

Why wouldn't everything you said apply to a drug that's easily titratable no matter who it's intended for?

Now, I think Larry's response to that was, well, we need a dose in the label somewhere. It's got to be based on some kind of evidence that people can refer to.

So I think, if I can rephrase your question, what I think you're saying, to put it in sort of a Bayesian context, is there's something about having studied dose-effect relationship in adults that for a drug that may be isn't too toxic and is easily titratable with respect to effect might mean we don't have to do any more than that for children, which means, in a sense, that there's a prior somehow on the doses that you ought to be trying in children because, after all, the implication of an easily titratable drug that's not very toxic is that you don't have to get the first dose very right. So you're saying I'll tolerate a much wider range. And there's something about having studied it in adults that gets me close enough to that range that if I were to work it all out with a utility function and put a prior on what I think I can extrapolate from adults to children, I'd find I don't have to do a study at all. I'd find my net benefit to society would be better served by going ahead and approving it for children than doing any more studies. I think that's how

I'd try to put it in a quantitative context.

All I'd like to say about that is we could do that. We could put it in a quantitative context so you don't have to sit there and feel a little uneasy about what other folks are suggesting and Larry doesn't have to feel a little uneasy about what you're suggesting. We can put it all together and, just as we were talking about earlier with Mats' presentation, take a look at it. What values do we need to place on doing -- what negative utility do we need to place on doing studies in children in order to make it worthwhile to extrapolate, given our sense of uncertainty from adults to children, and conditional on this drug, its safety profile in adults, et cetera?

What I'm trying to say is I think that the same way of trying to be quantitative about this could answer that question. I think it's just like the drug companies ask themselves now. Am I better off going right to phase III and skipping phase II? And there's a decision analytic argument that might say that that's really, in terms of net present value, et cetera, a good idea for this drug in this circumstance given what we know about it and its competitors and the fact that it's quite similar chemically and so on. And all that can be worked out and it doesn't have to be an opposition of people not really understanding each other or thinking that they're coming from different

places. They're not. They may value some things differently, but I doubt very much.

And that's a worthwhile exercise, it seems to me, to do as we think about going to this in the future to try to approach this whole issue from the way in which we're saying people might approach the much simpler and therefore easier metaphor of a dosing decision, this whole issue of do we do a study at all, because it's really all the same. It will get us into the habit of thinking that way in a place here where we've got some time to think rather than having to act right away.

Peter, I wanted to get to your question. It's really I guess my question about your question. It's really a question for Larry, and it's this one.

You called it a template and maybe words are important here. But it sounds to me an awful lot like a piece of software. At the sort of highest level, does the FDA want to be writing software for people who are then going to probably feel obliged to use that software because the agency says this is what you should do to design the study that you're going to do to come to us with? I mean, the agency I think has generally been pretty careful about that and has had best practices and guidances and suggestions and all that kind of stuff. But here's something we made and it's for you. That's a whole new

line, isn't it? And do we want to go down that line?

DR. LESKO: Yes. I think you're right in pointing that out. I don't think we're talking about, at this point, the software that we're either going to develop and advocate, advance for drug development. I think what we're talking about is a template that is based upon software that a reviewer might use in conjunction with a discussion with a sponsor that would prompt for the critical information that would go into making a robust study. It's intended to sort of be a starting point for discussion or designing such a study that would be, at the end of the day, generally acceptable to the people that need to accept it in terms of its review and utility.

I think the problem now is we don't necessarily see a consistency in advocating the design of these studies across different opportunities to do so. This represents a way of channeling the discussion into the critical areas that would lead to usable results.

DR. SHEINER: I like it. It's a good metaphor, but I think you may be getting too specific too fast. It's sort of like a guidance, a statement about what are good things, what you want to see, what kind of principles you apply, and it's got a lot of wiggle room in it. The thing about software is it hasn't got any wiggle room at all. You put the inputs in and it's deterministic; out comes the

answer. So you have got to be pretty sure that's the algorithm you like. So I'm not sure I would start there.

