DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

CLINICAL PHARMACOLOGY SUBCOMMITTEE OF THE ADVISORY COMMITTEE FOR PHARMACEUTICAL SCIENCE

Wednesday, October 23, 2002

8:10 a.m.

Advisors and Consultants Staff Conference Room
5630 Fishers Lane
Rockville, Maryland

PARTICIPANTS

William Jusko, Ph.D., Acting Chair
Kathleen Reedy, Acting Executive Secretary

MEMBERS

Edmund V. Capparelli, Pharm. D.
Hartmut Derendorf, Ph.D.
Mike Hale, Ph.D.
Richard L. Lalonde, Pharm. D.
Howard L. McCleod, Pharm. D.
Mary V. Relling, Pharm.D. (by telephone)
Lewis B. Sheiner, M.D.

GUEST PARTICIPANT

Richard M. Weinshilboum, M.D.

FDA

Peter Lee, Ph.D.
Larry Lesko, Ph.D.
Rosemary Roberts, M.D.
Arzu Selen, Ph,D,
J
rgen Venitz, M.D., Ph.D.
Helen Winkle

C O N T E N T S

PAGE

Call to Order:
William Jusko, Ph.D. 5

Conflict of Interest:
Kathleen Reedy 6

Welcome:
Helen Winkle 9

Introduction to Meeting:
Larry Lesko, Ph.D. 10

Topic No. 1
Consideration of Investigational Pharmacokinetic Studies to Identify Patient Populations at Risk: Methods Used to Adjust Dosing Given the Availability of Exposure-Response Information

FDA Presentation: Case Studies and a Model for the Future:
Peter Lee, Ph.D. 31

Evaluation of Methods and Clarifying Questions:
Richard LaLonde, Pharm D. 65
Lewis Sheiner, M.D. 77

Committee Discussion 114

Using Exposure-Response Relationships to Define Therapeutic Index:
Jrgen Venitz, M.D., Ph.D. 154

Topic No. 2
Use of Exposure-Response Relationships in the Pediatric Study Decision Tree: Questions to be Asked Using the FDA Pediatric Database

Medical and Clinical Pharmacology Perspective on the Pediatric Study Decision Tree and Experience to Date:
Rosemary Roberts, M.D. 195

Efforts to Optimize Pediatric Clinical Pharmacology Studies:
Arzu Selen, Ph.D. 216

Committee Discussion 231

C O N T E N T S (Continued)

PAGE

Topic No. 3
Scientific and Practical Considerations in the Use of Pharmacogenetic Tests to Determine Drug Dosage and Administration

Current Experience and Clinical Pharmacology Perspective:
Questions to the Committee:
Larry Lesko, Ph.D. 261

Assessment of TPMT Testing and Impact on Risk Management:
Richard Weinshilboum, M.D. 270
Mary Relling, Pharm.D. 301
Committee Discussion 314
Concluding Remarks:
Larry Lesko, Ph.D. 330

P R O C E E D I N G S

Call to Order

DR. JUSKO: Welcome everyone. My name is William Jusko. I am Acting Chair of this committee. We are calling to order the Clinical Pharmacology Subcommittee of the Advisory Committee of Pharmaceutical Sciences.

Dr. Lesko will be describing the functioning of this committee in a short time, but, as a way of beginning, I would like to have everyone introduce themselves. Let's begin over there with Peter Lee.

DR. LEE: I am Peter Lee with the Office of Clinical Pharmacology and Biopharmaceutics.

DR. LESKO: Larry Lesko with the Office of Clinical Pharmacology and Biopharmaceutics in CDER.

DR. VENITZ: Jrgen Venitz, Virginia Commonwealth University, currently on sabbatical with FDA.

MS. WINKLE: I am Helen Winkle. I am the Director of the Office of Pharmaceutical Science.

DR. DERENDORF: Harmut Derendorf, University of Florida.

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

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

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

DR. McCLEOD: Howard McCleod, Washington University, St. Louis.

DR. LALONDE: Richard Lalonde, Pfizer Global Research and Development.

DR. HALE: Mike Hale, GlaxoSmithKline.

DR. JUSKO: Thank you. We have two members who may be in contact by phone; Dr. Wolfgang Sadee from Ohio State University and Dr. Mary Relling from St. Jude Children's Research Hospital. The other member, Dr. Flockhart, was unable to attend today.

Kathleen Reedy will now read the conflict of interest statement.

Conflict of Interest

MS. REEDY: This is the acknowledgment related to general matters waivers for the Clinical Pharmacology Subcommittee of the Advisory Committee for Pharmaceutical Science on October 23, 2002.

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 today's meeting are issues of broad applicability. Unlike issues before a committee in which a particular product is discussed, issues of broader applicability involve many industrial sponsors and academic institutions.

All special government employees and federal guests 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 waivers to the following special government employees which permits them to participate in today's discussions: William J. Jusko and 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 12A30 of the Parklawn Building.

Because general topics impact so many institutions, it is not prudent to recite all potential conflicts of interest as they apply to each member, consultant and guest. FDA acknowledges that there may be potential conflicts

of interest, but because of the general nature of the discussion before the committee, these potential conflicts are mitigated.

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. JUSKO: Thank you, Kathleen.

Everyone on the committee has a copy of the agenda. The schedule for the agenda is laid out quite clearly. In relation to what is scheduled, at this point there is no one who has come forth to make presentations for the Open Public Hearing so will have the possibility of additional time for discussion or the possibility of moving lunch to an earlier time.

The first thing on the agenda this morning will be welcoming statements by Helen Winkle, Acting Director of the FDA.

Welcome

MS. WINKLE: Thank you. I would love to be Acting Director of the FDA. It is only of the Office of Pharmaceutical Sciences. Dr. McClellan might have some objections to that.

I do want to welcome everyone to the committee. This is really an exciting day for us. Larry and I have had the dream of having this subcommittee for quite a long time now and it is really good to see it come to fruition. We think that the committee will be an excellent way to discuss a number of really important issues that are focused on clinical pharmacology and other topics around that, and then be able to take those issues to our advisory committee for further recommendation and discussion.

I especially want to thank Dr. Venitz. Dr. Venitz has been on sabbatical with us for the last few months and has helped get this subcommittee up and running. When he is through with his sabbatical, he will then become an active member of the subcommittee. It is through his efforts and Larry's and others in his office that this subcommittee has been set up.

I am going to keep my comments extremely

short because it is a very, very long agenda here and I know you have a lot to accomplish and talk about. But I look forward to the discussion today and I look forward to future meetings of this subcommittee. So thank you all.

DR. JUSKO: Thank you.

Presenting at this point is Dr. Lesko, Director of the Office of Clinical Pharmacology and Biopharmaceutics.

Introduction to Meeting

DR. LESKO: I would also like to extend a warm greeting to all of the new members of our Clinical Pharmacology Subcommittee and also the guests that have agreed to come. We really appreciate your accepting the invitation to participate in this committee meeting and on the committee, itself. As I look around the room, I recognize the talent that we have assembled and the fact that all of you are busy in your own worlds, but to take the time and agree to participate in this committee is extremely exciting and we appreciate that.

[Slide.]

The Advisory Committee for Pharmaceutical Sciences has a number of subcommittees that focus

on specific topic areas. This one, of course is clinical pharmacology. It is the only advisory committee I am aware of that is focusing on these types of issues that have implications really across all of the therapeutic medical divisions in the center.

Clinical Pharmacology, as you know, is an office in CDER that is matrixed across these different therapeutic areas and a lot of the topics that we are going to bring forward to this committee will be of a general nature but with widespread applicability.

So it is pretty exciting and I hope that you will find that the topics we bring forward are important, relevant to you and the drug development and to regulatory decision making and we look forward to your input.

[Slide.]

I am going to set the stage for today's meeting and give a little bit of a framework for us. As Helen mentioned, we had planned to establish this committee for a long time and we discussed it publicly in May. We have proposed the formation of this committee which was heartily endorsed by the Advisory Committee for

Pharmaceutical Sciences.

What we said at that point is we wanted to assemble a critical mass of members along with guests that would provide us expertise external to the agency in the general field of clinical pharmacology.

We indicated there were three broad areas that we thought were important for us to focus on. These were not intended to exclude other areas in the future but, in the early days of this committee, we wanted to take a look at issues in pharmacometrics, pharmacogenetics and pediatrics, all three areas where clinical pharmacology plays an important role in the agency.

[Slide.]

The responsibility of the committee is very straightforward and, as I look at the people around the table, I am quite aware that we have interacted in many other settings and can appreciate what you can bring to the committee. What we are looking for in this committee is your advice and recommendations.

We hope to bring forward issues that revolve around the use of new data or emerging technology and ways in which we might apply that in

the regulatory environment in decision making and with regard to, of course, our public-health mission.

So we see the issues related to three broad areas within the Office of Pharmaceutical Sciences. We think this information from the committee will be important in regulatory decision making in our NDA reviews. We could easily imagine taking some of this information to policy under our good review practices and finally, because we are involved in regulatory research, we can imagine a lot of the issues and information filtering into our research program in the development of methodologies that can help in decision-making.

[Slide.]

Let me talk about what we plan for today and the topics and a little bit of background on them. The first topic is really the main course for today's agenda and we have allocated the most time for it. We want to look at the way we analyze investigational PK studies to identify patient populations at risk.

More importantly, we would like to think about methods used to adjust dosing in the face of this exposure-response information that comes in to

us. How is that best done? How is it best done in the context of limited information?

The context for this topic relates to the priority that CDER has in understanding the risk. For the purposes of this advisory committee, I will take risk and divide it into two broad areas.

[Slide.]

The first is risk assessment. I think of this as something we do in the context of our regulatory review where we attempt to get science-based estimates of a risk based by a special population who may be over and underexposed to a drug.

Of course, that can be a safety issue or an effectiveness issue. It is the responsibility of the office to look at this information and make proposals to the Medical Clinical Division in terms of dosing adjustments.

The second part of risk is risk management. Once we recognize a signal that may be relevant, how do we manage it? The best way we manage it is by looking at the need for a dosing adjustment and putting clear information in the package insert or in the product label.

[Slide.]

Risk assessment can easily be based on exposure-response relationships if that information is available and even if it is incomplete. We currently do this now in regulatory review. We have a range of quantitative methods we use to analyze exposure-response information. It may range from the simple methods, looking at mean values in a reference population and in a special population making a judgment about the differences and how important they are.

We also look at more complex methods. In the complex methods, which you are going to hear about today when Dr. Lee gets up here, is when we try to characterize both variability and uncertainty, in other words, try to bring a little more quantitative assessment to this risk in order to express it both internally to other disciplines but also to use in the context of do we need a dosing adjustment or not.

Variability, I have defined in this context as the true heterogeneity in the exposure or in the response. Uncertainty, I have differentiated that from variability. Uncertainty is the lack of knowledge about exposure or response and sometimes the two are intertwined in the types

of data that we see.

[Slide.]

Where we would like to go with this topic and is unrealistic to think we will get to there today, is to develop a standardized approach for our office in the risk-assessment area, particularly of safety.

We would like to develop standardized methods of identifying at-risk populations from clinical-pharmacology studies. The at-risk populations are the typical special populations that we evaluate; children, elderly, renally impaired and so on.

We would like to find a way to formulate the problem, identify the question, if and how dosing should be adjusted. And the third thing, as part of a standardized approach, is to specify the data, the quality of the data, that we need to look at and the methods of analyses. This has broad range of implications in what exposure information is important, what endpoints should be looked at, what assumptions and what models should be incorporated into this standardized approach.

I don't think I am saying we need a standardized method. I think we need a

standardized approach from which will stem different methods that reviewers would use on a routine basis.

[Slide.]

Let me give you an example. I have only picked this at random from the PDR. It is a resperidone label and it illustrates the issue that we will be talking about this morning. Resperidone, like many other drugs, has special-population information in the label. You can see that the way it is expressed is quite different from special population to special population.

In the first case, we are talking about a decrease in clearance, in the second case, an increase in free fraction and, in the third case, a change in half-life.

Is that the best way to express that information and how should that information be translated into a dosing recommendation. On the right-hand side, you can see the dosage and administration section of this label and what is recommended. In each case, with all of the different pieces of information included, the recommendation is the same, a decrease in dosing of

50 percent from 1 milligram twice a day to half a milligram twice a day.

I am not saying this is bad, or I am not saying it is good. I am saying can we make it better and be more specific in how we link changes in exposure to the dosing changes in the label and a way to do that.

[Slide.]

The method you will hear about this morning from Peter will take on the following features. It will start out by defining a response of concern. That might be a QTc prolongation. It might be a neutropenic reaction, whatever is relevant to the safety.

The next step is to identify a special population at risk based on changes in mean arithmetic exposure. But, beyond that, the proposal will be to look at the distribution of that exposure and/or the distribution of response and identify those patients at the high-end exposure using a critical cutoff value.

These would be the patients that would require a dosing adjustment, and we would like to look at a method to establish that cutoff value and identify those high-range exposure patients.

[Slide.]

We recognize that we don't always have ideal data in this circumstance. Oftentimes, and in particular with safety, exposure-response information is incomplete. This is in contrast to efficacy which is usually more complete in terms of exposure-response relationships.

So when we have this situation, the considerations that go through our mind in reviewing the data is to look at the frequency of adverse events at the available doses that have been studied. We look at the overall mean change in exposure in the special population.

In a little bit of the art, we look at the sensitivity or what we think to be the sensitivity of the patient subgroup and then come up with a recommendation on the dosing adjustment. This may not be as quantitative as we like it, but the data is incomplete.

Today, you will see some examples of this incomplete exposure-response information. One of the questions we are going to have is what are the best ways to deal with this in extrapolating beyond the known data when, in fact, the change in exposure in a special population goes either above

or below what we know to be the exposure-response data from the actual study.

We think there are ways to do this and we would like your input on that.

[Slide.]

We will finish off this morning with Dr. Venitz who is going to talk about a concept that I know many of you are familiar with called the utility function. In my mind, I think of utility function as a way of specifying the well-being of patients, but it also relates to the main theme of this morning and that is risk.

The two components of risk, I think, are the probability that an adverse event or lack of effect--we will call that harm--the probability that harm will occur and the magnitude of harm that results if the adverse event or lack of effect occurs.

So I think, again, it is a two-component part of risk as we look through these methodologies.

The other value of the utility function is an understanding of therapeutic index. I think we would like to understand that better and maybe even define it better because we certainly refer to

therapeutic index in several of our regulatory guidance for industry stopping short of saying what we mean.

So utility function brings in the notion of safety and efficacy or harm/benefit and it serves to identify as a visual method the maximum attainable levels of utility and, in some ways, is linked to dosing adjustments in special populations. So the two, while different, are interrelated.

[Slide.]

You will hear more about the specific questions and, after Peter and Dr. Venitz are finished, I will put specific questions on a slide. But, from my point of view, these are what I think the issues are for the program; are the proposed methods that you will hear today feasible and should the Agency pursue them further. How can the proposed methods you will hear about be improved in terms of a strategy and a way forward, or, what other methods should the Agency consider for dosing adjustments?

I am thinking of the work ahead of us and when we leave the committee what are the directions we are going to take.

[Slide.]

Let me move now to Topic No. 2 for today. If the first topic was the main course, these are appetizer topics because the time we have available for today don't do them justice. But we would like to bring them to the committee's attention to lay the ground work for subsequent meetings and we would like to get into this in a lot more detail.

The second topic is the use of exposure-response relationships in the pediatric study decision tree. You will see that today, and the issue for today is what are the questions that need to be asked of this database. It is extremely rich. It is loaded with good information, clinical pharmacology, clinical data. What would serve the public, the drug industry, the regulatory agencies the most in analyzing this data. It is a big task. We need to go in the right direction and we are looking for input.

You will hear from Dr. Roberts who is involved in pediatrics and has been for a long period of time and Dr. Selen from our office, also involved a long time. Both of them will be looking for your advice.

[Slide.]

To give you a little favor for this, the Pediatric Rule, or, as we refer to it now as the Best Pharmaceuticals for Children Act, despite the recent ruling of Henry Kennedy and the FDA's ability to ask for these studies, we have been using adult clinical data from controlled studies to draw conclusions about the efficacy and safety of drugs in the pediatric patient.

There is a logic to doing this. It avoids large-scale clinical trials in kids. It makes things faster. It expedites access to drugs for children. It is cost-effective. We are not doing big clinical trials and, for the most part, I think it has been successful and most people agree with that.

[Slide.]

We have a pediatric decision tree that we use in determining the pathway to bridging adult data to pediatric data. It is general. You have to read into it a bit but it clearly lays out pathways to extrapolate these data based on the different types of data; for example, clinical-pharmacology data, clinical-efficacy-and-safety data, and there are certain questions in that tree.

It is an addendum to our current draft exposure-response guidance. I think it was in the background. You will certainly see it in a minute.

The types of bridging studies that are utilized in pediatric decision is based on a key decision in the beginning part of this decision tree, the likelihood of two main assumptions being true. Admittedly, these assumptions are often deemed true or not true based on qualitative data, maybe subjective data. It is not always based on quantitative assessment but it based on judgment.

But, depending on the answer to those two main questions, the decision tree takes us down the path of doing safety and efficacy trials, PK or PK/PD studies. And it depends on what we know.

[Slide.]

Here is the tree. The two main questions are at the top. The key is is it reasonable to assume similar disease progression and similar response to intervention in the kids compared to the adults. You can see that if the answer to both of those is yes, one moves further down the tree to talk about exposure-response information.

It asks questions about are there PD measurements that can be used to predict efficacy

and, in each of those red boxes, the user of the decision tree focuses on a type of study or types of studies that would allow for bridging from the adult to the pediatric situation.

This afternoon, you will hear more about this. You will find out what drugs have been approved by what box. As I say, this tree has led us to a substantial database which has been systematically being organized. It is in the process by Dr. Selen. All of those on the Pediatric Initiative would like to know what can we glean from this database.

