FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

CLINICAL PHARMACOLOGY SUBCOMMITTEE

OF THE

ADVISORY COMMITTEE FOR PHARMACEUTICAL SCIENCE

8:34 a.m.

Tuesday, April 22, 2003

Conference Room
5630 Fishers Lane
Food and Drug Administration
Rockville, Maryland 20857

ASSOCIATED REPORTERS OF WASHINGTON
1523 North Carolina Avenue, N.E.
Washington, D.C. 20002
(202) 543-4809

ATTENDEES

SUBCOMMITTEE MEMBERS:

JURGEN VENITZ, M.D., PH.D., Acting Chair
Associate Professor
Department of Pharmaceutics
Virginia Commonwealth University
School of Pharmacy
Box 980533
410 North 12th Street
Richmond, Virginia 23298-0533

KATHLEEN REEDY, R.D.H., M.S., Executive Secretary
Advisors and Consultants Staff (HFD-21)
Center for Drug Evaluation and Research
Food and Drug Administration
5600 Fishers Lane
Rockville, Maryland 20857

EDMUND V. CAPPARELLI, PHARM.D.
Associate Clinical Professor
University of California, San Diego
9500 Gillman Drive
La Jolla, California 92093

HARTMUT DERENDORF, PH.D.
Professor, Department of Pharmaceutics
University of Florida College of Pharmacy
P.O. Box 100494, Health Science Center
Gainesville, Florida 32610-0494

DAVID FLOCKHART, M.D., PH.D.
Professor, Departments of Pharmacology
and Medicine
Indiana University School of Medicine
Division of Clinical Pharmacology
Wishard Memorial Hospital WP OPW 320
1001 West 10th Street
Indianapolis, Indiana 46202

WILLIAM J. JUSKO, PH.D.
Professor, Department of Pharmaceutics
State University of New York at Buffalo
School of Pharmacy
Buffalo, New York 14260

ATTENDEES (Continued)

SUBCOMMITTEE MEMBERS: (Continued)

GREGORY L. KEARNS, PHARM.D.
Professor and Division Chief
Pharmacology and Toxicology
Children's Mercy Hospital
2401 Gillham Road
Kansas City, Missouri 64143

MARY V. RELLING, PHARM.D.
Member, Pharmaceutical Sciences
St. Jude Children's Research Hospital
332 North Lauderdale, Room D-1052
Memphis, Tennessee 38105

WOLFGANG SADEE, DR.RER.NAT.
Chair, Department of Pharmacology
College of Medicine and Public Health
Ohio State University
5072 Graves Hall, 333 West 10th Avenue
Columbus, Ohio 43210

LEWIS B. SHEINER, M.D.
Professor, Laboratory Medicine
University of California, San Francisco
Box 0626, C-255
San Francisco, California 94143

MARC SWADENER, ED.D., Consumer Representative
2235 Dartmouth Avenue
Boulder, Colorado 80305-5207

GUEST SPEAKER:

MATS KARLSSON, PH.D.
Uppsala University
Sweden

ATTENDEES (Continued)

FOOD AND DRUG ADMINISTRATION STAFF:

PETER LEE, PH.D.
LARRY LESKO, PH.D.
NHI NGUYEN, PH.D.
HE SUN, PH.D.
GENE WILLIAMS, PH.D.
JENNY J. ZHENG, PH.D.

C O N T E N T S

P R O C E E D I N G S

(8:34 a.m.)

DR. VENITZ: I'd like to call the meeting to order, please.

Welcome, everybody. This is the Clinical Pharmacology Subcommittee meeting. We have a full agenda, as you can tell.

I'd like to open the meeting by introducing the individuals around the table, maybe starting with Dr. Derendorf, please.

DR. DERENDORF: Hartmut Derendorf, University of Florida.

DR. CAPPARELLI: Edmund Capparelli, University of California, San Diego.

DR. FLOCKHART: Dave Flockhart from Indiana University.

DR. SHEINER: Lewis Sheiner, University of California, San Francisco.

DR. SWADENER: Marc Swadener, Boulder, Colorado.

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

DR. VENITZ: Jurgen Venitz, Virginia Commonwealth University.

DR. JUSKO: William Jusko, University at

Buffalo.

DR. KEARNS: Greg Kearns, University of Missouri, Kansas City.

DR. RELLING: Mary Relling, St. Jude Children's Research Hospital in Memphis.

DR. SADEE: Wolfgang Sadee, Ohio State University.

DR. LESKO: Larry Lesko, Office of Clinical Pharmacology and Biopharmaceutics at FDA.

DR. LEE: Peter Lee, FDA.

DR. VENITZ: Thank you.

Our next order of business is the conflict of interest statement. Kathleen Reedy will read the conflict of interest statement.

MS. REEDY: Acknowledgement related to general matters waivers, Clinical Pharmacology Subcommittee of the Advisory Committee for Pharmaceutical Science, April 22, 2003, an open session.

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

The topics of this meeting are issues of broad applicability. Unlike issues before a committee in which a particular product is discussed, issues of broad

applicability involve many industrial sponsors and academic institutions.

All special government employees have been screened for their financial interests as they may apply to the general topics at hand. Because they have reported interests in pharmaceutical companies, the Food and Drug Administration has granted general matters waivers to the following SGEs which permits them to participate in these discussions: Dr. Edmund Capparelli, Dr. William Jusko, Dr. Gregory Kearns, Dr. Howard McLeod, Dr. Wolfgang Sadee, Dr. Lewis Sheiner.

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

In addition, Dr. Hartmut Derendorf, Dr. David Flockhart, Dr. Mary Relling, and Dr. Marc Swadener do not require special matters waivers because they do not have any personal or imputed financial interests in any pharmaceutical firms.

Because general topics impact so many institutions, it is not prudent to recite all potential conflicts of interest as they apply to each member and consultant.

FDA acknowledges that there may be potential conflicts of interest, but because of the general nature of

the discussion before the committee, these potential conflicts are mitigated.

With respect to FDA's invited guest speaker, Dr. Mats Karlsson reports that he has contracts and/or grants with AstraZeneca, Oasmia, Pfizer, and Servier. He also receives consulting fees from AstraZeneca, Ferring, Lilly, Pfizer, and Roche; and speaker fees from Johnson & Johnson and NovaNordisk.

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

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

DR. VENITZ: Thank you, Kathleen.

Before we proceed to the official business of today, I'd like to welcome a few new committee members. Dr. Swadener to my right is the consumer representative who also serves on the Advisory Committee for Pharmaceutical Science. We've got Dr. Shek who couldn't make it today who is the industry representative, also serving on the Advisory Committee for Pharmaceutical Science. And Drs.

D'Argenio and Davidian who couldn't make it today.

I would also like to thank the two outgoing members of the committee, Dr. Lalonde and Dr. Hale, as well as Dr. Jusko to my left, for chairing the previous committee meeting.

With that said, I'd like to turn over the meeting to Dr. Lesko who is going to introduce the topics for the next day and a half. Larry.

DR. LESKO: Thank you, Jurgen.

Well, good morning, everybody and again welcome back to Rockville. This is the second Clin Pharm Subcommittee meeting. We had our first back six months ago, on October 22-23, and as reflect back on that meeting, it was an extremely productive meeting for us. The advice we received at that time was excellent, and we've been thinking about it since then as we move forward with the initiatives we introduced at that first meeting.

I do want to say thanks again. As I look around the room, I recognize everyone here as very busy in their own work, and taking time to come to Washington and participate with us is extremely important. I think you'll find that the mission that we have for this committee is a noble one relating to some exciting times in the area of drug development generally and clinical pharmacology specifically.

Let me talk a little bit about today's meeting. Let me start with what's new since we last met in October, and there are really three exciting things that are new that really impact the topics that we'll talk about today.

The first is an FDA-wide announcement that our Commissioner made back in January called Improving Innovation in Medical Technology Beyond 2002. What this initiative entails is quite lengthy. There are several goals, but basically it revolves around improving the process, including the drug development process, for bringing medical innovations, treatments, and devices to the marketplace as quickly as possible that would benefit public health.

A second part of that, however, is improving the review process at FDA through a quality systems approach. This is the goal that we are going to sort of use to couch today's topics because a quality systems approach, if I think of it in terms of goals, has the goals I have on the slide basically.

Application of advances in science. Well, what are those advances in science? They're clinical trial design. They're clinical pharmacology study designs, statistical approaches, modeling and simulation, use of dose response in PK/PD in interesting ways.

Use of new technology. Use of new technology

embraces things like the quantitative methods we're going to talk about today. It embraces the integration of pharmacogenetics into drug development and review.

Rigorous analytic reviewer tools. This is the how to do it. What are the tools that we can make available to our reviewers to achieve this quality systems approach? We'll be talking about one of those in the first topic.

And finally, the overall goal of this initiative is to provide for high quality reviews. That's translated into effective reviews, efficient reviews, and consistent reviews.

Initiative number two that's occurred since the last time we met is an initiative under our Prescription Drug User Fee Act, PDUFA. It's the premarketing risk assessment initiative. We only recently had our first public meeting having to do with the risk assessment initiative, and Bob Meyer, our ODE II director, defined risk assessment as the process of identifying, estimating, and evaluating the nature and severity of risk of a drug product. You'll see some links between this and the first topic in today's presentations.

In that meeting on April 9th, Bob Temple described the ideal safety database that we ought to be striving for, and that included a complete characterization

of the clinical exposure-response relationship certainly for efficacy, but also for drug safety. And he made the point that this is important for making decisions about dosing adjustments, particularly in many of the clinical pharmacology studies when exposure goes up for a variety of reasons.

He recommended a good search for individualization factors. This obviously involves studying individual plasma drug levels, and he made the point that it's always necessary to assess polymorphic drug metabolism, and you'll see that this relates to one of our topics in this meeting.

Finally, he touched upon pediatrics, an important topic, and made the point that these pose special issues related to dose, PK, and PD, and we'll be touching upon that as well today.

The third initiative that's been launched since October is the one that relates to our FDA Science Board. We had this meeting last week on the same day as our risk assessment meeting. The theme of the Science Board meeting was integrating scientific advances into regulation with the emphasis on pharmacogenetics. Dr. Woodcock made a presentation at the Science Board and stated that genetic contributions to variability in toxicity include differences in drug metabolism, for example, thiopurine

methyltransferase, which was a topic of our last meeting, but more broadly recommended that we look at the use of genetic tests for metabolizer status to predict dosing.

So think of those three broad initiatives, and I hope it gives you a context for the discussion that we're going to have today and tomorrow.

We have four proposals, four topics on the agenda. The first is a proposal that we initiated our discussion in October and we've refined the proposal and we'll present with examples today the idea of a standardized approach to quantitate the impact that changes in exposure, related to efficacy or safety, result from changes in PK caused by a variety of intrinsic and extrinsic factors.

What we're trying to accomplish with this proposal and standardized approach relates to what I mentioned about the Commissioner's initiative, quality systems and review. We want to achieve a rational scientific basis for dosing adjustment that's quantitative and that links to the assessment of risk.

We have a goal of identifying individualization factors, and through our examples today you'll see some of those factors.

And finally at the end of the day, we hope to develop a standardized method that would rely on many

different tools, but a standardized thinking process that we hope would bring consistency to the label recommendations that we make in terms of dosing adjustments.

Our topic number two is pediatrics, and again going back to October, we opened the discussion of pediatrics very briefly the last time. Today we'll be making a proposal. The proposal will relate to a pediatric population PK design template. We'd like to recommend the use of this template for getting information about pediatrics during drug development. We feel that this approach is efficient in many cases. We feel it's under-utilized for a variety of reasons relating to perhaps a lack of understanding, perhaps related to a concern that FDA will not accept this type of study approach. But we'd like your input on the template that we'll be presenting.