But taking the metaphor of designing software to say that -- and that gets us to focus on all the key issues. That's, I think, a good idea so long as at least you're contemplating maybe not going the whole route and going into the software business.

Then I think fundamentally the questions are right, the first one being what's the minimum evidence I can gather that I ought to bother anymore. Is there any difference? But I don't think you'll get that from the same study. As I say, I think this is essentially sequential. That's the difference that I'd have between the way you put it. It's not going to be one study. There's something you do to figure out whether I need to go any further, and that may be quite different in design -- although I haven't thought this all out -- than what you do when you say, oh, I guess I better go further. I better pin down the key PK parameters in this population and how it varies with disease state and other things. My suspicion is that those will be two different activities.

DR. LESKO: I think the problem with the studies that have been done -- and it's probably why we haven't seen very many of this sort -- is the believability of the outcomes because these studies are not as well

understood as obviously a sample-rich study design and the issues that go into analyzing a sparse-sample study using NONLIN or something like that. Is the information reliable and how do I know that? It's having to explain that to people who have to make decisions over and over or having designs that would lead to an acceptable result is sort of where we're heading here.

That being said, if we have an optimally designed study to get at the questions we're asking about differences in pharmacokinetics for the purposes of dosing changes, can this method be confirmatory enough to stop there? I think Greg used the word "exploratory." I don't know if that was a suggestion that these studies at best can be exploratory for the purposes of designing another study or would they be confirmatory enough to say I know what the difference in clearance is between this drug as a function of age and maybe more age groups if I can do a sparse-sample approach, and thus I can recommend some different dosing for these age groups based on the study

cases limits the number of age groups that can be looked at.

DR. KEARNS: But, Larry, I think there are some examples on the books where it does work. The whole program on montelukast to me has been an incredible success

story because a very careful pop PK approach was taken. We went down through all the pediatric populations now down to 6 months. From my opinion, the appropriate variability was considered and the parameter estimates seemed to hang together, and when the data were taken the next step into showing proof of concept with respect to effect, the effect was there. Consequently, the labeling of the drug has been changed multiple times. It likely will continue to be changed based on that approach. It worked. It was done right. There was a need to do it so we know the dose.

But you're correct in that many other companies have not followed suit, so to speak, and for reasons that I don't completely understand.

On the other side of the issue too, logistically -- and certainly Dr. Capparelli knows this because he's kind of in the business -- for the most part, if you have a pediatric study, a PK study, and you go to the trouble of obtaining repeated blood samples, you're obtaining them through a catheter. If you're analytical method is such that you don't need a lot of blood, the bother in getting eight samples is no more than the bother in getting three or two. IRBs anymore, at least pediatric ones, do not allow you to stick children several times. So there are many times when a pop PK approach could, indeed, be used and it would be perfectly valuable and valid. But

a traditional approach is very achievable, more so than most people think.

DR. DERENDORF: Just as a follow-up, the other technique that is coming on strong is microdialysis where you can take as many samples as you want without taking any blood out. So I think that will change the ability of doing studies in children.

DR. SHEINER: Just a quick response, Larry, to your question. The essence of having a credible confirmatory analysis is controlling type I error, and controlling type I error involves essentially saying what you're going to do before you do it because you can't have feedback from the data to the analysis. That doesn't mean you can't get valid conclusions from doing that, but you can't control type I error if you do that. So I think that's the only issue.

If you do a well-done analysis, then you know how uncertain you are when you're done. That's an issue of design. That is to say, given the assumptions we're willing to make and the data that we get, the sparse data that we gathered, do we wind up with sufficient precision on these things to make the kinds of statements we want to make? That's an issue that unfortunately there isn't a lot of theory for because they're complicated designs and they're complicated analyses, and so you have to do it

through simulation.

But the key thing would be specifying beforehand exactly what models you're going to use, what procedures you're going to do, et cetera. All of us in the business of doing extensive modeling have always felt a little anxious about that because you can't take away from me my ability to look at the data and decide what model I ought to use, but you have to take that away from me if you want to control type I error.