[Slide.]

We have issued over 250 written requests. There have been approximately 600 studies in these written requests. These involve more than 34,000 pediatric patients, nearly 60 approved active moieties which have been given exclusivity because of the Pediatric Rule. I think you will agree that this database represents a gold mine.

But, like gold anywhere, we have to figure out how to extract the most from the source.

[Slide.]

So the issue for the committee today is what can we learn from this database. If you were

in our position, what would you think about it? What would be the questions that would benefit the public health, therapeutics, drug development.

Once we decide on a direction and we have some ideas, we are going to move forward with the analysis of the database and hopefully present this in subsequent advisory-committee meetings.

You will hear today a description of the data we are collecting. You will hear today also about some main objectives of research into the pediatric database. One can imagine this research then leading to a possible revision of our pediatric decision tree and the change in the paradigm by which these drugs are approved.

[Slide.]

Again, I will go back to the main theme of today which is a risk-assessment theme and go back to the issues that were on the top of that decision tree. This is the type of research we are thinking about conducting. The issue of is it reasonable to assume a similar PK/PD relationship in kids as we have in adults.

We would like to look at methods and standards for both drug-specific issues related to this question as well as drug-class decisions.

Part of this decision tree is to conduct PK studies. We do that using either full exposure profiles, standard traditional PK or sparse samples. We would like to see more sparse-sample strategies used in pediatric drug approvals, but the question is can we get to a standardized study-design template for these studies that everyone can agree is an appropriate one and the studies become efficient and effective.

I don't think they have been entirely efficient and effective to date.

[Slide.]

Then we conduct PK studies in the decision tree to achieve levels similar to adults for the purposes of dosing. We would like to delve into that data a little bit more and evaluate trends and exposure in kids due to differences in PK. What are the critical factors? Are there break-points in the maturation of enzymes?

Can we make some generalization about classes of drugs that may minimize the testing in pediatric patients? What specific questions would be worth asking? This is what we are thinking about on this topic.

[Slide.]

Now we move to the desert of our menu today. Again, we are going to scratch the surface of a very important topic to the agency and that is the scientific and practical considerations in the use of genetic tests, not to diagnose diseases, not to provide prognosis of disease but to determine drug dosage and administration. That is part of the clinical-pharmacology question.

[Slide.]

We are going to use as an example, because it is one of the most well-understood examples, 6-mercaptopurine. We know it is given chronically to maintain remission in children with acute lymphoblastic leukemia. We have data on the extensiveness of its use in this disease state. We also know, from our survey data, that it is widely used in adults with GI disorders. That, by the way, is an off-label use. We won't talk about that data today.

But 6-mercaptopurine is activated by conversion to 6-thioguanine. That is where its efficacy comes from. It is deactivated by the enzyme thiopurine-S-methyl-transferase, TPMT. We know historically there are TPMT genotypes in the general population that have either low,

intermediate or high activity of this enzyme, and each of those special populations defined by the genotype are at risk.

[Slide.]

Something to think about with regard to genetic tests for TPMT polymorphism, what do we know? We know the clearance rate of this drug differs by a factor of 4 to 10 among children with ALL. We know that 6-thioguanine leads to cytotoxicity if it is in excess, if the drug can't be metabolized via TPMT.

We also know that tests, while they have been historically available in academic research-hospital settings where this is a focus of the research of that institution, have now become more widely available and commercially available and one of the barriers, availability, is being broken down.

So this raises new questions, not only for 6-MP but for other drugs in the marketplace as the science of pharmacogenetics evolves and advances. At what point do we begin to include this information in the package insert for the purpose of determining appropriate dosing.

It is not only a question related to

approved drugs but new drugs as well, although one might think, from experience, that older drugs approved in the marketplace might be better candidates for revision of labels based on genetic tests because of the history of knowledge that we have through actual therapeutic use.

[Slide.]

I am going to pause at this point. The remaining slides I am going to save for this afternoon as we get into this topic. I will give an introduction to it in more detail, but we wanted to get you thinking about it as we set the stage for the meeting. We will also hear from Dick Weinshilboum who has been involved with this topic for at least twenty years and will present some of his experience.

As we go beyond TPMT, there are other areas that we need to be thinking about in terms of relevance of genetic tests. Think about the large number of substrates we have in the marketplace for the enzyme 2D6. We know that there are poor metabolizers in the population with a high prevalence. 2D6 tests appear to be reliable, widely available, and questions will revolve around at what point does the evidence meet a standard

that leads us to put this information in the label for a prescriber.

I recognize there are a lot of issues here, but we need to talk about it. It is a pending issue. It is going to hit us very soon and we need to get some good input on that topic.

So, with that, hopefully I have set the stage for the three topics today and I will turn it back to our chair of the committee.

DR. JUSKO: Before we go on, are there any questions of Dr. Lesko regarding the functioning and activities of our committee?

No? Thank you, Larry.

The next presentation is by Peter Lee.

Topic No. 1

Consideration of Investigational Pharmacokinetic Studies to Identify Patient Populations at Risk: Methods Used to Adjust Dosing Given the Availability of Exposure-Response Information

DR. LEE: Good morning.

[Slide.]

The first topic we are going to talk about today is consideration of investigational pharmacokinetics studies to identify patient populations at risk. Basically, what I wanted to

talk about is how do we apply exposure-response information for dose-adjustment recommendations in special populations if we see the exposure change in these populations.

What I will do is I will present several case studies and also present a proposed measure that we can use to apply exposure-response information for dosing adjustment.

[Slide.]

As you know, most of the NDAs may contain anywhere up to twenty or more clinical-pharmacology studies. In these studies, exposure or intrinsic or extrinsic factors may either increase or decrease exposure of pharmacokinetics and we need to have consistent approaches to determine the dosing adjustment in this special population and also interpret the change or experience change in these special populations.

[Slide.]

Here are some examples of intrinsic and extrinsic factors according to the ECH E5 Guidance. We have drug-drug interactions. We have disease states which include hepatic or renal impairment. We have age differences which may include elderly and pediatrics. We have sex, ethnicity difference.

We may have full interactions. High-fat foods, grapefruit juice, are known to affect the pharmacokinetics of the drugs.

We may have a formulation difference and dose-regimen difference which may also change the exposure of the drugs.

[Slide.]

Here I want to give one example of change in exposure due to extrinsic factors. In this particular NDA, we have about eleven clinical pharmacology studies. As you can see, the difference in the AUC between the reference and the test can range anywhere from 0 percent difference, which is no difference between reference and test, to 60 percent difference between the reference and the test.

So the question is where should we adjust the dose? Should we adjust the dose at 20 percent difference in the AUC or 30 percent or 60 percent or anywhere beyond that?

[Slide.]

Some of our guidance offers a solution to that question, when do we need to adjust the dose. The first guidance is the Exposure Response Guidance which we published the draft early this

year. In this guidance, we state that, "Exposure-response information can sometimes be used to support the use, without further clinical data, of a drug in a new target population by showing similar concentration-response relationships."

But the question is can we establish a standard to apply the exposure-response information and can we establish a criteria for dosing adjustment based on exposure-response information?

[Slide.]

Another guidance, Evidence of Effectiveness Guidance, which was published in 1998, also states that, "If there is a well-understood relationship between blood concentration and response, including an understanding of the time course of that relationship, it may be possible to conclude that a new dose regimen or dosage form effective on the basis of PK data without an additional clinical efficacy trial."

Again, the question is can we establish a standard to apply exposure response? Is that a standard criteria for dosing adjustment?

[Slide.]

Another guidance, the ICH Guidance on Dose Response, also stated similar things;

"Concentration response may be useful for ascertaining the magnitude of clinical consequences of PK differences such as those due to drug-disease or drug-drug interactions or assessing the effect of altered pharmacokinetics of new dosage forms or new dosage regimens without need for additional clinical trials."

We have a similar question here; what is the standard and what is the criteria?

[Slide.]

There are other specific guidance, For example, the Drug-Drug Interaction Guidance, Renal Guidance, General BA/BE Guidance and Hepatic Guidance also state similar things, we can apply exposure-response information for dosing adjustment.

[Slide.]

Recently, we have drafted a Good Review Practice MaPP which is an internal document. In the this document, we have listed a number of questions we typically ask during our OCPB review.

One of the major questions here is related to intrinsic factors. What it says here is, "Based

upon what is known about exposure-response relationship and their variability, and the groups of patients studied, what dosage-regimen adjustments, if any, are recommended for each of these subgroups?"

So this is very similar and consistent with the guidance that I just mentioned earlier.

[Slide.]

In the same document, there is another question related to the extrinsic factors. It has similar statements. So, based on all this FDA guidance and internal documents, we propose that we should use exposure-response information for dosing adjustment in special populations.

So the big question is how do we establish our standards and is there any criteria, or consistent criteria, we can apply for dosing adjustment in the special populations.

[Slide.]

First, I want to give another example. We thought that this is a good example of consistent dosing-adjustment recommendations based on intrinsic or extrinsic factors. In this NDA, we have four clinical pharmacology studies. We have four interactions; food, renal impairment, elderly

or age difference and the gender difference.

In this case, the four interactions actually reduce AUC by 20 percent. The label states that drug has to be given before a meal to avoid the food interactions. In the renal-impaired patient and in the elderly, the changing AUC is not clinically significant while in the gender-difference study, a female patient shows a two-fold or double the AUC than the male patients and it turns out that the drug doesn't work in the male patients, which is consistent with the PK of the patients.

Another important or interesting point I want to mention here is there is a 20 percent change of AUC in both the food-interaction study and the elderly studies. However, the label is slightly different or maybe very different.

In the food-interaction, we recommend that the drug has to be given without food. The reason is that we are looking at efficacy in this case because of the reduction in AUC. We are concerned whether efficacy may be reduced due to the pharmacokinetic change.

On the other hand, in the elderly study, we see a 20 percent increase of AUC. In this case,

we don't have any safety concerns for a 20 percent increase of AUC. So we are looking at two different exposure-response relationships. For food interaction, we are looking at the exposure-efficacy relationship. For the elderly study, we are looking at the exposure-safety relationship.

[Slide.]

This is another example we thought may illustrate an inconsistent dosing adjustment in the proposed label. This is the proposed label but we correct that later on.

There are six studies have been conducted in this NDA. The food-interaction study reduced AUC by 40 percent and the proposed label says that it has to be given before a meal to avoid food interactions. In the male and elderly patients, the AUC change is less than 30 percent and the proposed label says that it is not clinically significant.

For the clarithromoycin interaction, there is a 70 percent increase of AUC and the proposed label states that this is a significant drug-drug interaction in the Precaution Section.

The mild hepatic-impaired patients, we

have an even greater than 70 percent, close to 80 percent, increase in AUC. However, the proposed label states that this is not clinically significant. So immediately, you see some inconsistency here comparing the hepatic-impaired and clarithromycin interactions.

[Slide.]

So there are several issues involved related to dosing adjustment in drug labels of NDA submissions. First, inconsistency in dosing adjustment is frequently seen, as I have shown in the previous example, in the initial label language of NDA submissions.

Exposure-response information needed for rational dosing adjustment is sometimes incomplete or unavailable in the NDA submission and, as a result, additional exposure-response analyses are usually required and conducted by the FDA reviewer to address the question of dosing adjustment.

Because we had to conduct the exposure-response analyses, standard for analyzing and interpreting exposure-response data for the safety and efficacy assessment of drugs will be beneficial to the decision-making.

[Slide.]

In think there are several considerations in using exposure response for dosing adjustment. First, we had to recognize that there is a limited availability of exposure-response data in the NDA. According to our informal internal survey, about 40 percent of the NDA has some sort of exposure-response data or dose-response data. However, the rest, or 60 percent, of the NDA doesn't have that information. So we are working on limited exposure-response data.

Second, we also need to consider how are we going to select and combine different exposure-response studies in the NDA to establish the exposure-response relationship. We also need to consider the quality and the quantity of data so that we can get sufficient power to establish that relationship.

In addition, model building and verification are also very important processes for establishing that relationship. Finally, interpretation of the data and also the criteria for dosing adjustment are also very important.

[Slide.]

So, to improve the current status, we propose the following. We propose to develop an

evaluate a standardized approach for the reviewer to quantitatively assess the impact of the exposure change on either safety or efficacy that results from changing pharmacokinetics due to intrinsic or extrinsic factors.

[Slide.]

This is a flow chart that we proposed for using exposure-response information for dosing-adjustment recommendations. When we receive an NDA, the first thing we like to do is to identify or qualify exposure-response studies. Once we have these studies together, we ask the second question whether these pooled study is sufficient for determining an exposure-response relationship.

If the answer is yes, then we go to the right-hand box. We want to define the goalpost for dosing adjustment based on the pivotal exposure-response information. However, if there is no available exposure-response information in the NDA, then we propose to use the goal post set in the respective guidance. These are the guidance I mentioned earlier, Hepatic Guidance, Renal Guidance, BA/BE Guidance and Drug-Drug Interaction Guidance.

In this guidance, there is a default goalpost set for AUC and Cmax. At the end of this presentation, we are going to raise several questions to the committee for recommendations. The first question is related to three of the boxes in this flow chart.

[Slide.]

One of the goals here is to establish, perhaps, a standardized output. The reason for a need for a standardized output is that there are many exposure-response models with a range of complexity, as Larry has mentioned earlier. It can be as simple as a linear model and as complicated as a series of differential equations. So we would like to establish a standardized approach to interpreting the exposure-response data regardless of the complexity of the model so that we can better communicate useful and understandable information to other disciplines such as the medical officer here and the biostatistician and so that we can facilitate rational use of exposure-response information in regulatory decisions.

[Slide.]

This slide illustrates a proposed method,

a generalized proposed method, that we may use to present the exposure-response information.

Basically, we want to present the information in terms of probability. For example, if we have two published, and one is a test and the other is a reference, for the clinical pharmacology studies, we see a change of pharmacokinetics or exposure from the reference to the test--in this case, the test population has a higher exposure than the reference.

At the same time, we have exposure-response information. We also know the distribution of the exposure-response information. Then we can combine these two informations and estimate the distribution of the response. In this case, the distribution of the response for the test population shifts to the right as a result of the increase of pharmacokinetics.

Then we will need to establish a clinically significant critical value for the response and, beyond that critical value, the response is considered clinically significant which is the vertical line shown here. Then we can integrate the area under the curve of the distribution which are the red areas and divide the

area by the total area under the curve of the distribution. This will give you the probability of a clinically significant response.

Based on this probability of a clinically significant response, then we can make a clinically relevant decision on whether we are going to make a dose recommendation for the test population or not.

So this is a process of interpreting the significance of a PK change. First of all, the approach is usually limited to interpolation which means we will interpret a change in pharmacokinetics only within the exposure-response data and we don't normally extrapolate beyond the observed exposure-response data.

Then we will resample pharmacokinetics and response of PK/PD data to determine the change in response as a result of changing pharmacokinetics. Then we will estimate the probability in the patient population with a response greater than the clinically significant critical value. Based on that probability, we will make dosing-adjustment recommendations.

[Slide.]

In the next few slides, I am going to present two examples where we can illustrate--we

can use an example to illustrate how we apply the approach for dosing-adjustment recommendations.

The first example is an oncology drug. The effectiveness response is time to death and hematologic and cytogenic response. The safety variable here is neutropenia. There are three intrinsic and extrinsic factors that may influence the pharmacokinetics of the drug which include drug-drug interactions, body weight and age.

[Slide.]

This is the exposure-safety results based on nonlinear mixed-effect modeling and regression model. This was done in sets. Basically, we have already identified the critical value of adverse events, which is a Grade 2 change of neutropenia. We calculate the probability of this adverse event greater than Grade 2 in all populations as a function of steady-state drug concentration and the age of the patients.

As you can see, when the drug concentration increases in that direction, you have a higher probability of an adverse event intuitively. If you take two cross sections along age, one at twenty years old and one at sixty-five years old, then you get two curves for this

relationship in the elderly and in the young patient.

[Slide.]

This is what you get. You get one curve, PK/PD curve, for young patients and a PK/PD curve for elderly patients. We are further looking at three different groups at different body weights. What is observed here is, for the young patient, body weight doesn't have any important effect on the probability of an adverse event. However, in the elderly patient, body weight has a significant effect on the probability of adverse events; for example, from 50 kilograms to 150 kilograms, there is an increase of adverse events of greater than 10 percent.

[Slide.]

Similarly, we are looking at the effect of ketoconazole, drug-drug interaction on Drug A. We are also looking at two age groups. Ketoconazole increases the plasma concentrations. However, that increase of plasma concentration doesn't cause too much increase of adverse events in the young patient but it does increase the probability of adverse events significantly in the elderly patients.

So, based on this information, we can make a clinically relevant judgment on whether we are going to adjust the dose in the elderly patient or for body weight or for drug-drug interactions.

[Slide.]

The second example I want to raise here is an antiinfective drug which has nonlinear kinetics in clearance. Several intrinsic and extrinsic factors affect the pharmacokinetics. For example, the elderly have a two-times higher AUC than young patients, a 40 percent increase in AUC in the renally impaired patients. In addition, ketoconazole caused an almost 100 percent increase in AUC.

[Slide.]

The major safety concern here for this drug is QTc prolongations. This plot shows an exposure-response relationship linking the change of QTc to plasma concentrations. Apparently, there is an increasing trend of QTc, delta QTc, as a function of concentration.

[Slide.]

Based on that information, we calculate the probability of QTc change at several critical values because we are not sure whether a 10

millisecond increase, 20 millisecond increase or 30 millisecond increase is clinically significant. So we calculate the probability of change in all cases.

For example, there is about a 25 percent probability to have a 20 millisecond change in QTc when the drug is given to the elderly patients. There is about a 10 percent of the chance that the elderly may experience a 30 millisecond or greater increase in QTc when the patient is given the drug at a clinical dose.

[Slide.]