Related to that, at the last meeting I had mentioned that we have this database at FDA related to pediatric studies. These are studies that were done under our pediatric rule. We felt this database is loaded with information that we could capitalize on by studying it, looking for trends, and learning something about the pediatric clinical pharmacology situation. We'll update you on our progress. It's been slow. It's been difficult because we don't have access to an electronic database, and

much of our time is simply gathering and assembling data. But nevertheless, we'd like to share with you some of the things that we've learned so far, but more importantly what we'd like to do going forward and look for input on designing the studies of this database.

The third topic, which we'll talk about tomorrow morning, is what I'll call a work in progress. We'll all familiar with the human genome. We've been bombarded with information about it, particularly this month on the 50th anniversary of identifying the double helix structure for DNA. Certainly the dream of genomics is to develop new and better treatments for disease states. But we feel there's a lot to be gained. There's a lot of substantial improvement that could be made by integrating pharmacogenetics into the treatments that we now have and using this science as it matures for identifying more optimal doses for subsets of the population.

We'll continue to talk about this tomorrow. What we're going to emphasize is moving forward with the knowledge that we have on polymorphic drug-metabolizing enzymes that influence variability in drug response, especially toxicity. It's a challenging area. Many questions come up in the context of this, things like how much variability will genetics explain, how much of an effect on drug dose will genetics explain, how important is

it.

But I think more importantly where we're heading is to create a general construct for looking at improvement in existing therapies, existing therapies being approved drugs, to determine what criteria we ought to be thinking about that would warrant updating labels for products to optimize drug dosing using genetic information.

Last time we talked about specifically thiopurine, TPMT. Tomorrow we'll touch upon that, but we'll also be looking for a broad way to best program in this area and what we need to be thinking about in assessing data, assessing evidence to update labels. So a rational scientific basis.

Now, our fourth topic today is going to be a new topic. We'll actually talk about it tomorrow. We've been working pretty much over the last year in the area of drug-drug interactions. It continues to be a major problem if you read the current literature in JAMA and the New England Journal of Medicine about adverse drug reactions and the high fraction of those that are related to drug interactions. We have some ideas on revising our guidance on drug-drug interactions. We have some questions on transporter based drug-drug interactions. As was stated in our risk assessment workshop, some matters always need assessment in regulatory review, including new interactions

that we may not have paid as much attention to in the past, such as interaction involving glucuronidation and transporter interactions like P-gp.

So we're going to bring this topic forward tomorrow with some issues and questions. We'll talk about the use and extension of a classification system for 3A4 inhibition for single and multiple drug interactions. This is a classification system that we can see moving towards label language for bringing some consistency to how we report drug interactions in the label.

And a big question that we frequently get from sponsors during the drug development process and we ask ourselves is when and how should the role of P-glycoprotein in drug interactions be investigated. This is an emerging area and we're beginning to see clinical evidence that this important and we'd like to develop a path forward that's reasonable and rational.

Each of the presenters is going to have some specific questions on the topic, but I'd like you to think about some of the broader questions that we have for the session. For example, you'll hear many proposals during the next day and a half. Aside from the specific questions about the subtopics of today's meeting, think about the rationality of these proposals. Are they reasonable? Are they feasible? Overall, do you think these will enhance

the quality of drug development and regulatory review? Are these the priorities that we should be looking at in our clinical pharmacology program?

We have works in progress, topics number 3 and number 4. That means we need input as we move forward with these programs and some advice on whether these objectives are worthwhile. And in particular, what is the best way to integrate new science and technology, whether it's genetics, whether it's P-gp transporter information, into therapeutics and regulatory review?

Well, that's the introduction to the lineup for today. I look forward to the discussion. It was a terrific discussion in October. Again, as we look around the table, the expertise of this committee is really substantial, and we're looking forward to defining and expanding upon the proposals that we're going to make during this meeting. Thanks.

DR. VENITZ: Thank you, Larry.

Now we're moving to our first topic. As Larry indicated, we're going to talk about exposure response as a way of justifying dose adjustments. The introduction will be given by Peter Lee. He's Associate Director of the Office of Clinical Pharmacology and Biopharmaceutics.

DR. LEE: Good morning. The first topic we're going to discuss today is quantitative risk analysis using

exposure response for determining dose adjustment for special populations.

This topic has been discussed in our previous meeting back in October 2002. In the last meeting we talked about three main topics. We had proposed a standardized approach to estimate the probability of adverse events in special populations using exposure-response information. We also discussed a regulatory decision tree for recommending dose adjustment in special populations. And lastly, we also discussed the potential application of utility functions for risk and benefit assessment.

So today we're going to present several examples to illustrate what we talked about in the last meeting. We're going to present examples of a standardized approach for using exposure-response information to adjust dose in special populations. We also are going to present one example of using population analysis to obtain PK/PD information from the large clinical trials. And in the last example, we're going to present a methodology to applying utility function for an optimal dosing strategy.

Just to give a little bit of background information, as you know, many of the NDAs may contain up to 20 or more clinical pharmacology studies, and in these studies different intrinsic and extrinsic factors may be

studied and these factors may influence the pharmacokinetics of the drug in these special populations. Therefore, we need a consistent approach to determine the dosing adjustment requirement in these populations.

Here's one example. In this particular example, we have about 11 factors that have been studied in the NDA. As you can see, the area under the curve of the drug may change depending on what factors from 0 percent, which is no change, to a 60 percent increase in the special populations.

So the question is, how do we make the dose adjustment according to the pharmacokinetic results? Where is the cutoff? Do we adjust the dose at 30 percent increase of AUC, or do we adjust the dose at 60 percent increase of AUC?

So the answer is that we had to look at PK/PD information and determine what is the clinical significance of this AUC change.

So there are some issues related to the dosing adjustment in drug labels of NDA submissions. Quite often we have seen inconsistency in dosing adjustment recommendations in the initial label of NDA submissions. Exposure-response information, as is required to interpret the pharmacokinetic change, is now always available in the NDA submission. The FDA reviewer had to conduct additional

exposure-response analyses in order to interpret the AUC change. Therefore, we feel that a standard for analyzing and interpreting the exposure-response information will be critical and beneficial to regulatory decision making in terms of dose adjustment in special populations.

So to improve the current status, in the last meeting we had proposed to develop and evaluate a standardized approach for the reviewer and possibly for industry to quantitatively assess the impact of exposure change on either safety or efficacy that results from a change in pharmacokinetics due to intrinsic and extrinsic factors.

This is the standardized approach that we had proposed in the last meeting. In this example, basically we have seen an increase of exposure of the test population compared to the reference. Using exposure-response information, we can estimate the distribution of response in both reference and test populations. If we could determine what is the critical value of response, which is considered clinical significance, which is the vertical line here, then we can calculate the probability of a clinically significant response based on the PK change, as well as the PK/PD relationship.

So in order to interpret the clinical significance of pharmacokinetic change, we need to have

observed data in pharmacokinetics in special populations. We also need data of the PK/PD relationship. With this information, then we can estimate the probability of adverse events in the special populations with the response that is greater than the clinically significant critical value.

However, in order to determine the clinically significant critical values, we need to base it on a risk and benefit assessment of the drug therapy. Currently we're doing that on a case-by-case basis through a discussion between clinical pharmacology and the medical reviewer. But in the last meeting with the committee, we also proposed that we can use a utility function to assess the risk and benefit of pharmacokinetic change in the special populations.

In the last meeting, we also discussed a decision tree for dosing adjustment recommendations. Since the last meeting, based on the recommendation from this committee, we have made some modifications of the decision tree, and this is the current decision tree. Basically we ask a number of questions.

First, according to our current guidance, we ask whether the 90 percent confidence interval of test over reference is within the default no-effect boundary. A no-effect boundary could be, for example, 80 to 125. Now, if

the answer is yes -- we think it is the no-effect boundary -- then there's no dose adjustment required for the special populations. But if the answer is no, which means the 90 confidence interval is outside the boundary, then we ask the next question, whether we have a PK/PD relationship.

If we do have a PK/PD relationship, then we will take the standardized approach to estimate the probability of adverse events and probability of effectiveness in the special populations and ask the question whether that's a clinically significant change from the typical population. If it is considered clinically significant, then we will recommend a dose adjustment or precaution or warning in the drug label.

As I mentioned, there will be several examples discussed in today's meeting. The first two examples will be used to illustrate the use of the proposed standardized approach for estimating the probability of toxicity in special populations using exposure-response information.

And the next example will be used to illustrate the potential utility of population analyses to obtain exposure-response information from large clinical trials. We think this is a very important topic because the large clinical trials represent a unique opportunity to obtain exposure-response information from the studies.

The last example will be used to demonstrate a

method of applying utility functions to optimize a dosing strategy.

Today we have four speakers to present the examples. The first speaker is Nhi Nguyen from DPE I, and she will present the first example. The second speaker is Dr. Jenny Zheng from DPE III. She'll present a second example of the standardized approach. And the third presenter will be Dr. He Sun from DPE II, and he will present the population PK/PD approach. The last speaker, our guest speaker, is Dr. Mats Karlsson from Uppsala University, and he will present an example illustrating the utility functions.

After each presentation, we're going to ask one or two questions to the committee. I'd like to present the questions now so that hopefully the committee can keep those questions in mind when listening to the presentations.

The first two questions relate to the standardized approach. We would like to ask, under what treatment circumstances, for example, intrinsic or extrinsic factors or therapeutic areas, would this standardized approach not be applicable? We also ask a second question. Does the exposure response differ between special populations and typical populations? If so, how can the differences be detected?

The next questions will be related to the population PK/PD analysis, and there are two questions related to that topic. The first question is, what are the utility and general limitations of linking pharmacokinetics obtained from the population analysis to the response endpoints? And the second question is what are the general considerations in using exposure response for dose adjustment in special populations, especially using the population approach to obtain the exposure response?

The last question is related to the utility functions, and the question will be, can the presented approach of utility function be generalized to other scenarios?

So with that, I want to introduce our first speaker of the examples, Dr. Nhi Nguyen. She will present the first example of the standardized approach.

DR. NGUYEN: Good morning. This morning you will hear a presentation on a method of analysis and how it was applied in regulatory decision making.

This slide will summarize how we did the analysis for this NDA.

The first step is to develop your exposure-response models. The exposure-response models should be based on large, randomized clinical trials, trials that explore a wide exposure range and include a large number of

people and are of the longest duration.

The second step is to have an expectation for your target window of exposure. How much benefit does one need and how much risk is one willing to assume?

The third step is to example what happens to exposure response when various intrinsic and extrinsic factors are introduced. Typically studies in special populations include pharmacokinetic information and are underpowered to provide good response data. So with the appropriate assumptions about exposure response in these special populations, we took the data from the special population studies, the individual data, not just the mean data, and integrated it into the exposure-response models and then determined the probability of effectiveness and safety.

So probability is on the y axis here and the sum of these bars equals 100 percent. So you can see that we not only have a feel for the maximum likelihood of benefit or risk, but we also have a feel for the tails of the distribution.

This slide summarizes how we did the analysis for this review.

I will take one more slide to further explain the last step, determining the probabilities. And for this example, I've chosen the QTc interval which is a surrogate

for torsade, a fatal ventricular arrhythmia. A clinical trial that examines QTc prolongation may have a distribution of baseline QTc intervals that look like this. Modeling the data may result in concentration QTc slope distributions that look like this. So if we want to see what happens when an intrinsic or extrinsic factor is introduced, we took the results from the PK study, and for this example I'm illustrating Cmax and overlaid it into the known concentration-QTc relationship.