So I think there will be an interesting issue there of how you balance that and whether, in fact, controlling type I error is as important as you sort of said it is by bringing it up. I don't know. I think I'm probably with Mary and Greg on this one, that we've got a lot of priors behind us and I don't think I need to pin it down to a fare-thee-well.

DR. LESKO: One of the inputs into the model is the variability within the kids or within the age groups that are being studied, and I'm not sure how that's been handled. I can't recall the montelukast or, Ed, some of the studies you've done. But the variability associated with the -- what you would expect with your different age groups -- how was that generated and how important is that in terms of designing these studies?

DR. CAPPARELLI: Well, in terms of looking at

simulations and sort of real data that come out, the variability goes up as you go down the age group. So from a pragmatic standpoint, one starts with at least an adult value and goes up from there. But clearly there may be thresholds, some of them drug-specific and age-specific, in terms of dealing in HIV where we've got drugs that have major food effects, we've got formulation effects, and as soon as you cross certain thresholds, the variability is going to drop down. But it adds complexity both on the design standpoint and the analysis standpoint when you have observed doses where you've got your compliance and you see no drug. And this is in a CRC setting, but it's just the way that it behaves in this population.

So there clearly are needs for evaluating distributions much more intensively, especially if we're interested in sort of the outlying regions, which I think most of us are. But experience is that it's greater. It's just variable how much more.

DR. SHEINER: I've got to ask about that. I always thought the opposite. Once you line up kids by some maturational marker, whether it's gestational age or whatever it is, I thought they're all newly minted coins and they all look the same. In fact, I think I remember Bill Jusko saying that when you get real old, the variability goes up because you've run a longer race and

you're sort of stretched out there at the finish line. That's what I always thought the case was. There was more variability in old folks than there was in little babies, again lining them at the right maturational level. Obviously, a 3-month premature is not a term baby.

DR. KEARNS: No. Actually we're finding the opposite. It's quite interesting because if you look at what Dr. Sheiner said, if you look at a 3A4 substrate in a healthy adult, there's 20-30-fold variability in the processing. You look at it in a 3-year-old. There it's about the same. If you look at it in the 3-month-old, it's about the same.

The problem is that as that little beast travels into adolescence and adulthood, the shape of the acquisition curve, if you will, changes, not to mention changes in body composition which are quite evident. So there are a couple of moving targets that make it a particular challenge which, in designing a pop PK study, trying to estimate what your real variability is when you're up at the front, is not always an easy cookie to get.

But there's got to be a way to do it. I think what we hopefully will see, as we see the database, is some of the information that the agency is collecting is beginning to show us where these patterns might be, if you

will, or breakpoints might be, and that makes things a bit easier to deal with.

DR. DERENDORF: And if enzymes change like that, what makes us assume that receptors don't?

DR. KEARNS: Absolutely, right.

DR. FLOCKHART: The absence of a phenotypic probe for the receptor.

(Laughter.)

DR. LESKO: We'll talk about genetic solutions tomorrow.

DR. CAPPARELLI: I would also emphasize at least a lot of the variability experience where there are these major changes, besides the newborn, is when you get into oral drugs. So I think there are, at least from my experience -- again, the diet is different. Controlling for when they take it relative to food, all those things that may be a little bit easier to do in adults is much more difficult in kids. The formulations themselves -- while you can do bioavailability studies in adults and show similar formulations, it doesn't always extrapolate to kids. So you have those sorts of things, I think, contributing as well to the variabilities.

DR. VENITZ: Any other questions or comments?

(No response.)

DR. VENITZ: Then let's take our afternoon

break. It's now 2:35. So let's reconvene at 3:05, in 30 minutes. Thank you.

(Recess.)

DR. VENITZ: Let's reconvene our meeting, please.

Our next and our last presentation for today on the pediatric topic is Dr. Gene Williams. He's a pharmacometrics reviewer currently on detail at the Office of Clinical Pharmacology and Biopharmaceutics, and he's going to give us an update on the pediatric database. Gene.