Similarly, we are looking at a ketoconazole interaction. We also calculate the probability of delta QTc with monotherapy and combined therapy at a steady state. As you can see, the dashed line represents the probability of delta QTc at different critical values for the interactions and the solid line represents the monotherapy. It is clear that with drug-drug interactions, the probability of delta QTc, or QTc, increase is much greater than monotherapy.

So, based on this information, we can recommend dosing adjustment due to drug-drug interaction of this drug with ketoconazole.

[Slide.]

To summarize the above two examples. Safety assessment of intrinsic and extrinsic factors has become a routine part of the preapproval risk management. Exposure-response information provides a rational basis for dosing adjustment and estimating the probability of adverse events allows identification of the population at risk. A standardized approach for interpreting exposure-response data ensures consistent assessment across the review divisions and should improve the information in drug labels.

[Slide.]

This is a summary of current approaches for dosing adjustment in the FDA Guidance. The first thing we would like to is to set the "no-effect boundary." If there is exposure-response information available, then we will adjust the no-effect boundary according to the exposure-response data.

On the other hand, if that information, exposure-response information, is not available, then we will use a default goalpost such as 80 to 125 confidence interval, a 90 percent confidence interval, of the ratio between the test and

reference for AUC and Cmax.

The next step is, if there is a significant change in PK beyond that no-effect boundary due to intrinsic and extrinsic factors, then we will apply concentration-response relationship to determine whether there is a need for dosing adjustment. Should we have certain language in the Precaution or Warning Section of the label.

[Slide.]

To put it in the flow chart of both slides, this is what we recommend. The first question we ask is, if there is a PK/PD available. If the answer is no, then we will use the default goalpost for AUC and Cmax. If the answer is yes, then we ask the next question, whether that exposure-response information is sufficient to establish a no-effect boundary.

If the answer is yes, that will be great so we establish the no-effect boundary based on the exposure-response data. And then we ask if the 90 percent confidence interval of test and reference is within that boundary. If the answer is yes, then there is no dosing adjustment required for the special populations.

If the answer is no, we have to look at concentration-response data and see whether we need to do a recommendation on dosing adjustment put in the Precautions or Warnings.

There is a little box here with a question mark. That is when we have a PK/PD relationship, however we cannot establish a no-effect boundary based on the PK/PD relationship. The question is what do we need to do next. I will give an example in the later part of this presentation to illustrate the question here, and then we will ask the recommendation from this committee in terms of how do we deal with these type of issues.

[Slide.]

There are four remaining issues we would like to ask the committee for recommendations. I will go over one question at a time using several examples to illustrate the questions.

The first question is what are the acceptable study designs that provide reliable data to establish an exposure-response relationship for dosing adjustment.

[Slide.]

In the draft Exposure Response Guidance which we published early this year, we suggest two

different approaches. The first approach is to observe the plasma concentration attained in patients who have been given various doses of drug and relating the plasma concentration to observed response. So this is your typical dose-response study in which plasma concentration is obtained in patients. We want to relate the response to the plasma concentrations.

The second type of study is different. It is to assign patients randomly to the desired plasma concentration titrating doses to achieve them, which means to achieve the plasma concentrations, and to relate the concentration to observed response. This is usually called a concentration-response, or concentration-controlled, study.

The major difference between these two studies is that the first type of study randomized the patient to dose and the second type of study is to randomize the patient to drug concentrations.

I think, in general, we all agree that the second approach is better than the first one in terms of eliminating several potential biases in terms of data analysis and the results. However, the reality is that perhaps over 95 percent of the

time, we receive, in the NDA, the first type of study.

[Slide.]

So the question is, are there any specific considerations in terms of data analysis or study design for these two types of study that we should pay attention to so that we can eliminate or minimize potential bias due to the study design, itself.

I wanted to just present this table which is also in the Exposure Response Guidance. This table lists several considerations in terms of four different types of study design; a crossover design, a parallel design, a titration design and a concentration-control design.

I want to mention this table so that, perhaps, we can focus on some of the pros and cons of different study designs and see if there are any recommendations on special considerations so that we can eliminate, perhaps, the drawbacks of the typical study design we have seen in the NDA, which is typically a parallel-study design.

[Slide.]

The second question that we have here is how to model incomplete exposure-response data.

The first example I am showing here is a CNS drug. We have four different datapoints for this drug from four different doses. Theoretically, you can actually draw a straight line through these four datapoints.

It is also reasonable to connect the lowest point, the lowest datapoint, to the origin and to see a more complete exposure-response curve.

[Slide.]

The second example is just the opposite. This example shows also four datapoints, or five datapoints. But these five datapoints only illustrate the lower part of exposure-response curve. So the question is where does this exposure or the response lead to when the dose is increased beyond 40 milligrams.

[Slide.]

So the general issue is related to the previous two examples, because we see this type of data, incomplete data, a lot of times in the NDA just because there is a limitation of the doses that one can do in clinical development. So the question is, if we see an incomplete dataset, can we make any assumption in terms of the shape of this exposure-response curve, monotonous or

U-shaped, or can we make any assumption in the linear or nonlinear PK/PD relationship.

Also, when we see incomplete data, how do we make use of this data? Can we model the data? Can we make certain assumptions so that we can fit the data to an Emax model or do we always use a linear model? How about a sigmoid Emax model?

If we don't have a mechanism of action, can we use a polynomial just to feed the dataset?

[Slide.]

The third question is how to assess the risks and benefits of drug concentrations that are not contained with a known PK/PD relationship.

[Slide.]

This is the one example of cardiovascular drugs. In this case, AUC change due to different factors ranges from 200 percent to 80 times the increase of AUC.

[Slide.]

However, this is the only dose-response data that is available in the NDA at four different doses. The reference dose is 80 milligrams. So, you have a 20 percent increase in AUC, it will be 160 milligrams. But anything beyond that, we don't have exposure-response data to interpret or to get

the response based on the pharmacokinetic change. In addition, the critical value or the clinical significance of adverse events is beyond the dose that we have exposure-response data. So the critical value will be up here.

[Slide.]

So this is the question. What can we conclude for dosing adjustment if we don't have a complete exposure-response curve or we have a narrow range of exposure-response curve. In the previous example, the PK range of the exposure-response curve is less than the PK change due to different factors and the critical value is not within the range of known PK/PD relationship and the direction of the exposure-response trend beyond the observed concentration range cannot be determined or speculated.

Should we use the default goalpost in the respective guidance for these drugs?

[Slide.]

Basically, this is the question for this box. We have a PK/PD relationship. However, the PK/PD relationship is in a very narrow range of exposure so we cannot establish a no-effect boundary.

[Slide.]

So, what do we do? Do we use a default goalpost for dosing adjustment or should we request additional studies?

[Slide.]

The last question is how do we establish consistent criteria for determining the no-effect boundary or changing the pharmacokinetics for dosing adjustment.

[Slide.]

To establish a no-effect boundary, I think we need to do two things. First, we need to interpret the clinical significance of change in response and establish critical values. Second, based on the critical values, we have to estimate the probability of an adverse event and therapeutic response related to a change in exposures.

[Slide.]

So the question here is how do we establish this critical value? Is there any consistent way to do that and what are the criteria?

[Slide.]

Going by the example of the antiinfective drug where QTc prolongation is a concern, here we

have estimated the probability of QTc increase at different levels. So the question is what is the clinically significant change of QTc that would cause a safety concern. Is there any criteria that we can use to make that judgment?

[Slide.]

Here are some of the thoughts. Perhaps the criteria may depend on the severity of the adverse event. It may also depend on our experience on another drug in the same class or our experience on other drugs with similar adverse events. It may also depend on the sensitivity of the patient population to that particular adverse event. Finally, perhaps we can establish some sort of utility function to estimate the clinical significance of each adverse event and this will lead to the next presentation by Dr. Venitz.

[Slide.]

Finally, I want to thank the following people who have either provided examples in this presentation or provided their comment or suggestion on my presentation.

I think we have, perhaps, one hour after the break to go through the questions. Now, I want to give the floor back to the Chairman.

DR. JUSKO: Before we continue with the additional commentaries, perhaps there is the need for a couple of clarifying questions. I have one, in particular.

DR. LEE: Sure.

DR. JUSKO: In your slide where you say proposed standard outputs for ER results--it is about the eighteenth one in--you indicated that you would be dividing the distribution of AUC values from the high range over something else that would serve as the denominator and I wasn't clear what AUC values would serve as the denominator there. Would it be the total exposures for reference and test or just--

DR. LEE: The denominator is the total area under the curve of the exposure distributions. Let me go to that slide.

[Slide.]

DR. JUSKO: The way the slide is structured, it looks like you would be using only the test group.

DR. LEE: We would calculate--yes; the example is for the test, but we will calculate the same thing for the reference. But, in that case, the probability in the reference population will be

very small.

The example I am giving here is for calculating the probability of an adverse event in the test population, so this area under the curve will be the area under the curve of this distribution here. But we will do the same thing for the reference. In this example, the reference will have a very small probability.

So we will draw a line and calculate or extend this distribution to here and calculate the area under the curve beyond the critical value for the reference. As you can see, it could be very small in this case.

DR. SHEINER: It is just a fraction of the population that exhibits the response.

DR. LEE: Exactly.

DR. SHEINER: Or a greater one. I have a question about the same picture, or actually, I think it was the next one where you start to compute some kind of an optimal dose. Neither of the pictures there, the upper one which relates exposure to the frequency of adverse response and the bottom one which relates it to efficacy; is that right--on the left-hand side.

DR. LEE: This one?

DR. SHEINER: Yes, both; the one above and below, on the left, exposure versus--and frequency of something.

DR. LEE: Frequency of exposures. For example, it could be AUC.

DR. SHEINER: Ah; okay. Fine. Then, pretty much, the bottom one is this one that I have the question about which is that doesn't involve any uncertainty, as Larry mentioned earlier. So you are assuming that you know what the distribution of efficacy is and those dotted lines are inter-individual variability not uncertainty; right?

DR. LEE: It is inter-subject variability; yes.

DR. DERENDORF: Just another clarification. You also assume that they are the same for test and reference?

DR. LEE: Yes. That is a fundamental assumption. But when we do the review, we had to verify that, whether that exposure-response relationship holds true for the reference compared to the test published. Sometimes, it doesn't.

DR. DERENDORF: I think that is a very important issue because your decision tree starts

out with is there a PK/PD relationship available, that was the first question. That doesn't tell us anything about what it is. It can be available but it can look many different ways, particularly when you go to--the whole assumption, when you extrapolate from changes in exposure to response is that the exposure-response relationship is a given and known. If it changes, everything falls down.

DR. LEE: Yes; that is a very good comment. But, a lot of times, the reality is that you don't get different PK/PD relationships for different populations.

DR. DERENDORF: I think the reality is a lot of times, we don't know.

DR. LESKO: I was going to add to that because, if you think about drug interactions, a typical drug interaction is conducted in healthy volunteers and the healthy volunteers and, unless there is a reason to look at it, there frequently isn't any look at pharmacodynamics of any sort unless it is easily accessible or easily measured.

So the question could be how does that drug interaction translate into the patient who is the target patient for the drug in question and the drug that would be interacting. I am not sure how

we can deal with that, actually.

DR. DERENDORF: I think the focus of drug-interaction studies is mainly on the kinetics, traditionally. I think that is something really we need to look into if the PK/PD relationship changes as a result of a drug interaction or a special population. I think that is the challenge that we have, not just focus on exposure alone.

DR. LESKO: I think the art of this is to consider the protein-binding aspects and also the absence or presence of active metabolites in the test situation compared to the reference situation and then deal with that in a somewhat art way rather quantitative data on that information in terms of changes in exposure response.

DR. JUSKO: In one of your very last examples, where you talked about the cardiovascular drug with the incomplete range of doses, if you could show that one again. It is the third from the end.

[Slide.]

That one and the next one; in these studies, you clearly have an extremely wide range of exposures. The next graph that you show relates adverse effects in relation to dose. So I presume

there are no exposure data to accompany these studies because the obvious thing is to examine this relationship in terms of exposure which is the basis of a lot of what we are going to be talking about.

DR. LEE: You mean there is no exposure data in the dose-response study?

DR. JUSKO: Right.

DR. LEE: No, because this is a clinical phase II, phase III, study. We don't have exposure data available. This is a very rare event, so they require over 500 patients to get that.

DR. HALE: Peter, have you considered that the decision tree and the use of default goalposts might actually lead to the collection of less exposure-response data? Would there be actually some pressure just to see if we can show that we hit the goalpost on pharmacokinetics and don't worry about the exposure response?

DR. LEE: I don't know. If you use goalpost, then the criteria will be more stringent because if you exposure response, typically, you can widen that goalpost, so you will have, for example, in the label, less statement in terms of the drug-drug interaction. So I would imagine that

if you have a PK/PD relationship, you would like to use it.

DR. JUSKO: If there are no further questions from the committee, then let's continue with our presentations by committee members. This is meant to be evaluation of methods and clarifying questions. Richard Lalonde will be the first commentator.

Evaluation of Methods and Clarifying Issues

DR. LALONDE: Good morning, everyone.

[Slide.]

I have, I think, about fifteen minutes to offer some comments. I guess I will call them Points to Consider and, hopefully, this will lead to further discussion later on.

[Slide.]

Moving right along, I am offering some comments here on Peter's slides that I got a few days ago. Overall, essentially, the comment that I would like to offer is that the proposal, the general approach seems to be very logical. When I have discussed this with a couple of colleagues, we think that this is something that we would definitely want to support.

In response to one of the last questions,

we do believe that this opens up an opportunity to logically look at exposure-response relationship to set no-effect boundaries separate from the 80 to 125 which tend to be quite stringent. I think the argument of consistency across proposed labels from sponsors would be a definite benefit. We also see that in terms of consistency within the Agency. We certainly have observed, at times, difference of opinions depending on the groups that we deal with for dealing with labels and what is considered to be, let's say, an important pharmacokinetic alteration.

Once a consensus is reached on some of these key details, I don't know if this is the intent, but sharing this information certainly maybe as part of either a guidance or some other means would certainly help sponsors and FDA implement this in a more consistent fashion.

We have looked at some of these issues within our own drug development, so I think if we can speak the same language as we submit an application, I presume this would only help the different parties.

Just an interesting point, also, is that studies have demonstrated quite well that labels

are not very effective at preventing drug-drug interactions. I think you are all familiar with the terfenedine story, cisapride, mibefradil and the studies that have been done actually by different groups showing how, despite labels and "Dear Doctor" letters and a variety of warnings, that drugs were co-prescribed and this led to people really having significant adverse events.

So I feel this is a bit of the elephant under the table here. We are talking about the label and how we can improve the label. We should really think about does anybody else read this label except us and what we should do to increase the effectiveness of the dose adjustments that are recommended in the label.

I know the Agency is--obviously, this is a major concern in the proposed changes to the structure of the label, but what else can we do. It may be something that we can discuss later on. It is a bit off-topic but, again, I feel it is, as I said, the elephant under the table to a certain extent.

[Slide.]

This is the decision tree that Peter just showed a few minutes ago. I want to focus briefly

on a couple of points that were brought up already, but I think there are two sides to this.

As Peter indicated, to use the default goalposts on one side if we have appropriate PK/PD information to attempt to set a no-effect boundary. So, about these no-effect boundaries, with that adequate PK/PD data, the 80 to 125 would be used as per different guidance that are already out there.

[Slide.]

With PK/PD data, or exposure-response data, if you prefer, we would have the possibility of defining another no-effect boundary. As was pointed out earlier, the former is typically based on a mean change and the 95 percent confidence interval around this mean whereas the latter is based on the distribution of exposure and exposure-response relationships in the populations.

[Slide.]

This is shown in the slides that Peter showed earlier so this is the distribution in the populations and exposures and of response as a function of exposure.

[Slide.]

These include components of variability that are not included, if you wish, in the usual

criterion based on the mean. So there are some elements there that are different between the left side and the right side of this proposal. We can talk a little bit more about this, the idea, for example, that we are looking at a drug-drug interaction. Is there a specific population of people that may have a different response compared to, let's say, just the mean and the uncertainty around that mean.

The approach based on distribution of response seems to be very logical and I think, as Peter described, there are some examples. I would like to see some more because we have struggled with this also. We have not looked at it exactly the same way as the Agency but we have struggled with this and how to try to make some of these judgment calls in looking at the impact of PK variability and PK/PD variability on trying to provide some rational basis for no-effect boundaries, and the uncertainty, as was mentioned earlier, also.

This is Peter's slide also.

[Slide.]

Some other points; the question about some practical aspects of the proposed method. Peter

alluded to this, how to select the critical fraction of patients while taking into account the selected critical level of response. So how do we set that critical level of response, and also take into account the risk benefit for a particular drug therapeutic indication.

Keeping in mind that, depending on the area that we are concerned about in that tail of the distribution, we may or may not be able to estimate that very precisely depending on how frequent these occurrences are in the trials that we have in our database.

I believe we will hear more later on about utility function so the point I am making here is out of balance. For example, we will look at the increased risk. As we increase exposure, let's say, with drug-drug interaction or organ dysfunction, there may be greater benefit so how does one attempt to try to make that tradeoff. So I think we will talk a little more about that later on in terms of utility or cost function.

As I mentioned earlier, I think these are all interesting questions. Once we reach a consensus on this, it would be very nice to be able to share this across groups to foster a greater use

in regulatory submissions.

In response to, I think, some earlier comments also, this is something that I would say we do now very routinely to model exposure-response relationships for key responses in phase II-III trials. I think historically this approach was not as common. We would have looked at the population PK in phase II-III trials and maybe PK/PD very early in development. But now we definitely want to focus on exposure-response relationships looking at clinical outcomes--both of these are adverse-event effects--in the target population in the pivotal trials and we see this as an opportunity, as I said, to put a rational basis when we propose a label to say that here is the information we have on exposure-response, here is what we consider to be an important factor, here is why this factor may not be so important.