So, in essence, we sampled from each of these distributions and created a virtual patient, and we did this 1,000 times to determine empirically what happens with a concentration and achieving a specific QTc. So by doing these simulations, we were able to test combinations of the tails of the distributions that were untested.

So that leads me to our objective which was only to quantitate the risk-benefit of a drug. A decision about what to do about the risk-benefit should be made with the whole review team or the domain experts.

In developing the exposure-effectiveness model, for the primary endpoint, we chose the largest clinical trial which was also of the longest duration. The primary effectiveness endpoint and pharmacokinetics were collected at baseline, week 2, and week 4. This study also explored the largest exposure, and these doses have been changed for

the purpose of this presentation, but let's just say that 1 milligram a day is the sponsor's recommended starting dose, with titration to 2 milligrams a day. So this study explored a dose greater than and a dose less than the sponsor's recommended dose.

Next we developed the exposure-safety or exposure-risk models. For these models, we used the adverse event data from all the pivotal clinical trials. Now, you can imagine that large clinical trials may have hundreds of adverse events. So after discussions with other members of the review team, we prioritized these adverse events and focused on these six: dizziness, edema, liver toxicity, palpitations, tachycardia, and vertigo.

We also analyzed QTc prolongation because of drug properties suggestive of QTc prolongation. For this analysis, we chose the study that had the most information on the time course of QTc prolongation. You will note that this was a drug-drug interaction study, and the sponsor used half the recommended starting dose. ECGs were only measured up to 4 hours post dose, so there were some study design limitations. And the study contained 24 hours of drug concentration data.

So now that you've seen the exposure effectiveness and the exposure risks that were assessed, let's take a look at the models.

This is the exposure-primary effectiveness model, and the asterisk in the following slides will indicate the mean Cmax of the sponsor's recommended starting dose of 1 milligram. These lines indicate the mean Cmax for the 1, 2, and 4 milligram dose, and you will note that the increase in concentration is more than dose proportional. The maximum effect was 7.6 and the concentration that produced half the maximal effect was about .2 units.

In the following slides, I show a mean Cmax line to keep the slide clean, but really we are considering the entire distribution of individual Cmax's. So it's something that may look like this with some overlap between the 1 and 2 milligram dose. So that's the exposure-effectiveness model.

When you look at the risks, each blue line on this slide indicates one individual's concentration/QTc prolongation relationship. The QTc corrections shown here are individual corrections obtained by nonlinear mixed effects modeling.

There was a statistically significant relationship between concentration and QTc prolongation. However, you can see that there is a lot of variability. The starting dose in some patients results in no QTc prolongation. However, in other patients, it results in

substantial QTc prolongation. You will also note that we have very little data around the concentration of the mean Cmax for the 2 milligram dose.

For our analysis of other adverse events, we found three adverse events to be statistically significant and that was tachycardia, palpitations, and edema. However, since the analysis of all these adverse events was similar, I will only present one for the sake of time.

This slide shows the probability of tachycardia by the effective dose, and the effective dose is an adjustment of the actual dose to account for the saturable first pass of the drug. So, 1, 2, and 6 really correspond to 1, 2, and 4 milligrams of drug.

The probability of tachycardia was dependent on weight and dose. So in a 70-kilogram patient, you can see that there is a very small probability of tachycardia, and this probability does not increase with a six-fold effective dose increase. Whereas, in a 50-kilogram patient, there is about a 5 percent probability of tachycardia, and then this increases about 1 percent with a six-fold effective dose increase.

So now you've seen the exposure-effectiveness model, the exposure and QTc model, and then the probability of tachycardia by effective dose. So now we're equipped with the models necessary to interpret results of the

special population studies.

This slide shows results of two special population studies presented in terms of changes in AUC and Cmax. Ketoconazole with a half a milligram of drug resulted in a 13-fold increase in AUC and a 7-fold increase in Cmax, and grapefruit juice with 1 milligram of drug resulted in a 7-fold increase in AUC and a 6-fold increase in Cmax.

So the next step is to see how these data integrate into the known exposure-response relationship. This is the same figure you saw earlier, only it's smaller, of the concentration effectiveness. When a half milligram of drug is given, you would expect to see an effectiveness around 3. Taking ketoconazole with a half milligram of drug increases concentrations about 7-fold, reaching the Emax of about 7.6. Taking 1 milligram of drug with grapefruit juice results in about a 6-fold increase in concentration, and no additional effectiveness. Now, hypothetically if ketoconazole were given with 1 milligram of drug, we might expect to see a similar increase in concentration.

If we look at the concentration/QTc relationship, a half milligram of drug would result in about this amount of QTc prolongation. Taking ketoconazole with a half milligram of drug pushes patients from this

amount of QTc prolongation over to this amount. Taking 1 milligram of drug with grapefruit juice results in about a 6-fold increase in concentration, and the concentrations are then off the figure and we do not have QTc data there. And then hypothetically again, if ketoconazole were given with 1 milligram of drug, we might expect to see a similar response. So in this situation, we would not be able to make any conclusions about what happens at these higher concentrations on QTc prolongation because we do not have data there.

Now, I also want to remind you again that there is a distribution of Cmax's and slopes. So really we are looking at data that looks like this. So if we want to consider the worst case scenario, we have to consider both of these distributions, and the results of some of those simulations will be presented in the next slide.

Now, if we look at the probability of tachycardia, taking a half milligram of drug with ketoconazole in a 50-kilogram patient barely increases the probability of tachycardia, whereas taking grapefruit juice with 1 milligram of drug increases the probability of tachycardia about 1 percent in the 50-kilogram person. In a 70-kilogram person, you can see that this slope is pretty much a straight line and there is not much of an effect on the probability of tachycardia. Then again if ketoconazole

were given with 1 milligram of drug, we might expect to see a similar response as that with grapefruit juice.

So to summarize the data integration, ketoconazole with a half milligram of drug results in a 13-fold increase in AUC and a 7-fold increase in Cmax. And this translated into a 4-unit effect, and we did simulations to determine that 5 percent of the population may have a prolonged QTc greater than 32 milliseconds. Realize that these simulations are determined from the data in the ketoconazole study. So we could present this data in other terms, such as change from baseline or percent or the time above a certain threshold QTc.

And then the probability of tachycardia with a half milligram of drug and ketoconazole was barely increased or affected. Grapefruit juice with 1 milligram of drug resulted in a 7-fold increase in AUC and a 6-fold increase in Cmax, and this translated into no additional effectiveness, and we were unable to conclude anything about the effect on QTc because we did not have data at those higher concentrations. And then the probability of tachycardia increased about 1 percent in the 50-kilogram patient, and there was negligible effect in the 70-kilogram patient. And then again if ketoconazole were given with 1 milligram of drug, we may expect to see a similar response as that seen with grapefruit juice.

So at this point we would present the table to the review team and weigh the effectiveness and risks and realize the assumptions of our models. One is that it is the higher drug concentrations, not the intrinsic or extrinsic factor itself, that alters response. We recommended to conduct an appropriate QT study, one that explores a wider concentration range and one that collects QT data over 24 hours.

DR. VENITZ: Thank you, Nhi.

We have time for questions. Go ahead.

DR. SHEINER: I gather that the various parts of the various models were gathered from different data sets sometimes.

DR. NGUYEN: For the effectiveness, we used the largest clinical trial, and then for the safety and risks, we used the largest clinical trial and the other pivotal trials.

DR. SHEINER: I guess the question is this. You've got several models that are translating from A to B and then B to C and so on.

DR. NGUYEN: That's correct.

DR. SHEINER: And the question is, were enough of them gotten from the same set of people so that you could look at things like correlations? Does it turn out, for example -- not that it should -- that the people with

the high concentrations essentially -- in other words, your model is kind of assuming that this association that you see, there are no correlations. So it's not necessarily true that somebody who has, let's say, a raise in level and doesn't have toxicity will also not have efficacy or something like that.

You've got a set of relationships that translates from concentrations before you add ketoconazole, let's say, to afterwards, and then you map from concentration to effect, but there is no part of this thing that says, well, when you raise the concentration to the ketoconazole, maybe the relationship of concentration to effect is not the same. I'm not saying that it is, but you don't have any evidence to say one way or another. Is that right?

DR. NGUYEN: Yes.

DR. VENITZ: Any other questions?

DR. FLOCKHART: One thing directly to that point. There is some data -- but this might be addressable -- to suggest that ketoconazole itself can affect cardiac repolarization.

DR. NGUYEN: That's right.

DR. FLOCKHART: You might have data on that from the control arms of the smaller trials. If they show no effect, that would be reassuring. It doesn't completely

get to Lew's point because it's still possible that the concentration-effect relationship is different in the presence of ketoconazole.

DR. CAPPARELLI: Just a clarification so that I understand the terminology. When you say probability of tachycardia, you're speaking of sinus tachycardia in this case, not torsade de pointes?

DR. NGUYEN: That's correct, sinus tachycardia.

DR. CAPPARELLI: I just wanted to be clear because I think one question that I had is, did you take a similar approach to heart rate that you did to QTc? Because your sinus tachycardia is going to be relative to where you start from, at least the risk.

The one thing that I think was brought up as a question is, are there extrinsic factors that we need to think about in these models? Clearly, strictly from a PK standpoint, I wouldn't expect a 50-kilogram person to have a different response based on weight. So I think there clearly is an extrinsic factor that's linked to the 50-kilo patient rather than the 70-kilo patient. But I'm not certain that it's gotten here. So I have some questions about the classification scheme based strictly on weight, and maybe linking to where their baseline heart rates would be of help from that standpoint.

DR. NGUYEN: Actually for the tachycardia

analysis, we would have preferred to analyze it by heart rate, but the data were collected like that. So it was sinus tachycardia, yes or no. So we did a logistic regression.

DR. CAPPARELLI: Was there an age effect or other disease effects that you looked at in terms of heart failure? It's kind of difficult to look at this and see where you expect a large change in concentration such that the 70-kilo person at the highest dose is going to have much higher concentrations than the 50-kilo person at the lowest dose. And yet, you're seeing this differential PD response. Without understanding what's causing that, I think it becomes very difficult to extrapolate from this aspect of the analysis.

DR. NGUYEN: Probably the 50-kilogram person did receive more of a dose, milligram per kilo, than the 70-kilogram person. But the analysis -- they were given the same dose. So we didn't have concentration data to analyze data by concentration and probability of tachycardia. We only had dose data.

DR. SHEINER: Just a comment. I think we're getting at the fundamental problem that what you want to do is extrapolate to new circumstances, people having these other co-factors. You want to get some reasonable guess as to what's going to happen, what's dangerous and what isn't.

Yet, you're extrapolating from observational data based models, which is the hardest thing to do because you don't know where causality resides in those models.

One of the solutions in the past is to just not do it, and I don't think that's the right solution. But I do think there is no easy solution, and we have to be quite careful about things and recognize that we're talking about outer boundaries and recognize that we're talking about sort of worst case scenarios or maybe even best case scenarios. We can't be sure. We have to somehow get comfortable with the increased degree of uncertainty that arises in this activity. But as I say, I think we should do it because the alternative is even greater uncertainty.

DR. DERENDORF: There are a lot of straight lines in your concentration-effect and dose-effect relationships. Is there enough evidence that you have linear relationships between those parameters, particularly when you use them to extrapolate?

DR. NGUYEN: Which one are you referring to?