DR. WILLIAMS: Thank you, Jurgen.

This is the title of my talk, kind of a long title. The notion is to become better at predicting peds clearance and to take advantage of what we usually know at the time that we see peds submissions or proposals for peds studies, that is, child age, a lot of information in adults, and the knowledge of in vitro metabolism.

I'm going to ask four questions of the committee. It makes sense to show them first so you know where I'm headed.

The first one is sort of an overall scope and method. That is, is the general approach that I'm going to suggest in the presentation rational and logical? And the approach proceeds from a very empiric method to a more

mechanistic method.

Secondly, is there anything special that you think that I should be aware of, some difficulties that you think I'm likely to encounter, and if you can identify such, how can I avoid them?

The third question is of particular interest. Are there data sources you could recommend? One committee member has already been referred to as "in the business." We'll get there I guess.

The fourth question I have is, do you have any suggestions regarding the form of the non-physiologic-based PK mechanistic models? That will become clear as I proceed, I hope.

What brought on this project and what exactly are we talking about trying to accomplish? What we'd like to be able to do is construct a model that allows us to predict pediatric systemic drug clearance from, as I said, adult PK and in vitro microsomal metabolism data. That would be a short-term goal. Obviously, we have a longer vision. It seems like if you could construct such a model, it would aid us internally. It would also be of potential interest to industry scientists, and finally perhaps even health professionals in the community could make use of such a model.

It's probably appropriate to begin talking

about the data that we have because that largely drives how we'll be able to model.

The most fundamental data unit we're talking about using here is clearance, whether it would be from sparse or dense data, and age for each individual in the data set. A number of demographic data also we would use; that is, the weight and height for each individual, renal function for each individual, and gender and race for each individual. Finally, as I've alluded to earlier, we would also want to make use of what we know about in vitro metabolism data for each drug.

I should probably add here that, as many of you I believe have appreciated, FDA is well positioned to do this sort of work because the data that we have often is very specific. We not only see data summaries, but we also see individual data, which is a limitation that if you use literature data, you face, but we often don't face. We usually get fairly raw data where we do have all these values.

Our data set. I've taken this statistic from a website that I've included here. I believe it's publicly available. I don't think it's just on our intranet. In mid-March, about a month ago, we had 72 active moieties that had received pediatric exclusivity. As Larry said earlier, most if not all of those would have pediatric data

available.

As I've gotten ahead of myself a little bit, the data that we see is usually raw. It's actual measurements of individuals, not summaries across individuals. And for the models that we want to explore, that's of a lot of utility.

Further, our data is usually reviewable to an extent that literature data is not. The analytical methods, dropout, salient features of data accumulation and choices made in data analysis are often presented to us. So we can do a good job of assuring data quality.

However, there are some limitations of the data we see. First, studies are often not powered to compare PK across age groups. People are submitting data to us for regulatory purposes, not always to discern carefully small age effects. That's in distinction to studies sometimes performed by academicians where they're specifically trying to see age effects or, I should say, reasonably small age effects.

The ages with the greatest difference from adults, often the very young, are often most poorly represented in the data sets that we see. The data sets we see are motivated by the desire to treat, not necessarily the desire to see an age effect.

Finally, most of the drugs that we see are not

probe substrates. People, again, are not asking mechanistic questions. They're trying to get a drug approved. So the ability to tease out effects may be more difficult since we don't have good markers for each individual effect that may be present. As a result of this, it may be necessary to use some function of in vitro metabolism such as the Km, that is, for an enzyme, as a covariate when we do our analysis.

I'm now going to carefully consider a data set that I took from the literature. I did this for a number of reasons. As Larry said earlier, organizing our data set is a considerable effort, and since this data set was sitting out there, I thought I'd use it not because we want to analyze it, but because it makes a good platform for discussing the methods that we intend to use.

This is taken from Ginsberg, et al. There are somewhere 21 and 27 drugs represented here. The y axis is children's clearance relative to adult. You'll see at the bottom of the slide I've described the units that are used. These data are standardized for weight -- they looked at kilograms -- and age. Age here is not a continuous variable. Rather, it was grouped categorically, a decision the authors made. They took these data from the literature. It's not their own data, and I guess the data lent themselves or, for some reason, they organized for

categories in this way.