The recent, actually, approval of gapapentin for postherpetic neuralgia, I think, is another interesting example of the use of exposure-response relationship in regulatory decision-making.

[Slide.]

A few more points. This one here, I am

not sure if I know exactly what the Agency's plans are, so we will discuss this later on, I presume, but current labels generally report effects of intrinsic/extrinsic factors without necessarily making a recommendation about dosing adjustments. So, for example, we report a drug-drug interaction, say, the exposure increased 30 percent and it is not necessarily always accompanied with a dosage recommendation.

So are we looking to make a change and offer a dose recommendation for all studied factors, keeping in mind that the default 80 to 125 goalpost is quite conservative. People who do these kinds of studies readily recognize this, so this is probably fine if we are trying to claim that a dose adjustment is not needed using this equivalence approach conservative because, to remind people, in order for the 90 percent confidence interval to be entirely between 80 and 125, the mean change typically has to be in the range of 10 percent or less.

So many people who are not routinely involved with these studies don't really appreciate this. You don't typically see a study show no effect in having a point estimate of, let's say,

123. That is essentially almost impossible.

Some other practical aspects that we struggle with also when looking at this in the equivalence world, what would be the dose adjusted, if any, for the following situations based on the default goalpost, or any other goalpost for that matter, but when we have, let's say, a point estimate that suggests that, really, there is no mean difference but we don't have a lot of confidence in this number.

So we have not met the regulatory standard of claiming no effect but I would be at a loss to recommend a dose adjustment because the mean difference is really essentially 3 percent. So you could argue that this was a badly designed study--I made up these numbers, of course, but these things happen. At times, these are the data that we deal with maybe because of the limitations of doing trials in patients. Maybe this is not practical to study in healthy subjects.

Another situation would be where we have a change on average so we fail, again, to meet the equivalence criterion to say there is no effect. But the 19 percent change for most drugs would often not be considered important. So, again, I

think it speaks to the very conservative nature of the 80 to 125 criterion. There aren't too many drugs where we would typically say lower the dose by 20 percent.

There are examples, but relatively few. So these are challenges that we deal with at times.

Another factor that was touched in briefly in one of the slides by Peter, should the dose adjustment take into account the patient's current dose. If a patient is taking essentially the lowest dose that is recommended and there is an increase in experience of 50 percent, is that a different story, that someone is taking close to the maximum recommended dose in terms of risk.

So that leads now to should dose recommendations be based on the dose that the patient is taking as opposed to an arbitrary dose adjustment because of an extrinsic or intrinsic factor.

[Slide.]

Another interesting thing that we encountered recently that I want to comment on here, and I have no idea if this is an FDA policy or not, but dealing with pediatric dosing recommendations and so-called negative efficacy

trials. So I am talking about trials that are performed under the current Pediatric Regulations.

What I would like to propose is that sponsors be allowed to provide pediatric clinical PK information in an appropriate section of the label even if a pediatric indication is not approved.

We ran into some opposition here from the Agency to do this. I guess my proposal would be with appropriate wording about the lack of demonstrated benefit in children for a particular indication, that we include PK information and it could provide information to clinicians who choose to use the drug off-label.

I am not sure if this is completely impossible from a regulatory point of view, but I thought at first that at least there were a lot of similarities to other intrinsic/extrinsic factors in label for which we provide PK information without specific evidence of safety/efficacy, such as, for example, renal impairment. We just talked about the drug interactions, for example.

I just came across this paper recently. People in the audience here and on the panel who are working pediatrics probably know this very

well, that off-label use is very common in pediatrics so it seems that providing this information in the label would be consistent with the spirit of the pediatric regulations aimed at generating data to guide clinical use of drugs in children even if a particular indication was not approvable because, let's say, the drug didn't demonstrate the efficacy required to grant that approval.

[Slide.]

So, in summary, I am generally very support of the Agency's attempts to standardize methods for dose adjustments based on exposure-response data. I think there is a benefit, potentially, to the industry. I think it provides a rational basis for making these judgments as opposed to the infamous, "Let's ask one of our clinical colleagues and he will tell us that this is not clinically important," or, "This is clinically important."

I would like to see more examples to better understand the properties of the proposed method to define no-effect boundaries. I think, like a lot of proposals, the devil may be in the details. Maybe that sounds negative, but just to

try to better understand some of the properties and the subjective judgements that have to be made, the decisions about critical cutoff values, for example.

As I said earlier, keeping in mind that we are talking about the label here and that often this is not having the impact that we would like it to have, so what other measures should we consider to increase the effectiveness of the dose adjustments recommended in the label.

I think that is all I have. So, Mr. Chairman, back to you.

DR. JUSKO: Any clarifying questions needed of Dr. Lalonde? If not, we will proceed to Dr. Sheiner.

DR. SHEINER: Can I make a suggestion that we have a techno break, maybe move the break up, because it turned out that the media on which I brought my slides is not compatible with that machine so I have to boot up my machine and see if I can make it work. So maybe it would be more efficient for us to take our break and then come back.

DR. JUSKO: That would be fine. We are scheduled for a fifteen-minute break in the

morning, so we will do it now and resume at five minutes after 10:00.

[Break.]

DR. JUSKO: We will continue with our schedule presentations at this point. Dr. Sheiner will be giving commentary.

DR. SHEINER: Thank you.

[Slide.]

I want to echo Richard's sentiment that this is a very good idea, that beginning to think in a more formal way and a more careful way about exactly how we arrive at the doses we give and how we change those doses in light of differences among patients is, I think, long overdue and I think that we are poised at a point, in terms of both theoretical and practical knowledge that will allow us actually to make progress here.

So I commend you for being right on the forefront and asking the right questions and going after the right things. I think I am going to take the position I usually take which is kind of a theoretical one and try to give you a framework in which I like to think about these things.

However, I don't feel that the theory needs any apology because I believe strongly in the

statement that I heard once, I don't remember where, which is that the most practical thing in the world is a good theory. So what I think we have to realize is that dosing adjustment, based on exposure response, and dosage, based on whatever, are really part of the same thing and you can't separate them.

The issue just came up, for example, that are we really, here, supposed to be talking about the notion that, given that we have a desirable dose in some normative set of population and now people differ in their dose exposure relationship, are we asking the question what do we do about that?

That seems like a pretty simple question and we don't really have any problem with that. People differ in their PK and you know exactly where you want to be. Then you change the dose so that you compensate for the difference in PK.

But then we heard talk about no-effect boundaries and goalposts and suddenly, now, we are talking about what kinds of doses do we want to give people to make them better, not how do we want to adjust one person to get the same level or the same exposure as another person.

So I think we have to think about the whole thing and the special population just becomes part of it. So the question, I guess, that is being asked is are we ready for a standard approach, and to give my brief answer, I think, no; that is to say, I think there are ideas that we could have about approaches, about things we ought to ask for, but I think we are not quite ready to say this is how everybody ought to proceed lock, step, according to an algorithm.

Let me, though, paint the picture in the context and leave you not without hope because I do think there are some things that we can do.

I thought I would start with this. You have all seen these three questions that I always ask and I thought that, given that the ghost of Roger Williams still inhabits the place and he like these, I will start here.

There are three key questions that you ask before you do any inquiry whether it is dose-ranging or anything else. What do you want to know? How certain do you need to be? And what are you willing to assume?

If you can answer those three questions, and domain-specific individuals have to answer

those questions. Those are not technical questions. Those are questions about values and about what you want.

Then what happens is--there is another point here which is that the second and third questions, how certain you need to be and what you willing to assume, interact very strongly. The more certain you need to be, the less you can assume, in general. We will see why in a moment.

But the important point about this is once these questions are answered by the domain-specific people, by regulators, by physicians, by patients, some of them, then we can start to get down to that standard approach. Then we can start to get down to the technical aspects because all the issues after that are technical.

[Slide.]

So here are my answers for dose selection. What do you want to know? I would say you want to know dose response. I call that the response surface. Now, the distinction here is you want to know dose response, not exposure response. Dose is what you do so that is what you want to know about.

Exposure response turns out to be very useful in figuring out how to chose doses. I don't

deny that, but, fundamentally, you need to know what you need to do. And you need to know utilities. We have heard about these before and Jrgen will talk more about them later. My talk will serve as a bit of an introduction to that.

How certain do you need to be? I claim, not very. What are you willing to assume? I am going further than what you are willing to assume. I claim that you can't do this at all unless you are willing to assume valid scientific knowledge of PK/PD, unless you are willing to believe that there are mechanisms by which the drug acts and that you can trust that you know something about those mechanisms based on scientific inquiry which has preceded your activities in dose ranging.

So let me elaborate on these things.

[Slide.]

Decisions should maximize expected utility. There is a system, as you sort of heard already and will hear more, for making decisions that is a formal system. It tells us what we need to know and how we combine our knowledge in order to make those decisions.

I have a little notation. I am going to say D, are what I call decisions. So there are

many of them, so I have subscripted them. Y are outcomes and there are many possible outcomes. Utility is the subjective value of an outcome, it is what value you assign to an outcome, so that utility is a function of outcomes.

Expected utility is the average utility across all possible outcomes where each outcome is weighted by its probability under your decision. In other words, decisions affect probabilities of outcomes and the expected utility is just the average across all those possible outcomes, each one counted by as much as its likelihood under your decision.

If you change your decision, then the probabilities of different outcomes changes and so the utility of that decision changes. So there is a simple formula, the expected utility of a given decision, I, is the sum of the utilities of all the possible outcomes weighted by their probabilities under that decision.

The optimal decision is supposed to be the one, the decision, that maximizes that expected utility. So what is the necessary empirical information here? It is those probabilities. That is the empirical information. That is the stuff we

can all agree on.

The utilities, the transformation of outcomes to values is subjective. Those are, in principle, made by every patient, every individual who is going to make a decision for him or herself. Now, to some extent, especially in the health world, we generally imagine that we all more or less agree about utilities. You would rather be alive than dead, things like that.

So it is not too much trouble to assign sort of normative utilities, but the important point is that those are subjective. There isn't any data you can gather about what they ought to be. You can gather data about what they happen to be in a population.

[Slide.]

So the theoretical basis for combining these things in this way has been known for a long time and it has been known and presented in the drug-dosage literature for a long time, especially in a series of wonderful papers by John Wakefield and his colleagues. So it is all laid out there in exquisite detail. We have had this available to us. We haven't used it much, but there are some examples of where it has been used and I would

suggest that this is the place to start.

It is a complete theoretical framework. It is based on a Bayesian approach to things because whenever you are dealing with decisions, you have to be Bayesian. Testing is not part of decision making. Testing is a different function. It is checking out whether your notion about the world is right.

That is quite different than making decisions under uncertainty. You are not testing in that mode. In that mode, you are acting.

[Slide.]

So let's just talk about optimal dosage in a very simple example. We have a binary decision, treat or not. We have one binary efficacy so the drug is either effective or it isn't in any given individual and one binary toxicity, it is toxic or not. This, Jrgen and I did not co-consult here but I am using the same simple approach to utility that he is using. I am saying that the value of the single efficacy is equal and opposite in sign to the value of the single toxicity.

So, perhaps the drug saves your life but it might also kill you. The good things and the bad things that can happen are of equal value.

That is not too impossible but it is very unrealistic and idealized, and I want this to be an idealized example.

So, in that case, where the weighting, so to speak, the utility is exactly the same and they are only binary things, the natural measure of the amount of efficacy in this situation is the probability of the efficacy and the probability of the toxicity, and the difference between the two is the utility because they are each weighted equally.

So that is all we have to compute. What is the problem, then? The problem is that, of course, the probability of efficacy, given the treatment, is a function not only of the treatment but of the patient, of the dosage, of a whole host of other things that determine that relationship, and similarly for toxicity.

[Slide.]

So you have all seen diagrams like this. In fact, I often say that if you don't see a picture like this, then it isn't me giving the talk. Dose response is the probability of the outcome, given these various factors. So, on the left, I have a very idealized picture. The probability is going up in the vertical direction.

Patient factors, of which there are many but I just conglomerated them all on one axis, sex, age, weight, other drugss, et cetera, and dose is the dosage regimen, not just the amount but the frequency, et cetera, whether you take it with meals or not. It is whole program for how you take a drug.

So you can imagine that there is some kind of a surface. I have that thing in yellow which describes this probability of efficacy as a function of patient factors and dose. You have the same thing for the probability of toxicity.

Then you shift the curve of toxicity over to the efficacy one and what you are looking for according to the utility function here because the weights are identical. So I can just look at those curves, themselves. For example, for such an individual, a person who is at the origin of the patient factor, the right dose is the one that maximizes utility. That is the maximum difference between the curves so it is going to be right there and that is going to be the dose for that individual.

Notice if you go to the other end of the patient curve, the toxicity surface always is above

the efficacy surface. So, for that person, there is no optimal dose. The best dose is none.

[Slide.]

So the dose response and the curse of dimensionality. There are a large number of distinct dosage decisions, timing, et cetera. Each has multiple options. There are a large number of distinct patient variables that affect the relationship between dose and response and they each have multiple possible values. That generates a huge number of combinations.

The special-population paradigm is a kind of an attempt to reduce the combinations to a manageable number of homogeneous categories. So we have got renal function. We have got old people. We have got young people. And we imagine that, by doing that, we can actually make this problem tractable. We can actually figure out that there are only four or five categories we need to worry about and get it right for each one.

I don't think that is true. I don't think it is possible. I claim it is still impossible to study all the possible relevant combinations of dosage by patient type variables. You need something more than that. You need some kind of a

continuous function that maps from the space of patient variables and dosages to efficacy, toxicity and, ultimately, utility.

So the response surface that I showed you a picture of implies a kind of a parsimonious representation of dose response that smoothly interpolates and extrapolates between and beyond the necessarily limited data because you are never going to have the amount of data that you need to fill in every point. There is an infinite number of points on that surface.

So that is the goal. That is what I mean, the, by what we ought to be after, the big picture. Obviously, part of that picture is special populations, if you want to look at it this way. There are certain points along the patient-variable access, but the big picture is this whole picture. I think we have to keep that mind because everything that applies to choosing the dose for people that are not in special populations applies equally well to people in special populations. It is just that their surface has shifted.

The interpolating and extrapolating functions are assumptions. Now, they may be very good assumptions. They may be based on science.

They may be based on mechanisms. But they are fundamentally assumptions in the sense that we are not going to prove that the shape of that surface has a certain kind of a shape or the interpolation is correct on our data because that would mean filling in every point, and you can't do that. There are just too many.

So this certainty assumption tradeoff that I mentioned earlier hinges on the scientific validity of the assumptions. If the assumptions are right, then we have good certainty that we know that what we are seeing is what we are going to get.

If those assumptions are wrong, we could be quite distorted. So that is where this tradeoff occurs. So, if we need to be certain, if we claim we need to be certain, then we are going to have get a lot more data and prove a lot more things because we won't be able to make as many assumptions.

[Slide.]

So, now back to the second question, how certain do you need to be. Why do I say not very? Not very certain is okay because it is the current standard. We usually only test three or four doses

before we leave and one of these is almost always the one that is chosen to be the suggested dose. This is not in our special population. This is the original dose suggestion.

Preapproval special populations, as we heard, and observational dose-response studies are limited in scope and they are not often analyzed in a response-surface-compatible way, and we have some empirical evidence that a lot of labels, a lot of dosing, is wrong. There is a great deal of overdosing and I cite this recent work from CDDS.

For reasonably safe drugs, even though that is the case, that is not necessarily wrong either. For reasonably safe drugs, a wide dose range is tolerable so it is not a disaster that we can be a little uncertain about this. An unpredictable individual variation makes individual dose response uncertain no matter what.

A new person coming to you is always going to be different than what you expect to some degree so you have to tolerate that. You don't need to know, then, precisely what dose a person like that ought to get because you don't get any precise output.

Dose titration is also a standard part of

medical practice which limits the harm of the wrong initial dose. This is something that nobody speaks about but we all know it which is that we are not talking about getting the dose right in the label. We are talking about getting a good starting point and then we expect physicians and patients to monitor what is going on and to adjust on that basis, so the cost of getting it wrong is not very great.

[Slide.]

So what are you willing to assume? As I say, valid scientific knowledge of PK/PD. That comes in defining the response surface. So let me just raise a couple of technical issues in the response surface; the kinds of models, what are these interpolating and extrapolating functions? They have to deal with real clinical data problems because we are going to be estimating these things from real clinical data.

I have a little footnote there, that paper we wrote recently with Lee Ping Zhang, who is one of my fellows, illustrates this really rather nicely. What are the problems? The problems are dealt with these things called hierarchical statistical models. They deal with sparse data,

imbalance. Some patients have more datapoints than others. High noise because, in the press of clinical trials, we don't get everything. We don't write down all the times we did things right or do it exactly right, either.

These models allow essentially every patient to contribute to the overall picture rather than isolating each patient, estimating things from them all by themselves and then putting it together. So it is called borrowing strength.

Mechanistic structural models; this is where the science comes in. You put models forward that represent the science, the understanding. Those assumptions are ones that we can trust. When we use those kinds of models, then we can deal with other problems that clinical data arises, what is called informative missingness, that when the data are missing because of their value when patients don't show up to clinic, because they are sicker that day, and so they would have had measurements that we were supposed to take that were actually more abnormal than the ones of the people when they do come in. That kind of missingness can really mess up inference and, if we have good scientific models of what is going on, we can compensate for

that to some degree.

Use of biomarkers, knowing what to measure and how they relate to outcome and doing valid extrapolation, how do we go from situations that we have studied to situations that we have not studied. That is the whole point of the kinds of things we are doing here.

What else can we say technically about doing this? The measurements; they have to be highly informative. We have to measure clinical outcomes and they should be of all kinds. They can be categorical. They can be single. They could be delayed. We need to get good clinical outcomes. But biomarkers are going to be really crucial here. This is not the place to talk about it, but those are multiple longitudinal quantitative and dynamic.