DR. DERENDORF: The concentration-QTc prolongation plot and then also the one below the dose-tachycardia plot. You just have straight lines there that suggest that concentration and effect are linked that way. Do you know that?

DR. NGUYEN: No. I mean, that was the data

that we had. So like for the QTc, they measured ECGs at 0, 1, 2, and 4 hours post dose, and they had 24 hours of concentration data. So that straight line is the relationship between the concentration and QTc.

DR. DERENDORF: You think it is or you know it is?

DR. NGUYEN: Well, that's what it was in that population. They could have gone with a higher dose range, and then we could have seen what type of model the relationship is. So I don't know.

DR. DERENDORF: Then the other question that's related to the previous question that I really am puzzled with is this 50- versus 70-kilo situation where you have a 6-fold dose and a 70-kilogram person doesn't have probability versus a one-sixth of a dose in a 50-kilogram person. So there would be quite different exposures and very different risks.

DR. FLOCKHART: It just means to me that the relationship between dose and weight isn't a simple linear relationship. You're right. There may be something else involved, but that's not uncommon.

DR. SADEE: I have a comment also on the variability from one patient to the next. You're looking at two interactions. It's maybe one of the classic examples in pharmacogenetics where you have a variety of

different genetic variance from one patient to the next, and actually that could affect the interaction between the two drugs so that you may have specific cases in the single patients that are totally different in their exposure than others. So there would be no way of extrapolating from that because you're disturbing the very relationship with the dose response that you're looking at.

DR. FLOCKHART: My difficulty here is the FDA is faced with the problem of trying to make a rational prediction. We can sit as academics and ding them all over the place for it. You know, you can't do this and you can't do that. But the reality is you have to try. And I think Lew's point is salient. We have to try and include the error.

DR. KEARNS: And that's the point that I think I want to make. Back to your ECG slide. It's not to be critical. It's quite exciting to see 20 percent of people have a response that way and then one outlier at the top who really had one.

But I was struck by the recommendation that you showed on your slide which was, okay, we did this. Now we go back and recommend a trial with more concentrations, wider range. And as Dr. Sheiner was kind of getting at, there's a lot that's riding on this extrapolation. I think from a public side, there's always the question of how much

additional time is that going to take. From a medical side, there's always the question about that one or two outliers. Are those the people who die in that trial because they have some hERG channel defect that's not recognized at the outset?

So I think trying to go back and do the diligent thing is to wire up the model as best you can. For instance, if that was a pediatric population and you looked at baseline QTc's, you'd see quite a different dispersion based on age for no treatment than you would in an older free-living population.

So is the applicability of the model approach -- can we go across populations? It depends I think. But to jump to the study and to add the patients, to add the concentrations, to maybe add the risk until you've taken all the flies that you can out of the ointment could be premature.

DR. VENITZ: Thank you, Nhi.

Let's move to the next presentation. Dr. Jenny Zheng. She is a pharmacometrics reviewer in the Division of Pharmaceutical Evaluation III. She's going to give us a second example.

DR. ZHENG: Good morning. Today I'm going to present another example to illustrate how the dose-response relationship was used for recommending dose adjustment.

In our review process, it's very often to see the pharmacokinetics of a drug is influenced by intrinsic factors such as age, gender, impaired renal and hepatic function and extrinsic factors such as drug-drug interactions. In this situation, we have to ask the question what is the clinical significant of the changes in concentrations.

Currently the decision will be made based on the clinical assessment based on the clinical experience and totality of the evidence. But the disadvantage of that approach is it's not a quantitative and standardized approach. The assessment could be pretty subjective. In other words, the decision may not be the same based on who makes the assessment and from what perspective. Therefore, we propose from a clinical pharmacology perspective to use the exposure-response relationship to bridge the response and exposure data to quantitate the influence after changes in the exposure.

The example I'm going to present will focus on the drug concentration increase and the safety assessment. This is drug Z. It's a noncardiac drug. From both preclinical and phase I studies, it shows that the drug caused QT prolongation and this QT prolongation is concentration dependent.

The phase I PK studies showed three factors

increased the drug concentration. In an age study, it shows that the mean steady state maximum concentration was 100 percent higher in elderly subjects as compared with Cmax in young subjects. The renal study demonstrates steady state Cmax in severely renally impaired subjects was 50 percent higher as compared with healthy subjects. And drug interaction studies showed ketoconazole increased the steady state Cmax by 60 percent.

Knowing the concentration increase in this situation, the question raised is, should dose be adjusted in elderly, renally impaired subjects or when co-administered with ketoconazole? To answer that question, we need to understand the effect of increase in drug Z concentration on the QT prolongation which will rely on the exposure-response relationship.

The exposure-response relationship was obtained from several phase I studies. They were all placebo-controlled crossover studies. The doses included in the study were a clinical dose and two times of the clinical dose and three times of the clinical dose. The higher dose is important here to provide the wide range of the concentration which is the key for obtaining exposure-response relationship. From all these phase I studies, blood samples were collected for drug measurement. Also the QTs were measured.

The results of the analysis are shown in this slide. The QT prolongation is represented as delta QTc, which is the QT change from the baseline. So the association between the delta QTc with the concentration was described by a simple linear regression model. The linear mixed effect model was used for analyzing these data. The dashed lines represent individual regression lines. The solid line represent the population regression line. The wider band of the lines indicates that the inter-subject variability is quite high. The estimated slope ranged from 1.5 to 7.6, indicating that for some of the subjects, the delta QTc change is sensitive to the concentration change. In some of the subjects, the change is not quite as sensitive.

An outlier analysis is a very important part of QT assessment. We want to know how many subjects would experience the delta QTc longer, for example, 10 milliseconds, 20 milliseconds, 30 milliseconds, or 40 milliseconds. Unfortunately, the phase I study usually included a limited number of subjects which limits its ability for that type of analysis.

For the phase III study, even though hundreds of subjects may be included in that analysis, the QT measurement is not as intensive as the phase I study. So it's difficult sometimes to capture the outlier from the

phase III study. On the other hand, if you're interested in the special population, even a phase III study may not provide sufficient number, for example, severely renally impaired subjects.

So in order to make an outlier comparison between the population, a simulation exercise was conducted here. Most specifically, the phase I data, the concentration data was modeled assuming the logarithmic distribution in the PK parameter. Using that model, 2000 maximum concentration was simulated for young subjects, elderly subjects, for renally impaired subjects. The same approach is used to simulate 2000 Cmax for when ketoconazole is co-administered with the drug Z. So we have 2000 concentration in each special population, special situation. Then we used the exposure-response relationship as described in the previous slide to predict delta QTc.

The results of the age effect are presented in this slide. The data is presented as the percent of the subjects who would have the delta QTc longer than 10 milliseconds, 20 milliseconds, 30 milliseconds, 40 milliseconds. These results indicated that about 2 percent of young subjects would have delta QTc longer than 40 milliseconds and for the elderly, the percentage will increase to 7.3 percent. So for delta QTc longer than 30 milliseconds, in young subjects it's about 8 percent; in

the elderly, it's about 19 percent. So a similar trend could be seen for a delta QTc longer than 20 milliseconds and 10 milliseconds.

This slide presents the results for renal function. As you can see, most subjects with severe renal impairment would have longer QT prolongation than the normal subjects. For example, for the normal renal function subjects, 2 percent would have delta QTc longer than 40 milliseconds, but if you have severely renally impaired function due to the concentration increase, there will be 5 percent of the subjects who would have a delta QTc longer than 40 milliseconds.

The results in this slide show ketoconazole increased the percent of subjects who experienced a certain extent of delta QTc. Like if you're taking the drug alone, 2 percent of subjects would experience delta QTc longer than 40 milliseconds. But if you take drug Z with the ketoconazole, the percentage will increase to 6.2 percent.

The percent of subjects with delta QTc longer than 40 milliseconds is summarized in this slide. You can see that the risk of having a delta QTc longer than 40 milliseconds is higher in elderly subjects, in severely renally impaired subjects, and when the drug is co-administered with ketoconazole.

Examination of the creatinine clearance

indicated that the age effect might be partially attributed by reduced renal function. So in the age study, creatinine clearance was 50 percent lower in elderly subjects as compared with the young subjects. So it's believed that age effect would be reduced if the renal function effect was corrected by dose reduction.

Since the consequence of the worst event could be very severe, the question was asked in the review team, what would be the effect of ketoconazole in subjects with severe renal impairment? Not many subjects would belong to this group, even from a phase III study. So in order to make that assessment, a simulation was conducted.

First, steady state Cmax in severely renally impaired subjects was simulated as I described earlier. In the second step, the Cmax ratio of drug Z at presence and absence of ketoconazole was obtained from the crossover study so that the ratio actually characterized the ketoconazole effect. From that study the ratio ranged from 1 to 4. So the combined effect for both factors was simulated by just randomly multiplying the maximum concentration from step 1, which is the maximum concentration for severely renally impaired subjects, and the ratio from step 2 which characterized the ketoconazole effect.

The results are shown in this slide. As you

can see, 19 percent of subjects who are severely renally impaired would experience a delta QTc longer than 40 milliseconds when co-administered with ketoconazole.

This slide just simply summarizes all of the factors, the effects. It summarizes young subjects. It's the percent of subjects with a delta QTc longer than 40 milliseconds. For young subjects, it's about 2 percent. In the elderly, it's almost triple the percentage, up to 7.3 percent, and more than double that percentage in severely renally impaired subjects. And when drug Z is co-administered with ketoconazole, the effect could be very dramatic if the two factors are combined.

That analysis leads to the conclusion that the increase of concentration by age, severe renal function, and co-administration with ketoconazole resulted in increased number of subjects with a delta QTc longer than 40 milliseconds. The effect is more significant when two factors are combined.

Based on that analysis and the consideration of the nature of an adverse event, a dose reduction was recommended in severely renally impaired subjects. A dose reduction was also recommended when drug Z is co-administered with ketoconazole.

DR. VENITZ: Thank you, Jenny. We have about 5 minutes for questions.

DR. DERENDORF: I think it's the same issue as in the last case. You're assuming that the exposure-response relationship that you got from your phase I study is a constant and it doesn't change in elderly or in severe renal impairment. So the calculations that you're making are all focused on exposure, and then at the very end, you convert that into expected --

DR. ZHENG: I don't quite understand your point. Actually the delta QTc versus concentration, that is the relationship between the effect versus concentration. I don't think we make any assumption with a constant concentration.

DR. DERENDORF: No, not constant concentration. But the relationship between the exposure that you have in your different cases and the outcome -- you take that linear relationship that you have from your phase I study where you have concentration versus change of QTc and apply that to all of these cases assuming that this relationship holds true for all of these.

DR. ZHENG: Okay. So you have the problem with the extrapolation from the young healthy subjects to the elderly population.

DR. DERENDORF: I don't see any evidence that it holds.

DR. ZHENG: In one of the phase I studies, they

included not only the young subjects but the elderly. So we do look at the relationship there. We don't see much difference in terms of the slope, the relationship. So that's one of the information we could have.

In terms of the drug-drug interaction and the severely renally impaired subjects, yes, we don't have data to show that they are going to have the same relationship. But you are right. That's the assumption we have to make for these type analyses.

DR. CAPPARELLI: Just as a follow-up on that, because this is a recurrent issue -- I mean, this particular QT prolongation looking at drug concentration effects -- has there been a systematic look at several of these drugs looking at especially, say, with renal impairment where you're going to have changes in electrolyte abnormalities and looking really at sort of a population dynamic model to identify the covariates?