I've shown a line at 1. That would be where the child is exactly the same as adult. You can see at ages 12 to 18, that's accomplished.

I don't want to give much attention to this slide, but the question is likely going to arise as to what sorts of drugs they were. This is also taken from their database. This database is available on line for anyone who wanted to explore it. As I said, I don't want to discuss this, but the drugs represented a number of classes.

Before you attempt to model these sorts of data, it's necessarily to normalize clearance. The reason why is you want to consider each drug on its own and not have your analysis complicated when you compare drugs whose adult clearances differ widely. So the method we chose to normalize clearance, similar to the method that Ginsberg used, is to divide each individual pediatric clearance by the mean adult clearance.

Again, this is Ginsberg's data. The y axis is clearance ratio versus age. However, unlike the plot that I showed you from their paper, this data has had the element of weight removed. So this is no longer adjusted according to the representative body weight of the data.

The line shown here is a simple least squares

fit, no weighting has been performed. This is unlike what we intend to do when we analyze our database. We'll probably use NONMEN extended least squares.

As you can see, or perhaps as I need tell you, I have fit the effect of weight on clearance in this plot. So the equation is shown beneath the line and it's a simple exponential relationship. The maximum ratio I allowed to happen was 1; that is, where the ratio of child to adult would be 1. So essentially this is one parameter. I fixed the maximum ratio.

As you can see -- I was somewhat surprised -- it provides a reasonably good fit. This is a little at early ages, and this is sort of consistent with what we generally expect to happen, that during development and maturation, things may be a little different.

DR. SHEINER: Excuse me. Just to clarify. I guess I'm not sure what it is. You haven't fit the data on the y axis to the data on the x axis. Your equation there is in weight which is --

DR. WILLIAMS: Correct, yes.

DR. SHEINER: So tell me again what I'm looking at.

DR. WILLIAMS: Indeed. I have not fit age here. I fit weight. So what I did is, although I'm representing it age because that's the thrust of our

interest, before I went there, I wanted to isolate the effect of age as opposed to the effect of weight. So first, I fit the effect of weight, and in the next slide I will then add in the effect of age. I should have clarified. Thank you.

DR. SHEINER: So the brown line that I'm looking at there is the equation that you wrote in the lower right-hand corner.

DR. WILLIAMS: Correct.

DR. SHEINER: And the way you know where to plot it on the age axis is what?

DR. WILLIAMS: By converting each age to a weight based upon standard CDC pediatric tables.

DR. SHEINER: Okay, but the fit was actually to the blue points where you knew what those weights were.

DR. WILLIAMS: I did not know what those weights were. I had to go by the age. So if we back up a little bit, these are the ages I had, but I have summary data. I don't have individual data. So what I did is for each bar I took the mean age. Then I went to the CDC tables to get the weight --

DR. SHEINER: Transformed it to a weight.

DR. WILLIAMS: Exactly.

DR. SHEINER: So if we go back to that picture, we're really looking at a transformation -- a fit of the

blue points on the y direction to a to a transformation, defined by these tables, of the data on the x axis.

DR. WILLIAMS: Indeed. It would have been more straightforward to plot weight on the x axis, but the reason why I didn't do that is twofold. One is for continuity with the next example where I'm going to fit weight and age.

DR. SHEINER: Where you're going to have both pieces of data.

DR. WILLIAMS: Exactly.

DR. SHEINER: Okay.

DR. WILLIAMS: Is that clear to everyone?

DR. JUSKO: From the previous graph, that should start at .5, at the age near 0.

DR. WILLIAMS: We have birth, which is --

DR. JUSKO: The ratio at birth on that graph is around .5.

DR. WILLIAMS: Correct. But the y axis here is different because these are weight-adjusted. The y axis here is child clearance divided by kilograms, quantity divided by adult clearance divided by kilograms.