They have huge information content. The clinical endpoints generally, if they are single or categorical, have very low information content. You can't learn a lot from them, so we are going to need biomarkers and we need to know how they relate to the outcomes we care about.

But, again, it doesn't have to be certain because we don't need to be absolutely certain here. We have to learn from natural variation

which means that, in all the clinical trials we do, we have to measure compliance, measure pharmacokinetics, measure multiple outcomes even if we are not controlling them. That allows us to build these kinds of models.

So that is the kind of changes that we need in the industry in order to really deal with this issue if we want to deal with it.

[Slide.]

How can the regulatory agencies help that? I have a modest proposal. I chose that deliberately. I hope that the analogy, the reminder of Jonathan Swift and his modest proposals is not to come to mind too readily here. How about saying that the NDA must offer a reasonable decision analytic justification for dosage recommendation, not making a standardized procedure yet.

Let's just say to the manufacturers, you have got to come to us with a proposal for dose, dose modification, special populations, all that stuff, you have got to come to us with a proposal that fits the rules for decision analysis.

Now, what do I mean by that? What is a not reasonable one. 5 milligrams is safe and

effective. That is not a decision analysis. 5 milligrams is safe and effective and 1 milligram is not effective. That is not a decision analysis.

What is reasonable? At the minimum, as I sort of illustrated, one benefit, one risk and they should both be continuous versus dose. This is an important point. Probabilities are continuous. They go on the entire line between 0 and 1 so they are continuous. Even if it is a binary event, the probability is continuous.

I want to see an analysis of utility that says as I move dose continuously, I get a continuous change in response if it is a probability of binary event, if it is the level of blood pressure, whatever it may be. I want something continuous so I can know where to go.

If I make this whole thing discontinuous, 5 milligrams versus 1 milligram, then I have only got two choices, 5 or 1. You have got to be able to interpolate and that means we are going to bring the science and you are going to bring in the reasonable model.

So the minimum is one risk, one benefit and some utility function. The utility functions don't need to be complicated. It could be fraction

of time above the MIC for an antibiotic, or fracture of time within the therapeutic range if that has been well established for another type of drug, or just the probability of efficacy minus the probability of toxicity as I illustrated earlier and as I have an actual real-world illustration but I haven't got the time to show it. But maybe we will want to look at those later.

What are the benefits of doing it this way? I think one of them that I don't list is that we will get a lot of ideas about how to do this from the industry before we set down in stone any requirements. It will start to come in and we will see which ones seem to work and which ones don't.

I am suggesting a period of experimentation, a period of learning, by everybody involve, what works, what doesn't, what is a good job, and sharing of this information between the regulatory agencies and the manufacturers.

But, in particular, if we did this and if it became a regular part of a drug approval, then we would be exploring multiple doses between and within individuals. That is something that we don't tend to do. Yet, you need individual dose response in order to be able to do this thing

really right.

Variation will be better assessed which will lead to a better understanding of the causes of variation and, perhaps, better ability to adjust doses on that because the variation turns out to be absolutely crucial. The kinds of utilities that you are going to put forward will say, I want to sort of pay a price for everybody who is above a certain level, let's say, or has a certain toxicity.

That means you need to know how variable those things are. You need to know how likely it is that people will vary with respect to their drug levels and hence their effects.

Biomarkers are going to have to be used and so we will start to generate databases for validation of biomarkers as surrogates which is, I think, a very important thing as we go forward in developing drugs. We don't know where those databases are going to come from.

It will encourage a metaanalysis of all clinical trials in the dossier because you are trying to put together this information across trials. That is the only way to build up the whole picture and maybe it will actually lead to more

rational therapy and better and more effective doses.

[Slide.]

So what are some regulatory implications? Here are some that just popped into my mind as I thought about this. You may have to approve doses that have never been tested because the optimal point on the response surface is not any place you actually put a bunch of people when you did your studies.

That, I think, has problems possible for issues around formulation. I don't know an awful lot of formulations, but there is something about stability of formulations and you have to have them for a long time and things like that. You are going to run your trials when you are developing a drug with a formulation that allows you to give multiple doses, like capsules with liquid in them or something like that. Then you are going to have a problem translating that into an approved formulation.

Interpolation, obviously that is going to be allowed. But what about extrapolation? Peter sort of raised that issue of where you have missing data on your curves, can you really go to those

places and say, "That is where we ought to be operating, a place where there is no data to the right or no data to the left?"

I don't think any of have problems when we are talking about a place where we have data to the right and left and we are just kind of interpolating between points. Interesting questions.

Explicit use of utility; that is really new, I think, for regulatory agencies. It will deal with the consistency issue and, in fact, consistency of dosage recommendations is only achievable through reduction of all these things to a common scale and that scale is a utility scale. But, how do we establish an expected utility standard? Do we say we need to have a certain amount of expected utility for a given drug, otherwise you can't recommend it?

That begins to sound like we are starting to only approve drugs that do better than the competitor. So there are a lot of interesting issues here and that is why I don't think we are really quite ready for making these rules yet and we need time to think about it.

[Slide.]

A couple of points that just came to mind as I was preparing this presentation, formal decision, Bayesian decision analysis deals with a lot of the issues that he brought up. This consistency thing. As I said, utilities is common scale, risk-benefit goal posts, critical values, no effective boundary. These are all attempts to be dichotomous about utility judgments. Let's just face it. We have to deal with utilities. Let's do it in the right way acknowledging that it is not yes or no, that as soon as you cross a boundary, it is suddenly not bad and, before that, it is all good.

We need to have these continuous functions which tell us where we want to be located. Otherwise, as soon as we are below a certain level or threshold, we don't know where we want to be if we have these flat utility functions that are just step functions. We don't want those.

Pooling data from multiple studies; it is required in a sense. It is built into the Bayesian perspective here and yet it is not something that is done as much as it ought to be and is an issue that Peter raised.

Peter raised the issue of power and we

know a study is powered. That power becomes totally irrelevant here. That is about hypothesis testing. It is how much data have you got and what can you conclude from those data. A standardized interpretation--certainly, again, under this point of view, the standardized interpretation is the expected utility and it makes sense and it is translatable across different preparations, different drugs and even different diseases.

[Slide.]

So optimal dosage decisions maximize expected utility. Decision analysis is the only consistent and coherent theoretical framework for decision-making under uncertainties. Nothing has come along that does better. Nothing has come along that does better so let's not reinvent the wheel. Let's use what people have been working on for fifty, a hundred years, and put ourselves in that context and say what does that tell us.

That is one of things I tell my fellows. It is the best thing that can possibly happen to you is that you are working on a problem and you discover that some other folks have been working on that exact same problem. If you just change the names, then their problem is your problem and they

have been working on it for a hundred years.

That is the situation here. There is a lot of information about decision analysis and how you go about doing it. So let's stick with that.

Utilities are subjective values of outcomes. Expected utility is an average over outcomes weighted by the probability under each decision. The set of probabilities is the drug's response surface. It is a function of dosage regimen, patient features and it is derived through experiment and observation and prior science, I should say, because response-surface estimation is best viewed as learning, not confirming. It is a way of putting together information. It doesn't involve power. It doesn't involve hypothesis tests. That is not what it is about.

It means that you are trying to build in all of your knowledge to say what is the best current state of knowledge and make decisions based on it. My modest proposal is to require phase II to III to develop an empirical basis for optimizing dosage according to a decision analysis which they formally present and which would be based on a clinically reasonable utility function.

If we do that for a little while, I think

we will get to see just where the hard parts are and where the easy parts are.

I'm not going to show the examples. I am done.

DR. JUSKO: Would anybody like Dr. Sheiner to clarify any parts of his presentation? Larry?

DR. LESKO: I don't know if this is clarifying or just a question because it is something that we encounter in sort of a statistical framework of using exposure-response data. That was one of Peter's slides where he talked about randomizing patients in a phase II or phase III trial to dose and then looking at blood levels as opposed to randomizing to a blood level.

In the first case, that is often viewed from a biostatistical standpoint as being exploratory, hypothesis-generating, something short of confirmatory. The second case is viewed as confirmatory and that gets in the way of utilization of information when you have these different dimensions of statistics in clinical pharmacology.

I wonder, in the context of what you said, how fatal a flaw is that when we have, as Peter mentioned, most studies being conducted based on

dose randomization?

DR. SHEINER: It speaks to the "how certain you need to be" issue. First of all, let me say there is very exciting work within the last decade on causality. I think we really understand causality. I don't mean the huge philosophical issue of causality but I mean the practical, everyday, what you and I mean about causality, the drug causes the toxicity, the notion of causality, and how do you infer causality from natural experiments.

We know how we infer causality from designed experiments. We randomize people. Half the people get it and half the people don't. We know if the people come out differently, the cause was what we did, although even working out exactly how you know that, what kind of a theoretical framework you need to be able to say, "That works," whereas just watching doesn't.

But the point is there has been tremendous progress on this. So it turns out that if certain assumptions hold, then measuring the drug levels that arise in the course of the variability among people, even including variability in compliance which generates more variability in drug levels,

not only pharmacokinetic but compliance.

If certain assumptions hold, you can say that the observed relationship is approximately the same as the relationship that would obtain if you actually set the doses to those various amounts, which is what we want to know about. But you have to look and make sure those assumptions hold.

Then there are ways of designing studies in which you can be more sure that those assumptions do hold because as soon as you know what it takes to draw your conclusions, you know what you need to do to make what it takes have to happen.

That is the long answer. The short answer is those data are usable but they are harder to use and they need more thinking about exactly what assumptions we are willing to buy. But, if we are willing to say we don't need--as I say, the competition is we don't do this job well at all. So any improvement, it seems to me, is a good one. The other stock phrase I always like to say is let's not let the best be the enemy of the good.

We are not going to get perfect knowledge from observational data and most of our information about dose response and exposure response is going

to come from observational data in the sense that we are going to take advantage, we are going to have to take advantage, of natural variation to generate varied drug levels and various input patterns and see what the results are.

But I am very excited by the fact that there is some good, solid theoretical work, people who I thought would never ever be willing to deal with those kinds of data, a guy like Butch Tsiatis who has been a statistician, now at North Carolina but formerly at Harvard, who was very, very much just, "You have to do controlled trial," is now doing work in causality with Jamie Robbins at Harvard.

The reason why he always stayed away from that and the reason why many people stay away is because it just was a morass. You didn't know whether you were right or wrong. There was no good solid theory. Well, the solid theory is emerging.

DR. JUSKO: Mike?

DR. HALE: I have a couple of questions or comments. You won't be surprised that I think that utility is a definitely a very valuable approach to follow. Have you given some thought as to how we construct utility functions. Who does that? Is it

a public-health perspective? Is it pharmacoeconomics? Is it the physicians?

The second; have you also thought about risk avoidance? Is maximizing expected utility the way to go or do we need to think about maximizing the minimum payoff here?

DR. SHEINER: Mini-max. Let me first say, again, the thing that I always fall back on when I get hard questions like that is what is the competition. What is the competition? We are already--if you believe in decision analysis, if you believe that that way of describing what happens when you make decisions is right, then we are already using utility functions but we are being explicit about them.

So I say let's try to be explicit. We might be embarrassed to look when we write it down as to what we are actually saying is our value system but that is still better than just making believe that somewhere inside of us in some intuitive way it all comes out right.

That doesn't mean that intuition isn't very important. It is absolutely crucial. We need people to make it public. So that is my first statement.

My second is that is why I suggested in the beginning let's let the manufacturer come forward with the utility function that he thinks will work and run the thing out on his data, simple as it may be. Let's not be too critical. Let's spend some period of time just looking at what comes and maybe certain places and certain diseases and certain things will emerge.

Where therapeutic range is reasonably well established, why not just make it be some function of the distance that you are from the therapeutic range and make utility be minimum within that range. Let's start there. So I think there are ones where we can start. MICs for antibiotics seems like an obvious place to start.

The other reason why I like this is because it is going to encourage people to actually think about it and then they will have to start to think about, is it AUC? We keep talking about Cmax. I think Cmax is absurd. A, we can't estimate it without a model and we are not willing to take models so we estimate it by the maximum we observe, and that becomes a design-dependent parameters.

If we sample very five minutes, we get a

different Cmax then if we sample every hour. So it totally worthless in terms of an estimator and I don't know how many drugs Cmax is important for. I can think if a drug that is toxic to a rapidly perfused organ, then maybe Cmax is important. But how many of those are there?

Digoxin; remember that famous digoxin, which is deadly. Cmax is totally irrelevant because it takes about twenty minutes to a half an hour to reach equilibrium with the heart. But we stick with that because we have never written down explicitly what we are saying the cost is of a Cmax that is more than something or other.

I think the first time somebody tried to do that, somebody else would look at and say, "That is ridiculous."

DR. LEE: Dr. Sheiner, there are two components in the utility function. One is for effectiveness and the other one is for safety. I am wondering if you put it the context of special populations, and I would say probably over 90 percent of the time, you see an increase of exposures in special populations.

In that case, would it be possible to simplify your utility function and just look at the

safety part and not worry about the efficacy because, if you have an increase of exposure, you would anticipate that efficacy will stay the same or better, but then what you worry about is the safety.

If you simplify that, then you can even go one more step and say, let's not worry about the utility part of it. Let's just worry about the probability of an adverse event.

DR. SHEINER: I don't think so. Even if you said that efficacy is monotone, so if you are going to increase the exposure, you are going to increase efficacy, or it will just reach a max and stay there--even if you said that, you would still need what you are calling your threshold. You would still need to say when does toxicity get to a point where we say we can't accept this, or that is to say that we need to ask people to do some kind of a dosage adjustment which, presumably, is some kind of a bother, so it has got some negative utility associated with it.

You would still need to have a value, a utility function, on the toxicity and it would have to, presumably, be in the context of the efficacy. Again, I agree, if the efficacy was totally flat,

then it would go out of the picture. But you didn't know that unless you studied it.

The other point was the point that Hartmut made which is we are talking about a response surface. There is just no reason a priori to believe that things that change physiology in such a way that they change drug levels might not also change physiology the way that they change responses.

I agree they are probably reasonably well separated. There are many cases in which, if I had to make an assumption, I would say they are unrelated. If that is one of those I had to make because I didn't have the data, I would go ahead and say that. But it would be nice to have a little bit of information on that.

DR. LEE: Let me ask one more question before we move on. I saw, on the slide you didn't present, actually, an oxybutynin example. It brought to my mind another question and that is, let's say, a standard approach is to look at an area-under-the-curve change, given what you just said about Cmax, although we look at that. But you look at an area-under-the-curve change and you say, "Okay; this has increased 60 percent."

But, along with that, it is usually a change in clearance of a drug related to inhibition of metabolism, et cetera. The usual dose adjustment is to change the dose based on an area under the curve. What, in fact, is going on is a profile in the special population that probably hasn't been studied in any kind of efficacy or safety study, and how would that profile change and its possible implications play into the decisional-analysis framework that you presented?

DR. SHEINER: I think it would be a wonderful exercise to say, okay, if I believe that I ought to change the dose based on AUC, what other assumptions must I be making? Again, a formal kind of statement of, this is the efficacy I am concerned about, this is the toxicity I am concerned about, this is the kind of picture I think exists, this is the utility I am dealing with.

Then you can just see exactly what you would have to assume for an AUC adjustment to be the right thing to do. Then you can scratch your head and say, do I buy those assumptions; for example, that efficacy will proceed along the same curve for somebody who has got a different AUC or

that my data are sufficient to say what goes on when the AUC gets into this range, what was this original thing based on, et cetera.

So I am sort of arguing that we don't yet know exactly how we want to proceed in terms of being able to say to somebody, "You don't know anything. You follow these rules, you will be okay." I don't think we are there. But I think we are in a place, and I think Richard pointed out, the industry and he and others like him are really thinking about these things.

I think if we give him a chance and encouragement and tell him--we say, "You got to do this." It has got to be some reasonable rationale. And then you don't turn around and shoot everyone down so no drug gets approved. That is the other side.

DR. JUSKO: Thank you very much, Lewis.

Committee Discussion

DR. JUSKO: At this point, the schedule calls for committee discussion. It would be useful for Peter to put up his main slides, probably starting with the flow chart, the decision tree, and then we will go on to the specific remaining issues, questions to the committee that were posed.

It would be best if we did these in the same order.

DR. LEE: Dr. Jusko, do you to see the flow chart or the questions?

DR. JUSKO: We want to, in logical order, consider the main questions that the Office would like the committee to address. It is my interpretation that these questions to the committee are the secondary questions and your primary questions pertain, first of all, to the use of the decision tree and your standardized output method.

DR. LEE: Yes; these are more specific questions. So do you want to move to the flow chart, perhaps?

DR. JUSKO: Yes. It seems to me that the first question is for further commentary on the use of this decision tree for dosing-adjustment recommendations.

[Slide.]

Richard, you had some comments on the use of 90 percent confidence intervals? Maybe you could restate those.

DR. LALONDE: The point I was making was that when you go down the left side, and we use the 80 to 125 default no-effect boundaries as we

currently apply them, we don't take into account--maybe it is implicit in there, but we don't really think in terms of the variability across the population in the same way that we are trying to incorporate when we go down the right-hand side.

I think we kind of do, but it is not really stated the same way. So, when Peter showed I think it is called the desired output and he has the distribution of the population of pharmacokinetic variability and the distribution of the population of exposure-response relationship, and then you look at the tail of that distribution in terms of outcome, to say that beyond this tail, there will be concern about it, I am just saying that, while I think there is a very logical approach, I am just saying that there is subtle, or not so subtle, differences between the left side and the right side. It just may be the nature of the beast.

DR. LEE: I would agree with your observation. I think this flow chart is what is stated in the current guidance, that if you have exposure-response information, you can use that information to establish a 90 percent confidence

interval or a no-effect boundary.