So I think as Hartmut was saying, we're going forward with the assumption that the exposure-response relationship is totally uncorrelated with the changes in -- the disease states that are causing the changes in PK. I think this is actually a great example where maybe in some across-study evaluations, one could actually look at some of these populations not only at the variability in response in subpopulations, but maybe the electrolyte

differences in your renal failure patients may change that slope entirely. So adding these effects as we go along in the chain, it's nice along the way to test some of these assumptions.

DR. ZHENG: I think if we could have enough information, definitely that's a good thing to do. But I think here, unfortunately we just don't have that much information. So the focus here is simply the effect of concentration on the delta QT.

DR. SHEINER: Of course, the answer is get more information. But the answer to the problem when you don't have enough information is not that you have to make some assumption and go with it, but that you have very carefully display. It is sort of what I was indicating before. You have to carefully display the limits of your knowledge.

So if I look at, for example, the last page and the last several slides you showed, you've got these bars that are just heights, the amount of change with the elderly or renal failure, and there are no uncertainty intervals on them. Yet, this is exactly what you need to pay a lot of attention to, it seems to me, in this kind of a situation so that you can have a rational dialogue with other people.

And where do the uncertainties come from? And we have techniques whereby you say, well, look, I have this

linear relationship between QTc change and the concentration, but I could put in some uncertainties about, let's say, whether it applies to other populations. Then I can actually build that into my projections, and I can see that instead of having 30 percent of people above QTc of 40, it will be anywhere from 10 to 50 percent, or whatever the numbers are.

That's the point. You've got the computers to do it. It doesn't cost money. And that's the way I think to deal with the problem that there are so many assumptions that have to be made, sensitivity analyses and honest uncertainty intervals which involve model uncertainty as well as data uncertainty. And then everybody is talking about the same thing. It may well be that the conclusion stays the same.

DR. LEE: Dr. Sheiner?

DR. SHEINER: Yes.

DR. LEE: To follow up, if we don't know the true relationship in different populations, how do we build into the model the uncertainty due to population difference?

DR. SHEINER: Well, let's say we'll talk about the average slope. Let's say you're willing to assume it's linear, but it's the average slope that's different in different populations. So then you just talk to a bunch of

people and you say, how big do you think it could be, and you just build that uncertainty in. Now it spreads out all of your predictions, and it means there's a larger fraction of people who have low values, but there's a larger fraction of people who have high values.

So it's a matter of assessing the risk. It's what's the probability based on everything we know, including all the uncertainty, that the value will be greater than this. And that will be your most educated guess.

The point is it will be everybody's most educated guess, and anybody who says I don't think that will happen, you'll say, well, you're pointing to the 40 percent that's still below the line because we have uncertainty. And I understand you're betting on that 40 percent, but we're worried about the 30 percent. So that's the way we're going to go.

The point is you're never going to get the answer from doing the numerical calculations. All you get is an honest statement of what you know and that everybody can agree on. I think that's the big thing, is that everybody can agree this is the state of our knowledge. Therefore, if you're going this way and I'm going that way, it's because we're valuing different outcomes differently, and so the expected value comes out differently.

Another example here of a place for an opportunity for this is in the discussion -- well, actually, I'll let it go. But I think you get the idea.

DR. VENITZ: Let me just follow up to that, Jenny. Whenever you do an outlier analysis, which is really what you're trying to do, worst case scenario, what are the few that have a large change in QTc, distribution assumptions are very important in terms of what your final outcomes are. I look at your simulation slide. You're talking about a logarithmic distribution. I'm assuming you mean a log normal distribution.

DR. ZHENG: Right.

DR. VENITZ: How did you then actually simulate the changes due to the disease states or the drug-drug interaction? Did you just change the mean or did you change variances as well?

DR. ZHENG: Actually I just changed the mean because the model is fitted to the raw data. For example, the young subjects -- we modeled that. So we know the mean for that group.

DR. VENITZ: But what about the variance? I guess what I'm worried about, whenever you look at outliers and you have a change in variance, in other words your elderly or your renal population are probably more variable than your young reference population even in terms of

kinetics.

DR. ZHENG: We used the same model to model the data for young subjects and the data for elderly. I mean, the same compartment model.

DR. VENITZ: But in terms of your parameter variability, did you use the same variance in your --

DR. ZHENG: No. The data --

DR. VENITZ: So you used the actual data variance.

DR. ZHENG: Yes. I used the real data variance.

DR. VENITZ: Just to follow up on that, I'm assuming when you looked at your slopes, you assumed that the slopes followed normal distribution or log normal distribution?

DR. ZHENG: I did that analysis using NONMEM. So it's an additive model. So it's normal distribution.

DR. VENITZ: Well, based on what I've seen or based on the previous example, that may not be a good assumption. It could be that you just have a few outliers and have very steep slopes, but the rest of them have a fairly shallow slope. Whenever you do an outlier analysis, just as a general rule, the distribution assumption of the variances that you make really determine what your final outcome is.

In addition to that, I would reinforce what Dr. Sheiner said, and that is, I was missing the fact that you didn't really give us an idea of the uncertainty --

DR. ZHENG: Right. I think that's something I should have included. Probably I don't have enough information to speak to severely renally impaired subjects, what the relationship will be. But I do have the information about uncertainty of the parameter estimate. So I agree 100 percent.

DR. VENITZ: Just look at your three slides where you tell us what happens for the young individuals. Let's say the QTc of less than 10 is 41, 44, and 4 and 42.7.

DR. ZHENG: Right.

DR. VENITZ: Those are three different simulations. So you just do that a couple times and you know how much --

DR. ZHENG: Right, yes. The uncertainty of that estimate should have taken into that exercise. I agree.

DR. VENITZ: I think we have one more question.

DR. JUSKO: I imagine you're using the best available metrics on evaluating the exposure-response relationships. But you might consider using the availability of these data to examine additional

possibilities. For example, if you look at absolute changes in QTc, there might be the possibility that a change of 10 or 20 in the elderly is a bigger problem than a change of 10 or 20 in the young subjects. And perhaps something in relation to baseline values should be considered.

Secondly, you're using Cmax as the exposure index, and it would seem to me that in addition to that, the duration of time that a person has an abnormal QTc interval could be an additional hazard that could be factored in in exploring bigger sets of data as you may be doing.

DR. VENITZ: Okay. Thank you. Thank you, Jenny.

Before we go on a break, just an announcement. For those of you on the committee who haven't handed in your lunch orders, now is the time to do it or you're going to starve.

With that, we're going to reconvene at 10:10.

(Recess.)

DR. VENITZ: I'd like to reconvene the meeting please.

All right. While Peter is posting the questions that the FDA is asking the committee, are there any additional specific questions to the two presenters,

Dr. Zheng and Dr. Nguyen?

DR. KARLSSON: Yes. Just a question regarding this last presentation. The elderly showed quite a change when you looked at the distribution; 7.3 percent would be above 40 milliseconds. Maybe that would be mitigated by the renal impairment dose adjustment, although I guess even in the elderly population, it wouldn't be that many that's below 30 mls per minute in the elderly population. But I guess you could look at that through simulations as well.

But another question is, when looking at the percentage of a particular subpopulation that's outside, is it only the percentage within the population you're looking at, not the size of the population at all? Because I guess the elderly population is very large in absolute numbers compared to maybe severe renal impairment or ketoconazole.

Did I make myself clear?

DR. ZHENG: Actually could you repeat your second question?

DR. KARLSSON: Well, if we're looking at dose recommendations, is it only the percentage within the population that's interesting? Is it not also the size of the population as such?

DR. ZHENG: The simulation I did is 2,000 subjects. So it may change if you change the sample size to --

DR. KARLSSON: No. In essence, what I mean is that the elderly population is maybe like 80 million people in the U.S. and the severe renal impairment population is maybe 1 million. I don't know. 2 million. So that would come into play as well when making dosing recommendations.

DR. ZHENG: Okay. Yes, I think at the time we make a decision, we should consider the population who use the drug, the impact.

DR. LESKO: Mats, I'm not clear how you would consider it, though. Would you consider it in the context of saying that equal changes in a population that's larger number would get more weight in a dose adjustment scheme? Or how would you think about it as far as that issue goes?

It's like saying because the elderly population is so large, there's a greater overall risk to public health than there would be with patients with severe renal function. But it would seem in labeling a product, I'm not sure that would be taken into account for dose adjustment. Or if it is, I'm not sure how it would be.

DR. SHEINER: It would be if you were thinking about what dose sizes to make, for example.

DR. LESKO: Okay, from a manufacturer's standpoint.

DR. SHEINER: It's more convenience. If 90 percent of people are going to use this to make them safe

and 10 percent -- you know. They're the ones who are going to get out their pocket knives and hack the thing in half.

DR. VENITZ: Peter, do you want to review the questions for the committee one more time?

So those are the questions that we are asked to discuss with regard to the approach that we just two examples of using exposure-response information as a way of predicting probabilities of, in this case, adverse events as a way of deciding about dosing adjustments or not.

Lew?

DR. SHEINER: Did you want discussion on that?

DR. VENITZ: Yes.

DR. SHEINER: Well, I think we're back to where we were sort of in the very beginning. It's a good thing to do, but if it's not done with a little extra care, then maybe it's not a good thing to do. So I think it really comes down to that.

As I was saying to Jenny at the break, if you do a well-designed, even clinical experiment in which you know exactly the question you want to ask, you've got adequate data by good design, and you analyze it, in a funny way the statistics are relatively unimportant. The signal to noise is usually pretty high, and it's usually pretty clear what the result is. Yet, that's where most of the statistics that most of us have seen have been applied,

in making sure that type I error is controlled. And there's nothing the matter with that. It's a good idea. But it's not really where you need it. And it's not that you need sure inference here because you can't get sure inference when you're this uncertain.

But what we need is we need to have a good way of displaying what we know so that everybody is looking at the same thing and understands the uncertainties. It seems to me what we didn't see were two ways in which I feel that that needs to be done.

One, as soon as you generate a simulation model, you have to show me that that simulation model can simulate the data it was derived from. There are lots of different ways of going about convincing me of that. Some of them treat the data it was derived from as though they were new by leaving it out and then making the thing and remaking. There are many, many different techniques. But the fundamental idea is show me that the sorts of conclusions that you want me to draw about extrapolations are at least verified on the data that you built the thing from when you apply them to those data. So that's number one. I want to see a lot of that.

And then I want to see a real honest uncertainty in my simulation. We all understand we're not talking about anything that's sure here. But I want to

know how big the uncertainty is and I want to have some way of knowing where it came from. I personally really want to see model uncertainty as well as data uncertainty. In fact, I'm more concerned about model uncertainty than data uncertainty.

What do I mean by that? I mean if you have 100 patients from whom you've generated the data set, I understand the next 100 patients are going to have somewhat different numbers, and so you're going to get somewhat different conclusions. And we all estimate that uncertainty, and it's not that tough to estimate. And sometimes it isn't that large because we have a fair number of patients.

What's really uncertain is whether or not the relationship we discovered on this population is going to apply to that population. There we have no data if we haven't studied that population, if we're extrapolating to it. So there we need just some reasonable guesses. How different have populations been with respect to this kind of thing in the past with similar sorts of things? This is where the science comes in. This is where the judgment comes in. But you can build those model uncertainties in, and I can get to see how big they are. That's kind of like a robustness test. It's kind of like a way of saying how much will conclusions vary if I vary my assumptions.

Assumptions there will always be. I'm not against assumptions. What I'm against is making assumptions look like facts. It turns out that the things we don't know anything about we put the least uncertainty on, and that's very peculiar. We choose a form of a model. So it's a bi-exponential. And then we say, boom, that's it. No questions about that. And that's the thing we know the least well. What we do know well is the data we observed, and that we say, aha, that's got noise. So it's kind of like backwards. I want to see the model uncertainty.