DR. SHEINER: You're going to regret ever having shown that picture.

(Laughter.)

DR. WILLIAMS: So what I've done is I've taken

out the kilograms so I could fit weight.

DR. RELLING: What is your goal?

(Laughter.)

DR. RELLING: Why would you do that?

DR. WILLIAMS: The reason why I chose to do it this way is because you want to independently describe weight effects and age effects. You expect there to be weight effects, and you also perhaps expect there to be age effects. But you want to be able to independently address are there age effects that are not simply a consequence of weight.

DR. RELLING: Okay. Let's see what you have.

DR. WILLIAMS: I won't suffer from this difficulty when I have the FDA data because it doesn't initially present itself as normalized to kilograms. Is this making sense a little bit more now? Okay.

So in spite of the fact that the x axis is not saying so, I have fit the relationship between this ratio and weight here.

The next model I looked at is a combination of a weight effect and an age effect. The weight effect is what you saw on the previous slide. Here I've added in the age effect. The effect of the two summed together, each of which is a simple exponential relationship, is shown with the green line. The weight effect, which is what I

described previously, is shown with the line in the middle, sort of the pink dashed line, and finally, an age effect which is what's new on this slide.

Now, you'll notice that I fit 6 points with 4 parameters. My point here is not to show that I can draw pretty lines. The reason why I'm presenting this to you is because it shows the sort of strategy you might take when we have a larger database and how we might think about developing the models on our own data set.

Did this confuse everyone further? Can I aid anyone?

DR. SHEINER: Yes. You don't have any independent information in this particular data set.

DR. WILLIAMS: Correct.

DR. SHEINER: You have weight, which is a deterministic function of age, and then you have age, which is a deterministic function of age.

DR. WILLIAMS: Indeed.

VOICE: (Inaudible.)

DR. SHEINER: No. He got it from a table. So it's a deterministic function of age. So if I were to write your equation, it really is Rmax 1 minus E to the minus Kf of age plus Rmax 1 minus E to the minus K of age. So all you've done is done a shape thing. By restricting it to exponentials, you get more information out of two

exponentials than one. It's just like when we have time. And you said the right thing at the end which is this is just an illustration.

(Laughter.)

DR. WILLIAMS: I would agree. What we do have going here, though, is that the shapes -- I haven't looked at this specifically, but the shapes -- well, actually I have to an extent. The shapes are different. If you plot --

DR. SHEINER: No. I'm saying if I used a spline or some flexible function of age, I could only get one term in age because it would be as many parameters as I needed, but because you've broken it up into two exponentials, you can get two terms in age because they don't have the same shape because the function of age that weight represents is another shape change. So it's just like saying I have a polynomial in age. It's not a polynomial. It's a flexible function in age.

DR. WILLIAMS: I would agree.

DR. SHEINER: But it doesn't prove that you fractionated an age effect away from a weight effect.

DR. WILLIAMS: I would agree on that.

DR. SHEINER: Okay.

DR. WILLIAMS: Interestingly, I was somewhat surprised it turned out as consistent as it did because one

thing I did as a check -- like I said, it wasn't essential to my purpose because I'm just trying to show you the kind of strategy I would employ. But one thing I did do is I went back and switched the order in which I added the two, and interestingly I got the same relationship. I don't know if that's surprising, meaningful, or what, but it did happen.

DR. SADEE: Can I ask you about this? Using weight as a scaling may not be all that appropriate. So rather than saying there's an age effect, is there any information on body surface area which would just do away with this --

DR. WILLIAMS: No. The answer is no. In this database everything was normalized according to body weight, and other than the numbers as presented, which were always per kilogram, I had no raw data.

DR. SADEE: Well, could you translate this into body surface area which would provide you with a different scale and it may actually do away with the need to invoke age? Because body surface area changes with respect to weight and age, so it may account for both.

DR. WILLIAMS: Perhaps. When I actually do this on our own data set, the path that I intended to follow was, first, to describe the effect of weight or mass or BMI, ideal body weight, BSA. I would investigate a