But the two examples I showed actually didn't follow this flow chart exactly. It was calculated in the probability of an adverse event. So we haven't worked out the technical details of how do you get from the PK/PD relationship to the no-effect boundary. That is something we need to work out technically. How do you get that value and what types of assumptions do you have to make in order to get from intersubject variability of exposure response to a 90 percent confidence interval of the mean value between the test and the reference?

DR. SHEINER: Setting aside, for the moment, that I don't think that this is the way to go, and assuming that you do, take a look at what that is. That statement, the 90 percent confidence interval, is a statement about certainty. A confidence interval is a statement about epistemology, how well do you know that something is within a range.

The 90 percent interval loosely translated--my apologies to all frequentists who will find this objectionable--it, loosely translated, says something about the probability of

your degree of belief that it is within that range. Why 90 percent? What degree of belief do you need?

I just claim you can't do this. If you get down to this level of detail without having an overriding framework in which you have got a justification for all your computations, then, suddenly, you are in a place where you are doing things arbitrary like saying 80 to 125. It works; that is to say, you make the rule, they do it.

But it is just arbitrary. It has no justification in any way that you can get everybody to agree on. That is the same thing there. How can you put 90 percent down? Why do you need to be that certain? Why not 85? Why not 50 percent? Why not 99 percent?

You have got to show me some value in being 90 percent rather than 95 or 85 for me to buy that number. Now, the notion that you want to have uncertainty as well as variability in this whole process, that is absolutely correct and the Bayesian decision analytical framework has it right there and has it there and has it there explicitly and it does this right computations with it.

DR. LEE: Dr. Sheiner, what you are proposing is we go on two different paths. One

path is if you don't have a PK/PD relationship, then you go for the goalpost, 90 percent confidence interval. But if you have a PK/PD relationship, you don't think about the 90 percent confidence interval. You look at a utility function.

DR. SHEINER: No; I am going for one path. I am saying it is time to say to the manufacturers, "You present an argument within this theoretical framework that provides a basis for what you would like to recommend."

I am saying, in the beginning, now, as the regulatory agency, you be very generous about accepting those arguments. But the goal, eventually, is to have every dose have a rationale. Some will be better than others, but, again, there you would expect that you would want to be more concerned about those where the losses are greater.

DR. DERENDORF: I think the rationale may be to think, well, this is a similar situation as bioequivalence and, therefore, the rules that have worked there traditionally probably work here, too.

But it isn't the same thing as bioequivalence because it is a completely different scenario. If you have two patients with very different diseases, different physiology, that is a

different situation than a crossover study in a healthy subject.

So I think we need to clearly separate here the pharmacokinetic and the pharmacodynamic issues and we need to separate--even within the kinetics, we have to make certain assumptions that we may have different assumptions that we may have different disposition of metabolites that may be active or distribution issues that, if we compare between subjects, the simple ratios don't apply anymore.

DR. CAPPARELLI: I would just echo some of those concerns with the tightness, I think, that was brought up of the goalpost intervals. When I look at from the standpoint of pediatric subpopulations, if we took the data that we have for drugs when we are looking at pediatric dosing based on a milligram-per-kilo basis, for the most part, we would have a different dosage in almost every age group.

It would be very difficult to implement, without scientific rationale, for why one is making those sorts of distinctions. I think you would run into some problems, at least with that particular subpopulation group.

DR. HALE: This paradigm strikes me more or less as a static situation with regard to the data. In other words, you have got a package of data; what is the best you can do with it? It doesn't strike me as quite appropriate if you are in a situation where you can go do new studies, collect more data.

I agree completely this, at first glance, may feel like bioequivalence but it is so different in terms of, say, comparing a capsule versus a tablet. You are really talking about, if I give this patient A or B, are they going to expect the same AUC and Cmax. That is very different as opposed to having some kind of target AUC or Cmax. We don't know if those are the appropriate levels for a given disease condition.

I guess what is bothering me here, for instance, for example, if we find people with renal impairment have twice the AUC, is it an appropriate course of action to cut the dose in half. Well, I guess it depends on whether they have the same kind of exposure-response curve as other patients.

There would be a real temptation not even to go answer that question; in other words, maybe exclude those people from a phase-III trial and

just do a simple PK study to get what we need to know with regard to dose, if this is the paradigm.

DR. LESKO: I was going to say, these comments are well-taken. I would say, overall, the theoretical framework for a lot of this slide and a lot of the guidance that have come from the FDA over the last couple of years was an equivalence framework, equivalence approaches.

I think everyone acknowledged this isn't bioequivalence but the idea of an equivalence situation, not a tablet-versus-tablet, but a special population versus a reference population, sort of the fundamental approach here. These do appear in the guidance so I would put it in Dr. Sheiner's word, this is the competition and it obviously has some flaws.

To be honest, the way this has worked has not been very satisfying in practice because the default part of that, the box on the left, has only been useful in substantiating a claim of a need to not adjust dose. The reality is most of the studies that are done, whether it is drug interactions or renal disease or whatever, even if there is a modest effect or even a mild effect, you are going to exceed these so-called default

boundaries because of the number of patients in the study and the variability and so on.

So then it gets to sort of the other competition, how do you adjust the dose. It is nice when there is exposure-response data there. It is very satisfying to make a decision on adjusting the dose there but, when there isn't, it becomes basically the old way and that is looking at mean response differences and area under the curve and then thinking about the special population and the unique things that may make them sensitive in terms of that PK/PD issue, what may have changed. Then factoring all of that in, a decision is made.

But the reality is it only has worked well when there has been no interaction or no disease-state effects, or nothing uneventful.

DR. SHEINER: Larry, I have two questions. First of all, I am immensely sympathetic with the idea of cutting out little parts that you can do, getting some practice with it and then putting it together. So saying, let's address the simpler problem of we already have a good dose in people who don't have renal disease or hepatic disease or are not old and how do we figure out what the right

dose is for the old people and the people with renal disease or hepatic disease.

I think that is where this sort of comes from. I understand it. The only caution I would have is that, very often, as you start to work on one little piece of the pie, it turns out you just can't do it. So, for example, here knowing how much deviation from the usual exposure you will permit before you require a dosage estimate involves utilities. You just can't get away from it.

So, suddenly, you are back solving a problem that you should have solved in the first place when you set the original dose and maybe that is what we ought to be talking about at some point is let's go back to--maybe it is easier, maybe it is not easier, to do this little adjustment equivalence problem but maybe it will be easier in the long run to go back to the very beginning and say, "How do you choose a recommended dose? What do we require for that?"

That is what I am saying we want to have a nice decision-analysis argument even though it need not be totally complete or most modern or whatever. Then the rest of it, I say, will follow quite

easily rather than trying to come in from the periphery and finding that we run into these problems that we haven't solved because we were trying to avoid them.

But now I have just a technical question from what you just said. I don't understand, how does exposure response bear on the question of adjusting dose? If we believe we know exposure response, as I said in the very opening remarks that I made, then what we need to do is know dose exposure in each subgroup and then we will know what to do to change their dose to get the exposure that we have already decided they ought to have.

So, exposure response is irrelevant to adjustment of dose in special populations unless, as Hartmut is pointing out, maybe you have got a different exposure-response relationship in those groups.

DR. LESKO: I was, actually, thinking of this when Peter was doing his presentation because if you do a special-population study, your exposure measure is blood levels. When you fall back on exposure-response relationship, if you have PK/PD data, then you can interpret the PK part of it.

Often, however, and Peter mentioned the

statistic--I think he said something like 40 percent or whatever of NDAs have exposure-response information, that probably needs a little qualification as to what we are talking about there. But the bottom line is you have some sort of dose-response data on which you try to interpret the exposure changes in the PK studies.

So I guess that leads to another step in this process and that is do you take dose-response data from your phase II and phase II studies, but a little bell-shaped curve around the doses that have been administered and figure out what the average blood level ought to be or should be from that dose and also what the distribution is, and then use that sort of revised curve to interpret the PK data in your special populations, because, in essence, you have two different inputs on the exposure side that you are tying to blend, somehow, in making this decision on dose adjustment.

DR. SHEINER: My answer is simple. Measure dose and exposure. Set dose, measure dose and measure exposure.

DR. LESKO: Exposure being blood levels.

DR. SHEINER: Yes.

DR. JUSKO: The question about whether

there is a consistent exposure-response relationship across special populations remains a big frontier to be studied further. I sometimes give lectures where I point out specific differences, PD differences, in special populations. It is easily possible to come up with examples of gender differences, ethnic differences, differences in relation to obesity.

Pregnancy is a big factor that can cause marked differences in relationships between exposures and responses. So, while what Lew stated at the beginning, that a suitable starting assumption is that the exposure-response relationship is similar across populations, we really to do more work to ascertain whether that is true for drugs of particular critical importance.

DR. LESKO: To just add on to that, I think the topic this afternoon sort of will get into that on the pediatric side because one of those questions at the top is is it reasonable to assume I have a similar response to intervention. I think that is basically saying is the PK/PD the same in terms of disease progression.

That decision is often made--it is not entirely clear how that decision is made in each

and every case. We may hear about it more in the afternoon but it is almost like asking the question again, what is my default position. Do I assume it is the same in the absence of other information or do I assume it is different and now I need to be shown otherwise.

I think the same approach comes into play in special populations in general. I will assume it is the same in the absence of other information. I think that is reality. Is it perfect? No. I mean, we would like to do it differently and we need to figure out ways to get that information.

I think we do. In the cases of an easily measurable endpoint, in special-population trials, you will see some PD data. But if it is the longer clinical outcomes, we may not.

DR. SHEINER: I think the point that Bill just makes and that Hartmut was making earlier is absolutely--it sort of gets to the center of the issue, what are you willing to assume. I was saying, first guess, assume that PK and PD are indistinct. Clearly, we have many examples where that is not the case.

So sort of the right way to go about that is to build in that uncertainty, if you are

uncertain, into your analysis. You can either do that by looking at sensitivity--if I am going to suggest a dose adjustment and the PD might be this different, how wrong could I be? So you can do a sensitivity analysis or you can just build it in and say, okay; I am not going to make Lew's assumption and I am not going to say I know nothing. I am going to say, they are probably similar but they might be, and you ask the experts--they might be different by as much as X. Build that in into your model for what is going on and see what the utilities come out to say.

Does it still say it is worthwhile to adjust the dose in that case or does it say you might be hurting--you might now. So there are ways to do this within this context. That is what I am really trying to see is that there is a framework in which you can ask all these questions.

Then you invert the framework and it tells you what do you need to know? What is the crucial piece of missing information? At the moment, what is the thing to which your conclusions are most sensitive? That is what you need to go get information on.

DR. JUSKO: Before long, we are will be

hearing much more about practical aspects of use of utility functions. I guess the question that will come up then is how much of a retrospective could you do with the FDA's database to demonstrate that this or any other approach based on a decision analysis would be an improvement over the present approach.

DR. LESKO: My impression of what data would be needed to sort of take this down a path with a systematic sort of sound framework, I think that that is out there. And Peter has surveyed NDAs, knows better than I what is in it, but just thinking of an example I had picked at random from a lot of examples I could have chosen, respiridone. There is substantial information on dose response with that particular drug, something like six or seven dose-efficacy relationships from two or three controlled trials, lesser so on the safety side.

But it is typical. I think there are examples there. And there are also examples, perhaps more recently, where somewhat of a therapeutic range has been put into a label and that kind of information may actually be a good starting point, either something that has been approved in the past or something more recent where

there is, again, information on exposure and response that could be put into a more formal decisional analysis framework.

So, to answer the question, I think the data is there. But Peter has been looking at this a lot, too.

DR. LEE: I would agree that there are plenty of dose-response or concentration-response data available in the NDA database. I guess my question is what would be the systematic approach to assign a value to a particular, say, adverse event. How do you do that? Can the committee give us a recommendation?

If you see the QTc prolongation, do you assign a 1 to the QTc prolongation or 1.5 or 1.2? What is the criteria compared to liver toxicity? How do you do that?

DR. SHEINER: My answer would be if you don't know how to do it, then tells you who you--you are talking to the experts and nobody knows how, nobody will tell you, that a prolonged QT interval of this size is this bad, in some scale of good-bad--if nobody will tell you that, then you have discovered something fascinating, that we are making decisions based on total non-consensus.

Then you would start to ask the question, would you need to know that. The reason I like the example is because it is a biomarker. I think biomarkers are what is going to turn out to be crucial in this whole business, that we will be able to get a lot of PD data on biomarkers and not an awful lot on ultimate clinical responses.

So we are going to operate with those biomarkers and say essentially if the drug is interacting with its receptor in the way we think, then we are going to guess that that is the right dose even though the link between that and the ultimate clinical response is only based on moderate amounts of empirical data; good science, but not that much empirical data because it is going to be hard to get.

But I think just asking that question, just saying, what are the measures of the people who measure for toxicity and what relative value would be assigned to them. If you find you have no consensus, then it sort of makes you realize that you are in a morass, and there is a place to start.

DR. JUSKO: I think it is time for us to switch to another slide. I think our comments on all of this indicate that the committee feels that

this approach is wanting and is a very strong indication that we do need to explore these improved approaches as we will be discussing.

[Slide.]

So, as indicated on the slide, what are the acceptable study designs that provide reliable data to establish exposure-response relationships for dosing adjustments. Peter also followed this up by posing the typical designs of the typical dose-response study and the concentration-controlled study designs as ways that are currently followed with the first, the typical dose-response study, being one that is performed approximately 90 percent of the time.

Comments from the committee?

DR. SHEINER: Let me speak up again here. First of all, I think we have to careful about the question. Reliable data are data that are gathered when they were said to be gathered from whom, measured well, et cetera. So I don't think we have any problem with reliable data. That is sort of good experimental laboratory practices.

You are talking about reliable inferences, what designs will give you reliable inferences given that they are providing reliable data. I

said a little bit about that before, but I think the key point, absolutely key point, is that any design can provide, under a proper analysis, reliable inferences, and not only that, but inferences where the uncertainty is reasonably well assessed.

But the tradeoff there is the less rigorously designed, the more complex the analysis has to be and the more assumptions you will have to make. But that is all okay. You can make assumptions as long as they are explicit. But it gets tougher and tougher to draw conclusions by the seat of your pants from data that are lacking in certain design features.

However, the most important lack, it seems to me, is the one we need to focus on which is you cannot draw any conclusions if you didn't measure it. The things that we do not routinely measure are actual doses taken, although we have mechanisms available for that.

We don't measure all the relevant biomarkers or at least a large number of them. Among those, I would include drug concentrations. It is a biomarker of a kind of the drug-effect relationship. And relevant prognostic covariates,

and they vary in time. So I would say we would be a great step forward if, in every clinical trial, we measured those things and then attempted to make some sense of it. After that, we can talk about designs that make inference easier. There the basic rule is anything you can randomize, you can do a pretty good inference.

DR. HALE: I would like to offer a couple of notions here, one of them being always to look hard at who wasn't in the trial, who was excluded, and who was excluded unintentionally. That is always one of my concerns when I do these things.

If we are going to do this for undesirable effects, be it toxicity, tolerance, whatever, I think we have to think very carefully about a regimen to make sure we collect the right sort of data, kind of echoing what Lewis has said.

What happens is things like QT interval or liver function, we can schedule those well in advance, at Weeks 1, 6 and 8, or whatever, the people are going to come in and do these measurements. It is the self-reported things, it is the things we don't know about, that happen who knows when. It happens in the middle of the night or on Thursday and you are not scheduled to go to

the clinic until the next Tuesday, things like that.

If we are going to get serious about developing exposure response for those kinds of events, we are going to have to figure out a better way to make sure we can capture them reliably.

DR. LALONDE: Along the same lines, I think whatever we can do to promote evaluation of adverse events in a more, I guess I would call it, quantitative or continuous fashion. I think, often, there are summary statistics provided or an integration of the presence of adverse event over the period of weeks and months as opposed to using all the information that is gathered over time.

We have certainly learned that lesson a couple of times and we have discovered the important relationships when looking at, let's say, for example, if, as Mike said, maybe you have a more systematic way to collect the information, and look at it in that way, also, let's say daily scores of some adverse effect of the drug as opposed to, yes, no other patient had this effect over the last month.

You can look at time course and look at better quantitating, I think, the exposure-response

relationships. I think when you get to utility, the information has become more--it is richer so whatever we can do to promote that, I think, would be useful both for regulators and sponsors.

DR. SHEINER: Let me add just one thing. Richard reminded me of it. Longitudinal data is extremely valuable. It is a little hard to analyze and we may not want, if we are doing a confirmatory trial, to use the longitudinal aspects for our confirmatory endpoint.

But, in terms of the kinds of things you are looking at here, the variation over time tells you two things. One, it gets you more data so that just gets more information. But the other thing is it gets you causality. Causes cannot come after effects. It is a very important point.

So the grid, the fineness with which you measure things on a time scale, can make a huge difference. In the Helsinki Heart Trial--for example, compliance was measured and you had side effects measured and they were taking a--I don't even remember what the exact preparation was but it was a comestible type thing, there were a lot of GI side effects of taking it. ***If you look at the data gathered on essentially one-month intervals,

side effects are--and you look at that and compliance, it turns out that the people with the poorest compliance have the highest side effects. But that has got the timing wrong, is the problem. The problem is that the people with high GI side effects stop taking their drug. You can see that if you get the right time spacing.

So longitudinal data can be very valuable but you have got to get the kind of frequency right in order to be able to draw the conclusions that you want to draw.

DR. LESKO: Lewis, when you are talking about causality, are you talking about pharmacological causality in terms of an outcome or something broader than that?

DR. SHEINER: The temporal requirement for causality is very broad. I don't think any theory of causality, except maybe when you get to quantum mechanics and there are some weird things happen there--but, otherwise, if it happened first, it could be a cause. If it happened after, it couldn't be a cause. So that is very powerful for fitting mechanistic models.