This is not to be critical. I believe in modeling and I believe in trying to be quantitative about conclusions. But without that kind of thing, you won't ever get people around a table to agree on what you know, and if you can't do that, they won't agree on where you ought to go.

DR. VENITZ: Any other comments to the first question? Can we think of any specific examples or circumstances, therapeutic areas where this approach may not be applicable?

DR. KEARNS: Yes. I think one glaring one -- and Dr. Sheiner again speaks of model uncertainty. As I might understand it, it would be in the context of the facts of an experiment that we saw examples of. But as Dr.

Flockhart mentioned, for QTc studies where a patient may ingest a medicine that can have effects on its own, that's not necessarily part of model uncertainty. I don't know that you could predict the rate of co-ingestion of those drugs. And I would argue that with some combinations that are available, the relationships that you so nicely shared with us could look quite different.

So are there treatment circumstances that the approach might not be applicable as it was presented? Yes, and I think that's one example.

DR. VENITZ: What about the second question? I think that's something that we talked about last time. Differences in the exposure-response relationship between the typical and special populations.

DR. DERENDORF: Yes. I think that there are some examples in the literature where there are clearly differences in exposure-response relationship. If you think of benzodiazepines, for example, with the sensitivity and all the patient changes, so for the same concentration, you'll get a different response. But there's actually very little hard data available in the literature because it's hard to study. If you want to do it right, you have to do a complete PK/PD study, and just focusing on exposure is simply easier and therefore it's done more frequently. But I think that's what we need: more clean PK/PD studies in

different populations to see how much variability and how much systematic change we have in the exposure-response relationship.

DR. VENITZ: Larry.

DR. LESKO: There are actually two levels of uncertainty that we're dealing with. The first -- and I think it was the first example. We were talking about exposure-response relationship across various special populations. The second example illustrated a different problem and that was the assumption of exposure-response relationships between healthy volunteers and then patients. That's just the fact of the way, at least currently, drugs are developed. So we have to find ways to think about that, and it would seem there are two things I thought about.

One was in the pediatric decision tree or in the pediatric rule, we make the assumptions, or at least we ask the questions, about disease progression being the same in adults and kids and whether or not the mechanism of action in the exposure response is the same in adults and kids, and then we proceed down a path of logic that requires perhaps a dose being changed based on simply pharmacokinetic differences to achieve the same type of exposure.

It gets to the question, though, is my base

assumption that exposure response is similar in the absence of hard data, and then I look for reasons, perhaps mechanistic reasons, why it wouldn't be, or do I look and say, well, let me assume it's different and find mechanistic reasons that it should be the same? For example, in the pediatric adult area, you might ask the question, does a beta receptor's either density or sensitivity change and is it safe to assume that with a beta blocker I'm going to have similar exposure-response relationships?

It just seems to me that there's a way to think about it mechanistically if one understands the way the drug is working and the changes that are occurring in the special population. Like in the QTc, for example, if renal patients have altered potassium levels, then we know that affects sensitivity in terms of drug effects on QTc, and that could be kind of a rationale for including some assumption about heightened sensitivity or something like that or a change in the exposure-response curve. But in the absence of that kind of mechanistic information, it would seem we have to go with the assumption that these exposure-response relationships are the same.

I mean, does that line of thinking make sense?

DR. VENITZ: That would be the way I think. My default position is there is no difference between my

typical population and the special population unless I have either hard data to show that it is, which is rare, or I have mechanistic reasons based on the pathophysiology of the disease of that special population and/or the mechanism of action of the drug to suspect that it is. Then I either have to question the need for additional studies to show whether it exists or not or build it in as an uncertainty in my model.

DR. KEARNS: And Larry, I think another answer to your question that you posed is it depends on the surrogate chosen to assess effect. For example, if we look at studying a proton pump inhibitor in a child, there's convincing physiologic evidence that the maturation of the proton pump occurs very early and that the children respond to those medicines in ways that are very similar to adults.

But if you go pick a surrogate far from the tree of effect or drug action and you apply it and say, is gastroesophageal reflux in a 6-month-old the same as in a 46-year-old, and then try to make arguments about bridging, you'll find that the pillars that you've constructed the bridge out of are not worth traversing. So it depends on how close your surrogate is to where the medicine works.

DR. VENITZ: Something else I think we're going to talk about in a minute that I would also consider -- and I'm pretty sure you do that in your briefings with the

medical reviewers -- is what are the consequences of being wrong. In other words, what's the utility of whatever assumptions you may not be very certain about? Sometimes that severity or that consequence may be relatively inconsequential, and then it really doesn't make a difference. Forget the fact that you have statistical uncertainty associated with it.

DR. LESKO: There are many ways these kind of data are handled for purposes of dosing adjustment, and maybe that's one of the reasons we're trying to arrive at a standardized approach to doing it.

It would seem the safest way of doing it is to simply adjust the dose based on an area under curve change. The question then becomes what is the threshold level for that area under curve change to trigger that. And that's where the difference of opinion occurs because you don't have a method on the table that allows one, as people have said, to discuss this in a quantitative way.

So there may be, in essence, a lot of likelihood of not optimal dosing by doing it that way, either making dose adjustments when you don't need them or not making them when you should based on people's interpretation of the data without a methodology to discuss.

DR. VENITZ: But in the examples that you've

shown, the endpoints, as far as I can understand them -- QTc. That's a surrogate of fatal arrhythmia. So you're worried about a potential fatal consequence. There's a high, in my terminology, negative utility associated with it. On the other hand, things tachycardia or palpitations would rank much lower on the totem pole of my concerns. But I'm not sure how you quantitatively incorporate that short of using utility functions.

DR. FLOCKHART: This is an editing, small point. If you're going to talk about changes in AUC of a compound, I think particularly when you're talking about the QT -- but this may be representative of other things -- the area under the exposure curve is not necessarily the main thing. The time of exposure to a drug is not the trick. Parameters like the rate of rise to Cmax can be very important and the QTc max at a given dose can be very important. There are dis-relationships, blocks between the time-effect curve so the time of the concentration Cmax is absolutely not necessarily the time of the effect Cmax. It can be later. So it's possible there would be situations where a parameter other than a change in the PK AUC would be the appropriate parameter. It could still be a PK parameter, but it might be Cmax itself or it might be the rate of rise to Cmax. And that would be a drug-specific question.

If, for example, you looked at quinidine, quinidine has a very poor relationship to the QT interval. If you were able to talk about torsade, what really matters there is the rate of rise, how quickly you get to Cmax. And if you get to a very nasty Cmax very slowly, it's not a terribly dangerous thing it looks like, but if you get to the same Cmax very quickly, then it can be a very dangerous thing.

That's a drug-specific question, and I would caution about always using AUC. I mean, you can think of examples related to Greg's example too. Above a certain point, changes in the AUC of a proton pump inhibitor do nothing. They're meaningless. I guess that just emphasizes the point that you need to know the pharmacodynamic relationship first.

DR. SHEINER: True as what you say is, I quake at the notion that things as uncertain as area under the curves, which are essentially integrals and consequently smooth out error, and how you're telling us we're going to have to take derivatives, which augment error --

DR. FLOCKHART: Well, it's taking a smaller part of the data.

DR. SHEINER: We may never be able in sort of a naturalistic setting to estimate a derivative with any kind of accuracy.

It would work at the level of a preparation. If you had a preparation that was rapidly absorbed and one that wasn't, and that was pretty consistent, then you'd know that you'd have more danger from one than another in that sort of circumstance. But that we'll ever discover who are the people who absorb more rapidly by sort of surveying the world and then trying to put it together across several models -- and I'm the mad modeler.

DR. FLOCKHART: But, Dr. Sheiner, shouldn't that come out of some models? In other words, you would be able to see in a large population study whether people who get fast absorption get a QT longer.

DR. SHEINER: Well, I don't know the rate of rise because I don't know when their level was drawn. I put it on the graph at a certain point because that's what they told me, but I can have two levels that are 10 minutes apart and they're really 2 hours apart.

I guess what I'm trying to say is that -- and I was just being facetious, but I think we do need to temper the kinds of conclusions we hope people to draw from sort of messy clinical data. I'm just mentioning that derivatives are really hard.

DR. VENITZ: But that's the empiricist talking. As a pharmacologist, I say maybe I can understand something about what's responsible for orthostatic tachycardia and it

may well be that my rate of change in concentration or my Cmax is much more physiologically important, and I know that without having to do an empirical study.

DR. SHEINER: Right. What I'm saying is the implications of what I'm saying, to be serious rather than just making trouble, would be if that's the kind of thing you want to know, if you're not sure about it, then you need to do a very well-controlled experiment. You're not going to be able to learn that from the same kind of data you might be able to learn that area under the curve was the determinant.

DR. VENITZ: Either that or you have some mechanistic understanding how the drug concentration leads to a response. That's the point that I'm making.

Hartmut.

DR. DERENDORF: Well, I think this discussion shows that it's really impossible to even try to have a standardized approach in terms of a parameter like a bioequivalence approach, that you have a single criterion that would summarize it all up. I think each drug, each class of drug is different, and each situation is different. I'm not sure if we can find a standardized approach as we're asked to.

DR. LEE: I guess by standardized approach, we mean a standardized conceptual approach, which means we

always like to calculate the probability of an adverse event, rather than saying that we're going to standardized an Emax model as the method to be used or the magical AUC or Cmax. So, again, we're trying to standardize the conceptual approach.

DR. SHEINER: I think it's really worthwhile focusing on the positive here, which is this is a difficult problem and the very fact that people involved in regulation are acknowledging that it's worthwhile to try to be quantitative about things that are extremely uncertain and to try to come up with a better way to be more quantitative about a problem where there will never cease to be disagreement about any particular case because you'll never nail anything down close enough -- you'll be saying this is what we think we ought to do in terms of dosage recommendations. And it will be based upon an information base which would allow a rational person to say, no, you don't need to do that. That's where we're going to wind up, and to wind up anywhere else would be so prohibitively expensive that it would not justify the effort.

So I'm extremely encouraged. It's not as it has been in the past, hands being thrown up and we can't do this well enough, so we won't do it at all. That's not the right attitude. And that you're seeking advice on how to do this difficult thing I think is a very good thing.

But I do think that some kind of standardization, for example, about turning things into probabilities and utilities in an honest way so that everybody can be on the same page -- they can all understand what you know and what values you're applying.

DR. LESKO: A lot of the context for the discussion in the case studies so far have been what we've seen, what has come in in an NDA, but is there a way to translate the methodology we're talking about, let's say, to a drug development program in order to get studies designed that would provide for information that would reduce some of the uncertainty that we work with in the absence of some formal recommendation to do studies a certain way?

In other words, let's say a standardized method evolves over time, and let's say that that could perhaps evolve into a guidance on study design that would provide for information that would be better apt to provide the information that we're asking here in terms of dose adjustment and quantitation of risk. Is that a logical follow-through on the path we're on in the minds of people?

DR. VENITZ: But is the uncertainty and the consequence of the uncertainty that we currently have so large that we really need to do a whole lot more experimental work, short of what you're doing right now,

which is on a case-by-case basis, evaluate whether the information is sufficient for you to assess the risk-benefit, and then as in Jenny's case, recommend to the sponsor that they would have to do a larger study to look at high exposures?

DR. LESKO: I guess I was sort of asking the question -- in Jenny's case, for example, QTc was obtained for the first 4 hours. Would a study design that looked at a longer period of time -- wait a minute. Was that your case? Well, it was one case. Sorry, Nhi. I should know these data.