DR. JUSKO: It seems to me that this is a very difficult issue to be very conclusive about.

Very typically, the phase II studies yield very rich PK/PD information that is very helpful in establishing basic relationships that we are after, but it is the phase III studies that provide the broader incidence of patient--the greater number of patients studied and the opportunity to identify low incidence of adverse effects.

It is difficult to avoid the present approaches to identify those relationships through any other kind of paradigm.

So I think we can move on to the next topic area basically concluding that we need good rich data and present approaches, at least experimental approaches, are difficult to obviate.

Could we go on to the next question?

[Slide.]

Peter showed some examples of incomplete exposure-response data and is now posing the question of how to model those situations. Comments from the committee?

DR. LALONDE: Just stating the obviously, I guess I find this--I don't know how you can deal with this from a regulatory point of view, to be honest with you. Internally, what we would do is try to look at the previous knowledge have about

the particular therapeutic area, compounds, if it is in the same class, and maybe try to build information to help us make certain types of judgments as we move forward.

But in the regulatory world, where you need to make a recommendation, I am at a loss, to be honest with you, as to how to--I mean, you can come up with methods, but I don't know how you would want to make strong statements about extrapolating above a certain dose range that you have never observed. But I would love to hear other comments.

DR. LEE: We usually don't extrapolate beyond what is observed. But my question is to make use of existing data, which is the incomplete curve, can we model it--for example, one example I show is apparently missing the data of the upper curve. Now, with this incomplete data, how do we make use of the information?

Can we model it? Can we use a polynomial equation or--what would be the recommendations?

DR. SHEINER: No; you can't use a polynomial. It is like Richard says, if you really want to--divorcing it from the regulatory context, divorcing it from the situation that you have to

defend what you do more than most people have to defend what they do. That, I think, is sort of what Richard is saying is it is a big deal.

But you have to make some assumptions. Where you have no data, you have to make some assumptions. That is what extrapolation is about. It says, in one area, that area is connected to the other area, but, in what way is it connected? Does it project off-linearly? Does it project off some other way?

So, for example, where you have that upper bound where you don't know anything more, I would say if you really want to be pretty hard-nosed and make an assumption that most people will buy, all you can assume is monotonicity. All you can assume is that, to the right, as you increase the dose, the toxicity will only get worse. But, whether it will go on a straight line, whether it will go up suddenly, whether it will go flat, you cannot say.

If your conclusions are sensitive to the shape of the curve in that area, then what you have learned is you need those data.

DR. CAPPARELLI: I think it, also, though, stresses some of the points that were brought up earlier about better utilization or more increased

utilization of biomarkers and linking some of those to some of these clinical outcomes because I think you are dealing with low frequencies and it is not just what happens to the curve out there. It is your confidence of those values out there is low.

So you are looking at relationships between biomarkers and with the eventual linking, or trying to validate them into surrogate markers and looking at a more continuous, which I think would be more powerful, scale is of importance.

The other thing is, while you did present that as dose data, you may actually get some additional information if one looks at it from the exposure point of view because you will, within your own dosing, cohorts have variability that do have exposures. But, again, if your endpoint is categorical in that nature, the power to say anything is going to be very limited.

DR. LALONDE: Just a quick follow up. I may be missing part of the point here, but if I recall the example you had, I believe a ketoconazole, or some type of interaction that increased exposure by twenty-fold.

DR. LEE: That is the next question.

That is the next one? I am jumping ahead.

Okay. I thought you were trying to bring those data back in the range of observed ER data that you had. I will just wait, then.

DR. LESKO: Again, going to that same question, I wonder how reasonable it would be to use data from a class of drugs that are fairly well understood and where you might have more complete exposure-response information already available and borrow some of that data in incorporating it into the assessment of an incomplete exposure-response dataset; for example, H2 blockers or something like that where there is fairly well-known pharmacology, the biomarker data is pretty well-understood in terms of its relationship to clinical outcome and the drugs don't differ a heck of a lot in potency.

DR. JUSKO: That seems to be extremely reasonable. Also, it gives you a perspective on the physiological or pharmacological limits of the system. Oftentimes, in those scenarios, you can define the limits of what will happen and that can be used, at least on the Y axis, on one of these graphs to know where you are heading with higher doses.

DR. SHEINER: The beauty of doing that in a Bayesian context is you can add in uncertainty;

that is, you can, okay, this is what we know about another drug but the fact that it is another drug and not statistically this drug means we will widen, essentially our spread on that as we apply it to this drug.

You can actually debate with people how much you ought to do that. At some point, of course, you add in so much uncertainty that you have made it worthless. But, again, you can see the sensitivity. So that is exactly the kind of thing of what are you willing to assume. Those assumptions have to come from science. Those are subject-matter assumptions. They are not based in statistics.

DR. McCLEOD: It is also an area that you can model based on your current data. There are going to be a lot of classes of drugs where they are new or you just can't do that modeling. In oncology, much of that modeling, the data is not going to be solid enough to do because of the differences within a supposed class of drugs whereas your example with the GERD drugs, generally, there is a common physiology that is being measured fairly close to the real thing, to the actual dynamic endpoint that allows you to do

some of that modeling much more appropriately.

DR. JUSKO: Perhaps we could move on to the next question here.

[Slide.]

This question is how to assess the risk and benefit of drug concentrations that are not contained within the known ER relationship. Richard, you were concerned with that ketoconazole example.

DR. LALONDE: I thought it was linked to the previous one, too, in terms of extrapolating the exposure response. But I still think that, again, from a regulatory point of view, this is a very tough one. The part I was missing, I guess, I thought was the ketoconazole interactions are like a twenty-fold increase in exposure, a very large increase in exposure, well above the range that you had studied, and I think you showed the ER relationship, I think, for a certain risk, if I recall.

The part that I am missing, I guess, is that without having other type of information, I think the solution has to be that the dose recommendation for that group, unless you have some other data, has got to be brought in within the

range of exposure that you have studied.

Surely, you are not trying to come up with an exposure-response relationship in that twenty-fold-higher range to show that that is an unimportant drug interaction. Is that intent here?

DR. LEE: In general, the drug is pretty safe. But then it does have this rare adverse event which could be fatal. In this example, the data we have is only up to two times the clinical dose. Of course, drug-drug interaction data we show has up to twenty times the increase of AUC. Of course, for the extreme cases, we don't intend to bring that twenty times down to the normal level. That means you are going to have a dose of 6 milligrams, or whatever, 8 milligrams, which is not possible.

But then, I guess, the question will be how about those with three times the increase of AUC or four times the increase of AUC, which is a little bit greater or beyond the exposure-response data that we have. And then we are not certain whether, when there is a three-times increase of AUC, whether that will cause any clinically significant change in total probability of an adverse event.

So that is the gray zone. How do we make a recommendation in those intermediate areas?

DR. McCLEOD: It seems to really get back to what Lew Sheiner was mentioning about you are not missing the data. You are missing the exposure information to realize you have the data because the variability in AUC is there. It is just that you haven't quantitated it or the quantitation is not available at these given doses.

Just because you only have a two-fold range in dose doesn't mean you have a two-fold range in AUC. So you are kind of taking--I don't know what the right analogy is. It is not an apple-orange analogy. It is a red apple-green apply analogy in trying to say things about all apples.

You have to go down and have information about what seeds you are dealing with. If you haven't modeled in the variability that is possible, you can't draw these conclusions. So, in the context of the phase III studies where you are not going to go back and get exposure information on the adverse events, all you can do is model what variability you would expect to see based on your phase I and phase II studies.

It is not that you are missing--what you are missing is the ability to go from dose to exposure to endpoint. I guess Dr. Sheiner can comment about whether that is ever going to be attainable in the practical sense.

DR. HALE: It seems to me that we have got a choice here between two courses of action. Apparently you know something about the pharmacokinetics in this subpopulation since you know that we are outside of our concentration where we have a relationship. So the question is, we have got a subpopulation. Do you take that subpopulation through a demonstration of effectiveness and/or safety so that we know something in that subpopulation or do you make an assumption?

It seems we have got a choice; either show it or assume it, getting back to what Lewis said earlier. So the question is do we have good science to back up the assumption and, if we don't, we don't have many choices left, do we.

DR. LEE: Or, in this case, it is going to be very difficult to show it because it is a rare adverse event and you need, like, 500 patients or more to show that adverse event in the special

populations. So I guess we have to make some sort of assumption that the dose-response or concentration-response relationship holds true for the special populations.

DR. CAPPARELLI: It is not even that big an assumption because if you are looking at it strictly from the safety standpoint, and you can target within the range, if you are talking a three or fourfold range, your dosing adjustment, more than likely, is not going to bring them even down to the level of the typical value. It is the assumption that they aren't this much more insensitive than the typical population.

In a lot of these situations, I don't think the assumption is a huge one where we can't actually validate it. I think that the is not are these more sensitive issues. It is are they less sensitive and are they less sensitive to actually a pretty large magnitude.

DR. JUSKO: That cardiovascular drug example we have been discussing is particularly fraught with concerns that might have led to a contraindication because a couple of these drugs that cause the marked change in AUC are also on Ray Woosley's list of drugs that change QT intervals.

So you probably have a double interaction there, a kinetic one and changing metabolism as well as a possible dynamic one and both agents having the possibility of changing QT intervals.

But, in any case, it is a difficult situation to resolve and it certainly would require a marked cautionary note if not the need for more explicit studies in lower doses.

DR. LALONDE: I have got to come out and say this. I am not sure I understand the controversy here. If there is no drug interaction, would you allow someone to propose in their label to give twenty times the dose and, if not, I would say even as just a pure contraindication to this combination, then we don't have the data to support this and it is up to the sponsor to provide this not to the Agency to try to create this.

DR. LESKO: I agree there isn't much controversy here. This would be a drug that would be handled through labeling. It is not a labelable situation in terms of a dosing adjustment. I don't know what the real example was, or what the real label says, but my guess is this would be a contraindication for these drugs to be given together.

But let's step back a minute and let's say it wasn't quite 2000 percent. Let's say it was more like 100 percent or 50 percent, something that goes above the plasma levels that you know are associated with an approved dose. Maybe in the absence of other information, you just do a proportional dose reduction and leave it at that.

Whether you need to do that or not, or whether that is necessary, is another question. What if a 50 milligram strength is the only strength available. The question becomes relevant because if the special population has a blood level that requires a downward dose adjustment based on exposure alone to 20 milligrams, how do you handle that situation.

So I think there are other examples where this issue comes into play in terms of extrapolating beyond what you know to have some more data to input into that decision. This one is a little bit at the extreme, but there are others that are less extreme. That is kind of where the difficult comes in.

DR. JUSKO: Perhaps we can move to the last question.

[Slide.]

This one is how to establish consistent criteria for determining the no-effect boundaries for change in pharmacokinetics for dosing adjustment.

DR. SHEINER: You can't do it without utilities, either implicit or explicit.

DR. LALONDE: Since we have talked about utilities quite a bit, I am curious as to what the experience has been around the table with that concept, maybe especially within the agency. Very briefly, we have looked at this for some compounds. Depending who we talk to within Pfizer--we talk to some very quantitative people and they say, "Oh; this is very interesting. Let's incorporate this. Let's see how we can use utility to make decisions.

To the other extreme of, "What planet are you coming from to think that you can incorporate all this complex information into a simple utility function?" That would be, let's say, the typical clinical perspective to say kind of I know what is useful for the patient because I know and I make those judgments all the time."

But it is almost like the opposite of the definition of the judge who couldn't define pornography, I guess; "I know it when I see it but

I can't put it on paper."

So we have had this very wide range of responses and we are still trying to be as quantitative as we can. A lot of a colleagues within the Agency who would have a key role in making these dose would be your clinical colleagues. I am just curious as to, as you advance this concept of utility, as Lewis and others have mentioned that this is the way you need to.

We are making these judgments right now but people are not coming out and stating their assumptions explicitly. I am curious as to how this is being received with the rest of your colleagues in trying to advance these concepts.

DR. JUSKO: I would like to intervene at this point and ask you to restate that question immediately after Jrgen presents his topic that is scheduled at this time.

The program calls for a presentation on using exposure-response relationships to define therapeutic index, a preliminary approach based on utility function. So we can all learn a little bit more about what utility functions are all about and then discuss them further.

Using Exposure-Response Relationships to Define Therapeutic Index: a Preliminary Approach Based on Utility Function

DR. VENITZ: I would like to get started by saying that, Lew, you have stolen most of my thunder already and not coincidently because, for those of you who did get the background, I did include an article that he coauthored twenty-five years ago that actually looked at the use of utility functions. This was the only article that came up when I did a MedLine search on risk and utility.

[Slide.]

So what I want to talk about today is actually how to use utility in the big picture of risk assessment.

[Slide.]

You all are clinical pharmacologists so you are familiar with the world that we live in where we are looking at dosing regimens and we are trying to optimize clinical outcomes by reducing the bad outcomes, toxicity or harm, and by increasing the likelihood of good outcomes, efficacy or benefit. We have also variability that we have already talked about today that relate

dosing regimens to exposure, things like compliance, kinetics, exposure to response, dynamic variability and then the relationship between those biomarkers or response markers and clinical outcomes.

[Slide.]

The context that I started working on this had to do with the definition of narrow therapeutic-index drugs. So how can we come up with the framework that allows us to assess whether a drug or a compound, or product, I should say, is a narrow-therapeutic-index drug.

The analogy that Rich gave is the most common definition; "Well, I know it when I see it." So there wasn't really any kind of framework. There are some definitions, or at least tables, listed in FDA guidance but they are relatively outdated.

So this is looking at a dose-response curve. Now, with this paradigm of kinetics, dynamics and clinical outcomes, you are looking at dose-response curves. Blue is the efficacy dose-response curve. Red is the toxicity dose-response curve. You are looking here at clinical outcomes, so you are looking on the Y axis

at the percent of the people or the patients receiving the drug that show those outcomes.

You can see that this is nothing but a cumulative-density function, a probability function. Typically, one of the definitions that you find in the literature for narrow-therapeutic-index drugs is, well, we are going to see how far those two curves are apart, so we are going to look at the ED50. For example, in this case, the ED50, I think, is 60. We compare that to the TD, the toxic dose, where 50 percent of the patients show us toxic effects. In this case, that number would be 120.

So this would be an example where the two curves are very close together.

[Slide.]

What my contention is, and that is not, really, what, in most people's mind makes a drug a narrow-therapeutic-index drug, but it is much rather what happens if you are over- or under-dose; in other words, what are the consequences of toxicity or efficacy.

So my personal definition is the fact that a drug is a narrow-therapeutic-index drug or not is primarily determined by the severity of the

toxicity of the severity or the lack of efficacy, so what happens when you underdose. The example that I like to use for that is warfarin. I think it goes back to, Lew, you mentioned in your presentation that negative consequence and positive outcomes kind of outweigh.

Warfarin, either you bleed to death or you stroke to death. Either way, by underdose or overdose, you get a very bad clinical outcome.

Something to consider that I don't think we have talked about a whole lot is it really depends on how we dose those drugs. A lot of those narrow-therapeutic-index drugs are not really given as fixed doses. But we individual them or, most of the time, we actually dose-hydrate them.

The most commonly used definitions, I have listed them here. Look at the separation of the dose-response curve or the effect-concentration relationship.

[Slide.]

What I would like to add to that is this concept of utility function that you have heard about all morning long. Here I am saying that the utility value that you achieve depends on the likelihood of having efficacy or toxicity

multiplied by a utility factor.

So the utility factor, or cost function if that is the term that you find in the literature, describes our preference of lack of preference for a certain outcome. For example, clinical efficacy, then, would be defined as how likely is it that the drug is efficacious for a certain dose, so it depends on the dose on the exposure response, and what are the consequences.

In this case, the negative consequences would be a drug that is subtherapeutic. A positive consequence would be the drug actually has the efficacy that it is supposed to have.

On the other hand, if you look at clinical toxicity, you would look at how likely is it that you have toxicity occurred and what are the negative consequences; how bad is the toxicity that you get.

Then you can look at the therapeutic index, the term that is part of the NTI, as a composite of the two. For example, what I am using for a simulation I am going to show is the difference, the mathematical difference. So this therapeutic index, then, follows an exposure response because both toxicity and funicular*

toxicity follow an exposure response and it is affected by our assigned utility values.

As you have heard before, those utility factors are not empirical values that you can do studies for, but they are judgmental entities, things that we assign based on our personal preferences.

[Slide.]

So this is a simple model just to illustrate the point. Now we are stepping back and kind of trying to put that into play. So here I am setting up a pharmacokinetic dynamic model that blends outcomes to dose regimens. I have sources of variability--so we are looking at the different sources of variability. We have variability in terms of compliance, that the dosing regimen actually gets translated into an actual dosing regimen as opposed to the nomina.

You have got pharmacokinetic variability in terms of clearance if you are assuming that it is steady state. And then I have a pharmacodynamic that just says I am trying to get into a therapeutic range, and that therapeutic range is defined by effective concentration and the toxic concentration. Both of them can introduce

variability from patient to patient.

I am looking, then, at the outcomes, the lack of efficacy and the adverse events as the two negative outcomes. So, in the scenario that I am going to walk you through now, I am going to look at dose-dependency studies, administration every 24 hours. I assign certain clearance values and those would population means, and this would be the population mean therapeutic range.

I am simulating here what most people would consider to be a narrow-therapeutic-index drug because there is a twofold range between the effective and the toxic concentration. Then I can add variability in each of those components; compliance, kinetics and dynamics.

[Slide.]

This would be the result of a Monte Carlo simulation where I am looking at dose-response curve on the top and I look at the therapeutic utility curve on the bottom. You have already seen this therapeutic and the dose-response curve for efficacy and for toxicity.

On the bottom here, this is the utility curve for efficacy and this is the utility curve for toxicity. You can see I am assigning a 1,

meaning a maximum positive utility for efficacy and a negative 1, that means maximum negative utility to my toxicity group.