There was one case where the QTc was obtained for only 4 hours and blood levels were obtained for a longer period of time. And would a different study design have provided a better basis to make the recommendations that people were trying to make with the analysis of the data? That's sort of where I'm heading with that.

DR. FLOCKHART: I think the answer to that is obviously it depends. 4 hours might have been long enough for that drug, but there are other drugs -- haloperidol comes to mind -- where that would not have been enough.

To go back to my point, I think actually it is, Dr. Lee, very generalizable. I think there is a generalizable conceptual approach. My point about bringing up just sticking to the AUC was just to be educated about

that. I suspect that the AUC would very often be a valuable parameter, but you have to be open to using other things when that's biologically and pharmacologically appropriate.

DR. VENITZ: The only thing I would add is, as you've heard the committee talk about last time, as well as this time, I think there's a lot of favorable sentiment. As Dr. Sheiner likes to point out, it's better than what we currently have. It beats the competition.

The one thing that I would reinforce, though, is that it's very important to communicate it appropriately, and that has to do with all the assumptions that are being made. Are they verifiable to some extent or not? Do you want to err on the conservative side or on the more liberal side? So that the people that deal with the clinical pharmacology reviews interact with the medical reviewers. They may not understand the technical side of it, but they're the domain experts and they can follow those kind of thoughts. So it's really a matter of risk communication in my mind more than it is the actual process.

DR. KEARNS: And to pick up on a point that Dr. Flockhart mentioned too, it has to be driven by biological or pharmacological plausibility. To use an approach, a guidance across the board can create information that is

not factual.

For example, I had the occasion to look at a new molecule just last week with a sponsor to talk about a study design. Of course, they had received some input about that study design, which included multiple ECGs over time that was coincident with the sampling time for the pharmacokinetics. When I inquired as to the preclinical data about the ability of the molecule to prolong QT, about the only way that I could be convinced it could happen is if the structure could be inserted in the chart of the alphabet and somehow got between the letters Q and T.

(Laughter.)

DR. KEARNS: So what will we see when we do the trial? We do multiple ECGs, in this case, on children. What happens if we see a relationship come out of that that can be described by a host of models with all the appropriate variability nested in? Will we have proven something that wasn't shown by prudent preclinical testing, or will we be finding ourselves in the midst of yet another epi phenomenon that has implications about how the drug might be used?

So I think one has to use caution in making sure that when we do these things, we have good reason to do it based upon what we know. I'm not saying that we will always know everything up front. We clearly, clearly

don't. It's an imperfect science in an imperfect world. But to just cast it out there as indiscriminate use of an approach carries with it some liability that might not serve the public at the end of the day.

DR. DERENDORF: Just a follow-up comment to what Dr. Sheiner said. I fully agree that it is worthwhile doing it and it is a good thing to do it. But the standardization -- really my point was that it stops at the point where we say each drug is different and the more you know about the exposure-response relationship for that particular drug, the more we can use it to make predictions and the better they will be. That's a trivial conclusion, but I think that's where the standardization ends. Then from there on, really each case is different and needs to be dealt with individually.

DR. VENITZ: Would it be helpful, as far as the internal workings are concerned, to come up with a list of questions that you typically consider when you go through this process and for the committee to have a look at them? I'm not sure whether the approach is something that can be unified, but maybe the kind of questions that should be asked every time you do this can be found consensus on. Does that make sense to the committee? So maybe at a future meeting, the questions that you would ask, what surrogate markers do you have, what relationships do you

have, do you use areas or Cmax, those kinds of things that you go through every time that you have to review an NDA based on your experience.

DR. SHEINER: I think you can go a little further. I think there are sort of best practices. Maybe that's the way to think about it in doing this kind of thing. Since I don't know anything about anything in particular, I've been dwelling on the generals of showing clearly what you know and what you don't know and somehow checking your models against your current data and so on. So I think there are best practices in this and I think there are some things you can say in general, although I agree with you, when you get down to putting the labels on the x axis and the y axis, then suddenly you're in the domain area and you've got to talk to the right folks.

DR. VENITZ: Any further comments by the committee or any further questions from the FDA staff? Larry.

DR. LESKO: Yes. Maybe this is a deeper question and there isn't time to discuss it, but it does lead us down the path if we develop a standardized approach, the question that I have, in terms of labeling, comes into my mind. Right now we put in the label descriptive information, for example, in the clinical pharmacology section about a change in an area that

describes the, let's say, drug interaction or a special population change, and then if it warrants, a change in the dosage and administration section as to what to do about it. But if you have more data in hand, i.e., the likelihood of a risk or the probability of a risk or the probability of an increase in risk or other things that might come out of a standardized approach, the question would be to what extent would this information be helpful to prescribers or would it be a distraction to the prescribers and how can we enhance labels. Because we now know there are certain pieces of information that go into labels that at least the consumers, public and physicians tell us are not helpful to them and they can't interpret, and drug interaction seems to be one of those areas we frequently hear about.

So we're thinking of ways of improving labels in terms of consistent language, the scope of information that goes into it, and with this kind of standardized approach, it could lead to some interesting ways of revising labels to convey different information to prescribers and patients.

DR. VENITZ: Any comments?

(No response.)

DR. VENITZ: Okay. Then let's move to the third example for today, which is going to be presented by

Dr. He Sun. He is a pharmacometrics reviewer in the Division of Pharmaceutical Evaluation II.

DR. SUN: Good morning. I will try to discuss some general questions in my discussion, and we may switch the specific detail in numbers to general issues.

These will be the questions I'm going to ask after the presentation, but I just put it up front to get some initial feelings.

The first question is, if we get adverse reaction data from clinical studies, these data can be treated as either a continuous variable or a categorical variable. Then, what should we do? Do you have any preference, and why? I will show some examples to illustrate it further for this.

The second question is, in phase III clinical trials, lots of subjects, but we may not have observations in special subjects. Therefore, the population PK approach may give us the opportunity to either simulate or predict the exposure parameter for the population who don't have exposure observation but do have response observation. So what's the limitation and utility of this approach?

Now, if we have a PK model based on the above approach, we get some kind of conclusion on side effect versus drug concentration relationship. How do we make a dose adjustment recommendation for subpopulations?

There's some limited information here to present what data distribution pretty much looks like. On the slide here, in this corner it shows what the data distribution may look like. It can be dense data from phase II trials or it can be sparse data from phase III clinical trials, or some kind of a combination with an unbalanced situation.

Now, the safety information can be either a single critical key adverse reaction parameter which is a continuous variable, like QTc variable or high blood pressure and so on. But it can also be a categorical variable like pain or "yes or no" for liver toxicity and so on.

But these two actually are switchable. Let's say blood pressure. You can set up a cutoff marker that says if above such and such, it is abnormal, below such and such, it is normal. So the continuous variable actually can be changed to a categorical variable.

Now, a categorical variable, although it can be a "yes or no" situation, but for the group with "yes," you can also give a score of 1, 2, 3, 4, 5 or from 0 to 10. So it becomes some kind of a continuous variable.

So these two actually have no clear cut. That's why I ask this question in this presentation. If you have this situation, which one do you prefer? Of

course, this will also change your data analysis process.

So you can also have a combination of both with multiple ADR observations. But for phase III trials, pretty much we have this kind of situation: mixed types, multiple ADRs, and unbalanced. Then that is why population approach can play here.

Let me first show some data sets. Then we go back to see what analysis process we can apply. This data set is used just to illustrate the question or the process I mentioned before. Forget the exact numbers and the true terms. Sometimes I have to modify this.

Let's say we have two clinical trials, very big size, 1,500 evaluable treatment patients. And we also have multiple dose levels from X to 4 times higher. And the patient plasma drug concentration was measured, although for some are dense and for some are sparse. Therefore, the total data set is kind of unbalanced.

Then we have endpoints for safety measurement. This can be some blood chemistry variables which usually is a continuous variable at the beginning, but a clinician can define some value as a cutoff point shows this variable as normal/abnormal to claim at such situation there's no ADR and others will be ADR.

The ADR can be also present as a "yes or no" situation for some, like headache, liver toxicity,

phototoxicity values. And this variable again can be changed to different scores for the extent of headache, like mild, moderate, or severe.

Now, PK results. Let's say drug-drug interaction causes AUC to increase by almost 300 percent. The Cmax changes by 150 percent. AUC and Cmax may also be changed by age, gender, or so on and so forth, even between ethnic groups.

The safety results. We will not focus on efficacy in the presentation. We will only focus on safety parameters. Safety parameters usually are very, very small in percentage and very sparse. So there are some cases where you never have any so-called "maximal effect" for side effect terms.

The efficacy results. Let's make this discussion a little simple. We see efficacy has no such exposure-efficacy relationship detected although we see there's a demonstration of clinical efficacy in total.

Now, with these data sets, let's come back and see what process we usually can do. First of all, of course, there are managing/editing data processes we can use. For this part, we start to have a question: shall we treat the data as a continuous variable or should we treat the data as a categorical variable?

Then we can conduct a population PK analysis

based on exposure data such as building a base model, add variability, add covariate, and so on and so forth. Now, there's one problem: if we cannot find a significant covariate in the PK model, then the next step for predicting E for the new population or new individuals will havew a little problem. But let's say we have the model built and the model validated. Then we can go to the next step and determine individual exposure or subpopulation exposure parameters. There are two parts here. We can do post hoc for the subjects who are already included in the study, or we can do a simulation trial to determine exposure parameter for the population who was not really in the trial or the observation was not available in the particular patient.

The next step will be to derive secondary exposure parameters, such as AUC, Tmic, effect compartment concentration and so on. And another important factor here I want to emphasize is that the accumulative exposure parameter can be estimated and determined, like say what if after multiple dose or long exposure situation.

With these exposure parameters available through the above processes, we can determine exposure-response relationship for individuals or for special populations. We have lots of methods here. We can use classification method. I will give you some examples later

on. We can do classification or regression tree analysis, logistic regression for binary data, and so on.

Then we consider the accumulative exposure time, then perform statistical analysis on response data. Now, this again correlates with what do we do with the data? If our data is a continuous variable and we have odds ratios with uncertainty built in, we can do statistics. But sometimes if the variable purely is categorical and is divided to either above the mean or below the mean, the statistics will be hard.

Now, with all this situation, the next step, finally we will make a dose adjustment. What are we going to do especially if we have multiple variables? Let's say the exposure-response depends on age, gender, body weight, and blah, blah, so on. If we take all of this condition together for making a dose adjustment, it may be too complicated in drug labeling. Shall we only consider the one which is critical, or shall we consider the one which has most frequently occurred, or some other method? I want to hear some discussion on this.

Let's see some results. If we're dealing with this process, what result can we get?

Classification. The first part we can see is to divide the whole population into some equal populated segments. Let's say every 25 percent subjects from low

exposure to high exposure. Or we can divide this whole population into equally distanced segments, the percentile; that is, the first 25 percent in concentration, the second 25 percent, and so on and so forth. Then we count what's the frequency of ADR. For example, the results can be presented as total ADR is 18 percent if AUC is greater than the mean and only 5 percent if AUC is less or equal to the mean value. That makes the whole discussion for this kind of classification results. Of course, there are lots of pros and cons. I really want to hear a discussion later on.

Now, the second one is that we can do a classification analysis based on severity. Let's say we reclassified R values, the response values, as different class or different scores, as severe, moderate, or mild, and so on. Then we see the example. Severe ADR occurs if Cmax is greater than 10 but only mild ADR is apparent if Cmax is less than 2, although the frequency probably between these two has no significant difference. But in this situation, we see it looks like 10 is some kind of cutoff value to avoid severe ADRs.

Then we can throw this data into a computer to search for the best maximum split, maximum split distinctions between R values as by a classification tree or regression tree. One result I present here, for

example, the first split on ADR frequency was at AUC equals to 871. So when the split occurs on ADR, it was 23 percent if AUC is greater than 871 and would be 2.5 percent if AUC is less than 871.

Next, we will do modeling work. We can do modeling work for the same data set for different ADR or the same ADR parameters. When we do modeling work, there are several ways. I do not want to discuss further on this part. I only want to show the possible ways. We can base on purely statistical models to do the modeling work, or base on some kind of physiological-based, meaningful models. There's a lot of discussion on the pros and cons for each. But let's go to the next one. Our model can be a linear model or a nonlinear model. Of course, there are uncertainty parts when building nonlinear models, adding fixed effects and random effects, again, with this model.

So two examples. We can do a simple regression analysis based on so-called logistic regressions. So, for example, we can get a result with even a 95 percent confidence interval for the logistic regression results for odds ratios. For example, we can see the odds ratio for acute tissue rejection increases 23 percent if AUC 0-24 decreases by 10 percent. That's one way to present this data.

Then we can treat all the data as a continuous

variable, do as the next few examples for modeling work. We can get an equation that says the HDL drops below normal on day 95 if the concentration, average concentration, is greater than 10 micrograms per ml on day 95 for patients with high body weight and low initial HDL at baseline. So there's a covariate effect built in and it also has this kind of a drug concentration curve profile.

The back pain we treat as 0 to 10 scores, some kind of semi-categorical variable. It's nonlinearly correlated with plasma drug daily AUC and the number of treatment days and the dose regimen. So three factors. The result here found was b.i.d. actually has less incidence of back pain. T.i.d. will have more, although the total daily doses are equivalent. And the number of treatment days significantly correlate with or predict back pain scores.

QTc prolongation. Now, we can find some models. The relationship between QTc and the drug concentration can be described by an Emax model. Then you can find some E-R parameter like E0, Emax, ED50, and so on. But these parameters can be, as we discussed before, correlated with either gender or age or some other subpopulation variables.

Phototoxicity. Now, we only have two variables, yes or no. We get a "yes" value on day 10 if C

average on day 10 is greater than 8. This will occur only in female subjects. So this can be due to either data limitation or this really is a true result that female subjects are more sensitive to the drug in terms of having phototoxicity occurring.

Liver toxicity. We can get some kind of frequency linearly correlated with C in the initial few hours of time of exposure. It may not correlate exactly with the Cmax, but it correlated with the initial range of concentration average.

Blood chemistry. Now, blood chemistry actually is a continuous variable, but we can treat the blood chemistry variable as a categorical "yes or no," normal or abnormal, or give them a score from 0 to 10. Now, if we have a score from 0 to 10, we can get some kind of correlation with plasma concentration, either AUC or number of treatment days.

So these are examples I want to show. So one single variable can be treated by different ways, but in one study we have lots of different variables, and different variables may be treated by different ways, and use different data sets as different base for information.

And there are others, probably we never find any relationship, never can find a cutoff point. Descriptive. Here I give some examples. Some we just can

have a definition of normal/abnormal values, but may not see any difference between different subgroups like these four examples I show here.

So now we come to the end. We have all the information, by all different methods, with different bases, from different data sets, so on. We definitely already used population PK, did it two times. One is predicting exposure parameter for subjects who do not have observation in exposure but do have a response measurement, and in the second part, use a population approach, nonlinear mixed effect modeling, to show whether the E-R relationship depends on some other co-variables.

So with all this population of subjects and information, now we will make a dose adjustment. See, for example, we can do this: the average upper therapeutic limit is probably around 10 for Cmax and 871 for AUC. Remember, these two variables are gathered from previous toxicity analysis. And the female subjects seemed the most sensitive to phototoxicity, and the concentration average should be less than 8. Then we see a b.i.d. dose regimen is preferred because it reduced one of the particular toxicity results.

Now let's go back to my questions with all the data we have seen in the examples. First is what are the utility and general limitations of linking PK obtained from

population analysis to response endpoints? And what are the general considerations in E-R based dose adjustment for special populations? And should we treat this data as a continuous or categorical variable? What's the preference?

Again, I really want to hear a discussion more focused on the general concept and ideas based on the experience you have and let us know the pros and cons for each situation rather than focusing on the numbers because I have modified the values somewhat to make the presentation smooth. Thanks.

DR. VENITZ: Thank you, He.

Any questions about Dr. Sun's presentation before we delve into his proposed questions?

On one of your slides, you mentioned Cint. Can you tell me what that meant?

DR. SUN: This is the initial concentration exposure.

DR. VENITZ: Oh, initial concentration. Okay.

Do you want to pose the questions and then let the committee bat it around?

DR. SUN: Okay.

DR. VENITZ: The first question regarding the limitations of linking PK from a population analysis to response endpoints. Do you want to elaborate on that question what specifically you had in mind?

DR. SUN: Okay. This question was the utility and general limitations of using population PK for population for this analysis. As I mentioned, we have two places we can use nonlinear mixed effect modeling work for doing data analysis for this data. The first part is in clinical phase III trials, we may not have a concentration exposure measure for every subject, but we do have a response measure for every subject. In this situation, if we have sufficient information, use the population PK, get the model, then we can predict or estimate concentration or other exposure parameter for the population we see in the phase III trial. Or in some situation, patients only have one or two trough measures and we want to determine the total exposure and the time of exposure. So this is one place population PK can be used.

The second part is after we have E data and R data, either categorical or continuous variable, now we can use mixed effect modeling to see whether these two variables are related to each other based on some covariate factors.

So that's my question. What are the utility and general limitations on this if we do it this way?

DR. SHEINER: I can't decide if what you're asking is, is there a manual for how you treat any given set of data to come up with the conclusions that you're

going to find most believable. We can't address that. So when I look at that first question, your description of what you might do sounded sort of like something I might do.

The only thing I can say, the only serious general limitation about which even good data analysis cannot help you is the problem of confounding. Both the PK and the responses are endpoints, are outcomes, and whenever you try to relate outcomes to outcomes, you have the problem that you can't tell which way causality runs. And you base that conclusion, if you do, on external information in the way of science or other things. You can't tell it from your data. So that's the limitation.

Now, that doesn't mean we don't proceed every day, based on observational data, to make the most important decisions in our lives. We do. But we have to understand that in a regulatory context, there are other forces operating. You want to be cautious in certain ways. And that's the serious problem.

Most of the other stuff, it seems to me, that you brought up were technical issues. And I'm not sure that we really want to spend -- even though I'm a real techno-wonk when it comes to model-based analyses, I don't think you want to hear me dilate on that.

I think the basic thing there is if you get

different conclusions when you treat your data as categorical versus continuous or when you use one kind of a model or another, then there's something wrong. So you ought to get all the same conclusions. What happens as you turn data from continuous, if it's got a lot of information, to categorical, you're losing information. So some things will drop out. Some things will appear no longer to have relationships that did appear before to have relationships. That's got to happen as you limit the information in your data.

But other than that, I think you want to basically use the data representation that's most relevant to the people who are going to use it and that keeps the information, et cetera, all the good rules of modeling. But I don't think we can get into too many details.

So maybe if you have particular instances where you think, looking at data in different ways, the same data in different ways led you to very different conclusions, I think that's something that I might be interested in hearing about. Otherwise, I don't know what we can say in general.

DR. LEE: Can I rephrase the question a little bit? The reason we're making this presentation is because phase III studies actually present a unique opportunity for us to look at exposure and response, especially the safety

endpoint that we don't frequently observe in a phase II study which is too small to capture a rare adverse event. That's why we wanted to ask the committee whether a population approach would be a good approach to look at exposure response in the phase III studies.

However, there may be some limitation in terms of study design. For example, we may not have enough plasma samples or maybe the sampling time between the PK and the pharmacodynamic endpoints is different.

So this is the type of question we're trying to ask the committee, whether internal study design, whether there's any limitation to conduct such type of a population PK analysis. And if there are certain limitations, can we recommend to the sponsor to design the study differently in the future so that we can get a better quality PK/PD relationship out of the phase III studies?

DR. SHEINER: Let me just say one more thing about that. So you're talking about wanting to use this confirmatory study for learning purposes. And there are certain kinds of learning data elements that don't interfere with your design that would make your life a lot easier. Whether they're worth it or not, you can only say afterwards. But measuring things serially, whether it be toxicity or efficacy or both, rather than just the 6-month and 1-year endpoint, or whatever it is that was the primary

thing; measuring compliance, that is to say, what drug did they actually take; measuring PK.

I think, by the way, my guess is that adherence is more important an influence on outcome than PK is for most cases. But when they say measuring PK, you want to measure that in the case where adherence is assured so that you have those as two separate variables. And so on.

Basically the idea is measuring biomarkers, whether they be adherence or chemicals, you know, along the causal path from the prescription to the effect, and measuring them serially over time. That's the best you can hope for without changing the design radically. And if you want to be able to do these kinds of analyses, that's the kind of data that you need.

But techniques for dealing with missing data, techniques for dealing with other problems that arise, with mixed kinds of data, both continuous and categorical, and so on, those I don't think are essential. They exist. They make your life a little tougher or a more interesting, depending on where you come from. But they're there and you should be able to get the information out of the data.

DR. SUN: He was asking whether I have an example regarding how the data can be treated either continuous and categorical.

DR. SHEINER: And reach different conclusions.

DR. SUN: Yes, reach a different conclusion.

Let's say this situation. If you treat data as a "yes or no" situation and you divide the concentration distribution to be above the mean and below the mean, do you will find the only conclusion you can get is the frequency of "yes" when concentration or AUC above the mean is 23 percent. If lower than the mean, it will be 5 percent. That's all you can get.

If you treat it as a continuous variable, you can get some kind of sigmoid models, correlation between scores of adverse reaction versus the concentration. Then from the curve, you can pick up -- say you want to limit less than 10 percent of subjects has a score less than 2 -- a concentration. So this becomes a different decision. And the curve becomes a smooth curve. You pick up a point at which you limit two factors. Percent of subjects reach a score of XYZ. Compared with the first one, you only can get a result if above the mean will be such and such, if below the mean will be such and such.

Then in labeling, it will be different. In labeling, when make a dose adjustment, say due to drug-drug interaction, if the population Cmax change, still somewhat below the mean values, for the overall population, or you can get a feeling what the frequency of side effects due to drug-drug interaction will be. But in the second

situation, if you have a continuous curve, you can estimate when concentrations switch a kind of 10 percent, what's the percentage of patients will have a score of 2 or 3 increase by such and such.

So when we recommended these two suggestions to clinical or to the labeling committee for NDA review, these two really makes different. That's why the question comes. Any parameter really we can treat as one of, or we can switch between the two. And what really we want to do?

DR. DERENDORF: I think as was said earlier, whenever you move from continuous to categorical values, you throw away information. I think it comes down to a compromise that you have to make with the information that you have and how you want to communicate it. If you make it too complicated and include everything you know, nobody is going to use it. So you have to find a way to focus on the important things, but still come up with an accurate conclusion.

You have a great example that you can overdo it and make it too simple. One of the conclusions you have here, the total ADR is 18 percent when the AUC is above 1,200 and 5 percent when it's below. That may be true for the data set that you have, but it's totally useless for someone who wants to extrapolate it for a certain situation because obviously you can have very, very low AUCs that