The composite of the two, what I am referring to as a therapeutic index is the mathematical difference between the two and you see, now, it is this kind of a curve, in green here. It has a U-shape and you can tell based on what Lew Shiner mentioned early on, there is a dose right here where you are maximizing utility.

So, if you give this dose, you are optimizing utility relative to toxicity and efficacy.

[Slide.]

If you look at this same scenario now, and we are looking at a case, an ideal case which obviously doesn't exist where we have no variability at all. So here we have no compliance issues. We have no kinetic and no dynamic variability. What you get are those two dose-response curves. They are basically step functions.

More important, if you look at the utility curve, the utility curve now tells you there is a range from 60 to 120 milligrams where you get 100

percent. You will get your maximum utility in every patient. As soon as you are outside that range, you have zero utility. That means your clinical efficacy is completely offset by toxicity.

[Slide.]

You start introducing variability. The first source of variability now is the 20 percent COV variability introduced to compliance. All of a sudden, you see that dose step function, the dose-response curves, get spread out. You can also see that now the utility function gets spread out as well and you don't get 100 percent utility anymore. You are now even at the optimal dose, here around 90, you don't get 100 percent utility.

So some patients, even at that optimal dose, have more clinical toxicity than they have efficacy.

[Slide.]

If you introduce kinetic variability, only. Here we have only kinetic variability, none of the other sources contributed. Again, you can see the spreading out of the dose-response curve, this kind of inverse U-shape looking utilization curve that tells you there is a maximum utility.

[Slide.]

The same thing happens if the only source of variability is dynamics. So, now, 20 percent COV in my effective and toxic concentration. Again, you see the inverse U and you see the spreading out.

[Slide.]

If you put all of this together, you end up with the dose-response curves that you have seen before. So this is what you have already seen before. Now, what I want to change, because that is really what the main gist of my presentation is, I want to change utility factors.

In other words, the dose-response curves do not change. From now on, we have the same dose-response curve that you have seen at the very beginning. If you assume that this is, or at least my definition of, a narrow-therapeutic-index drug where it is very good to have efficacy and very bad to have toxicity.

Then what you would see is the utility curve that looks like this; inverse U. There is a range of maybe 30 to 230 or something like that where you would have a positive utility. You have your maximum utility value at around 90 milligrams dose.

Now, for the same dose-response curve, now, I am deciding that my utility values are different. I have a drug that has a marginal therapeutic benefit, so 0.2 out of 2.0. So it is one-fifth less important for me to have clinical efficacy. At the same time, I am concerned about toxicity because I am assigning it a negative 0.8. So I think there are pretty bad potential outcomes as far as toxicity.

What you get, then, is, again, if you look at the green curve, you now see a very narrow therapeutic range, a very narrow range of doses where you have positive utility. You can also see that even at the optimal dose, still around 90, your maximum utility that you get is very small. So this would be a marginal efficacious drug with significant toxicity and you probably wouldn't want for this drug to come to the market in the first place because it provides very marginal efficacy given the fact that it has such significant toxicity. Even dose optimization is not going to help you.

On the other hand, if you look at this drug, this would be a drug that has significant efficacy. I am assigning a large utility value to

it. On the other hand, the toxicity, the consequences of toxicity, is relatively insignificant, negative 0.2. Same dose-response curve. Now, look at the utility curve. Now the utility curve goes up. It peaks at around 90 to 100 and then it remains positive for a large dose range.

So this would be a drug, even though the dose-response curves are twofold separated--so it would meet the conventional definition of narrow-therapeutic-index drugs, if you look at the utility, there is a wide range of doses where you would have a large degree of utility. So a lot of patients would benefit regardless of where you are on this dose response.

[Slide.]

As you know, I am on sabbatical with FDA and this is the project that I am working on, just to give you some idea where this is going to lead to before I am going to ask you for some additional input. Right now, I am looking at additional simulations where I separate the variability into different subpopulations, something that I am really excited about. It would be the second direction and I have some stuff, and I have done

some stuff--it is not ready for prime-time yet--but to look at strategies to deal with narrow-therapeutic-index drugs, things like dose titration.

Can I deal with the fact that I have a source of variability by using dose titration either on a kinetic endpoint like a plasma level, or some surrogate markers. And then, down the road, potentially look at more complex PK/PD models even though I am not sure how much they contribute for the proof of concept.

Something that I do look for guidance from you; are there any ways that I can get actually to real-life data that allow me to show in a real-life example how this would work.

[Slide.]

Now, the discussion that I think--Rich, you asked that question about utility, how do you come up with utility factors. Let me give you some general ideas that I think we might want to consider, maybe come up with utility factors. So utility factors describe our perception of what the consequences are of either not being efficacious or being toxic.

The first thing to consider; can we

actually monitor clinical outcomes, or is the first clinical outcome a dead patient? If you can monitor, then the utility function would potentially be less, or the utility factor, I should say, would be potentially less, or can the patient diagnose that there is some clinical outcome.

Can the physician diagnose it or is there a special testing that is required? At what setting does the outcome occur; self-treatment by the patient, outpatient, or does the patient have to be hospitalized if something bad happens either lack of efficacy or toxicity.

Specifically, to the efficacy, what kind of utility considerations would we have when we try to assign efficacy utility values? What is the impact of the treatment, the drug, itself, on the disease? Are we preventing the condition? Are we relieving symptoms only, or do we cure disease--that would tell us how important it is to have clinical benefit.

What is the severity of the disease? And are there any alternative treatments available and how would they compare to the treatment of interest?

On the other hand, if you look at toxicity, or the harm that you can cause, is that reversible harm or is this something like patient death? And what is the impact of this toxicity on the quality of life or the activities of daily living?

[Slide.]

What I want to conclude with, and the reason why I think we had this discussion early on, that using utility functions, you are actually combining clinical pharmacology-type information, exposure response, that we can reduce, as Peter is proposing, to probability-density functions, basically, for efficacy and toxicity.

We are combining them with therapeutic judgment. The therapeutic judgment is implemented by assigning utility values in order for us to come up with a therapeutic index. I believe that that is going to be useful for us to come up with a consensus of how to define narrow-therapeutic-index drugs, and the narrow therapeutic index, in general for other drugs as well.

[Slide.]

So the question I have for you as a committee, in terms of feedback, what do you think

of this general approach, what specific modifications or additions do you suggest, what would be an approach to come up with a consensus on those utility factors, the very question that you asked, and what are specific classes of drugs that I ought to look at a little more closely.

Thank you.

DR. JUSKO: Maybe I could begin with a question. What is the typical range of utility factors? You used negative 1.0 to positive 1.0

DR. VENITZ: It is arbitrary. I have just defined, for the purposes of my presentation, that positive 1.0 would be the best possible consequence that I can have. I am saving somebody's life. Negative 1.0 would be the worst possible outcome. I am killing somebody. It is arbitrary. You can assign any range that you want.

So, for the definition, the way I have defined it is it ranges from negative 0.1 to positive 0.1. But you could assign any value that you would like.

DR. LALONDE: What is important, I think, is the relative weight.

DR. VENITZ: Exactly.

DR. LALONDE: The relative weight that you

put on these. Are they equal, as you said, in your example or are they not equal.

DR. VENITZ: In one of the examples, they are equal. And that is the point that I was--if you look at this, here I am assigning equal weight, toxicity and clinical efficacy. What you get is a utility curve that looks like this which would suggest there is a range of about 30 to maybe 230, we have a positive utility.

On the other hand, with the same dose-response curve, if I now say I have marginal efficacy--in other words, my efficacy really is not very important clinically speaking, I still have very important clinical, or clinically significant toxicity that, all of a sudden, my utility curve is much smaller.

So you see the change from here to there just by assigning different utility values. But it is arbitrary judgmental way of looking at the consequence, the positive or negative consequences of over or underdosing.

DR. SHEINER: Let me just clarify. The scale is absolutely arbitrary and no computations come out different when you change the scale and variant. The last thing it says is an arbitrary

way of assigning clinical value could be heard as that utilities are arbitrary. They are subjective, but I wouldn't say they are arbitrary.

DR. VENITZ: The numbers that you assign are arbitrary. The values that they reflect are not arbitrary. They are judgmental values on looking at benefit and harm.

DR. JUSKO: The relationship between efficacy and the utility factor, or toxicity and utility factor, is it typically a linear function or it can be any type of arbitrary function.

DR. VENITZ: It can be any function. What I am assuming here, it is just a factor. I am just multiplying the likelihood of having clinical efficacy by some factor that tells me how good is it for me to have this kind of efficacy. Here I would say it is very good. I am saving lives. That is my clinical efficacy utility. Here maybe I am treating hay fever and I am preventing somebody from sneezing.

DR. SHEINER: No; I think that Bill was getting at a different point and it is an important point. If you defined all of your outcomes as categorical, so there were three levels of efficacy and there were two levels of toxicity and so on,

and you had lots of them. Then, for every unique combination, in principle, you would have to assign a utility and that would be what is called a saturated model and nobody could argue with it because you get to assign utilities any way you like.

But if you, for example, talk about blood pressure which is continuous, and you talk about some insomnia which is continuous, then you need some model for combining those separate utilities. Do they interact or do they not interact? Certainly, for multiple toxicities, you can imagine total degree of discomfort is greater than the amount you might assign for one toxicity and another if you have both at once, if you are both nauseated and vomiting, that is worse than either--well, I would say that. But it may not be any worse than vomiting alone.

So you have the same problem in modeling that you have in modeling anything. As soon as they become continuous, do you want to combine them or your endocombinator*ics blow up. I didn't mention, and obviously this is one of the problems, and you didn't mention it either so we ought to state it out here is that that is much tougher to

model because we don't have the same kind of empirical data. In principle, utilities vary from person to person.

DR. VENITZ: You look at them as personal preferences of outcomes and they could be different between you and me. They could be different between you and your patient. So they are subjective. But the numbers that you assign are arbitrary because, as Lew pointed out, there is scale and variant. You are looking at relative changes.

DR. LEE: Dr. Sheiner, you mentioned that you model a utility function. So, when you model it, what would be the required data that you model?

DR. SHEINER: Assuming that you are willing to make the assumption that everybody's utilities are about the same, so you would have to dealing with big things. Most people would feel the same about it. But that is a tough assumption which is not an assumption about the natural world. We really do assume that the natural world doesn't change as we move from place to place and from time to time.

But preferences do. If we are willing to assume that everybody is basically the same, then

the way you elicit utilities is you have a dialogue with people in which you say--there is a whole literature on this--but in which you essentially say, what is your equilibrium point. If you had to walk with a limp for the next ten years, would that be about equal to living five years longer, or whatever the number is.

They have spent a lot of time figuring out how to elicit utilities from people. So the experiment you do is conversations with people in which you pose them hypothetical situations and essentially you get them to talk about things that are even odds, and that is how you get your weightings. When they are indifferent about two things, then you say they have the same utility.

So it requires interviews. Probably we would take the paternalistic point of view that we would start out eliciting utilities from doctors, not patients, and so we would have to interview health-givers.

DR. HALE: I think there are some things we could probably learn from our pharmacoeconomics people. They have been doing this sort of thing for years. They typically look at Regimen A versus Regimen B rather than having an underlying

continuous input such as dose or exposure in terms of pharmacokinetics.

But it is a methodology that has been around for years in that arena for sure. They often look at things like length of stay in hospital, quality of life, et cetera.

DR. VENITZ: I looked at some of the literature. Most of the time, their utility function is cost; in other words, they are looking at dollars which are pretty unambiguous to actually empirically come up with. It is much more difficult to come up with utility values that look at preferences, as Lew pointed out, because they vary from doctor to doctor, they vary from doctor to patient.

DR. HALE: The thing about utility is that you have a common scale, that everything basically translates, whether it is quality of life, medical outcome, dollars. Basically everything goes through a utility function and put on a common scale. There are these things called multi-attribute utility functions where you have lots of inputs or dimensionalities to worry about.

DR. LESKO: I have to come back to a regulatory-world reality. Approving drugs is a

benefit-risk assessment. There are always efficacy questions. There are always safety questions. At some point in time, utilities are probably unconsciously being thought about in making the benefit-risk assessment.

The next step is to say, now I am going to put a number on this. That makes people very nervous. As a prior step, one would have to figure out a process, even just agreeing on a process by which one could establish utility values. It seems to me, at best, one could establish relative value. I am speaking of this in the context of Drug X and what it might cause on the harm side versus Drug B and what it might cause on the harm side as opposed to absolute values.

Whenever I hear the variability across medical or the variability across physicians, it just reminds me of how difficult this could be to establish in the context of regulatory decision-making. I am trying to look for advice on a way forward in that sense.

DR. SHEINER: Again, you don't want to make the best be the enemy of the good. You have got a nice example here in the sense that it is a relatively limited question. It is not, what do we

do for the next thirty years in this country. It is, what dose of this drug are we to give for this indication.

Let's even get away from the issue of it might be different for every patient because we can't do that. So we could, then, begin to talk about cost because it becomes a societal kind of a thing. We don't necessarily have to start comparing it to other drugs because that is not generally what the FDA sees itself as doing, as approving something that is better than anything else out there. It is just, does the balance here--and, as I say, in the beginning, we can start with very few effects. Jrgen used just one efficacy and one toxicity. We can start there.

I think just starting down this path with the simplest kinds of things will take us to some very useful places. We will start getting explicit about things we never got explicit about before. But I really like it for the dosing thing because this is a containable problem. It doesn't suddenly start to have tentacles going out into everywhere and we have to decide what the next ten years are going to look like in the politics of Iraq or something.

DR. VENITZ: I am going to just add to that. I have been with FDA now on sabbatical for the past three months and I have attended briefings. You have heard Peter talk about how difficult it is sometimes to assess the impact that changes have in area under the curve, let's say. Usually, there is an implicit utility value that clinical pharmacology reviewers and medical reviewers use to decide whether 50 percent or 75 percent change in area under the curve is relevant, meaning is it a precaution, is it a warning or is it a dose adjustment.

There is a utility value already being used. We just don't call it that way. So we can't really argue. So, all of a sudden, you have two people disagreeing. This person says, well, 20 percent is important. The other person says it is not important.

What they are really not arguing about is the extent of change but what the potential negative consequences are, usually. So this is just an explicit way of putting that on the table so we can have a discussion on it. We might not agree on the utility values but at least I know why Rich and I don't necessarily agree on the

particular scenario.

DR. LALONDE: I completely agree. When we try to sell these types of concepts to colleagues who are skeptical, we say, well, these judgments are being made right now. The difference is that you are not stating your assumptions. You are just basically leaving them up here and saying, "I am saying that we can't use the dose, or this is not clinically important or this is very clinically important."

What we want to do is, basically, with models that you can state your assumptions. You put them on the table and then you debate the assumptions. I think this is what these weights and factors are really all about.

In response to a comment that was made earlier, what we have tried to do is include several people into that assessment so that you don't talk to one expert but maybe have a collection of so-called experts, go around the room and say what is the average figure that you would come up with after the people have a chance to just say their preference.

I would like to come back--again, I think this is very similar to the kinds of discussions

that we have had here and internally where I work. I am just curious as to Larry and Jrgen how the people who are less familiar with this type of approach, who may be very familiar with making these judgments but don't think of it in a quantitative way in terms of the utility function, how far do you see this going in the next six months, twelve months? Is this something that is going to take ten years to move forward? Is there mainly skepticism, because these medical reviewers are the ones who are at the heart of some of these decisions also.

DR. LESKO: Maybe that is an answer we need to save for another advisory committee meeting. I would say we haven't tested the waters there. I don't think there is an answer. As everyone realizes, we make these decisions all the time and that is how labels get out there.

We were approaching this, and are approaching it, from the standpoint of bringing more systematic ways of doing that in order to both improve the labeling of the product as well as to bring consistency to the interpretation of these changes.

This is one of the approaches that is out

there. I think we need to advance it further and then ask the question about how do other people react to it. In fact, I would like to see us advance it with a specific drug and some specific examples to show how this would work. Conceptually speaking, these are hard concepts to advance within the agency, in my opinion.

But, with some examples in model drugs, I think it would be much easier. My sense is, in the overall framework of risk assessment, because of the priority this has been sort of elevated to in the Center, I think people will want to look at this. But it has to be presented in the right way.

DR. SHEINER: As I said, doing it in the general case is very, very tough. But there are very straightforward examples. One of the examples I was going to show is not the oxybutynin, which is a complex one, but just a recent study we did on use of magnesium infusions in preeclamptic hypertension.

We were able to get a PD model with the level of magnesium associated with blood-pressure fall and everybody agrees that you don't want to go above 4 because you start getting seizures and ugly things like that in terms of a level. We didn't

get any toxicity data.

Then you fit the population model for the variability in response and the variability in PK, simulate out the patients under various dosage regimens and you get to find out that there is reasonable expectation that the currently used dosage regimen has a problem in that it gets to where you want to go but too slowly and that you ought to regularly have a loading dose which has been used by many people.

Sometimes, they give it IM. That has got its problems. But the point is it is a simple analysis that says here is a regimen that somebody ought to try and it might be better. That is where you go from there. Now, to approve that on a label is quite a different thing than saying in the course of drug development, "Oh; we ought to try this and use that in our phase III study and maybe try some variance to show that it makes a difference."

It is that kind of encouragement, if they knew that they had to do that kind of justification of the dose they offered at the end, maybe, at the time when you can do smaller experiments and get richer data, you would start to get to see what we

would have.

But I really feel strongly that we are not at the point now where we are ready to say, "This is how you do it."

DR. LESKO: That was kind of my reaction, to pick let's say a negative utility value for something everyone agrees is bad. You can start out with the QTc, for example, as a bad thing. Everybody is concerned about it. It is probably one of the bad things we have some continuous dose-response data for some drugs--and take a look at that. That would be where you would expect the easy case to be made, and then maybe go into some of the more complex.

But having the prototypes would help, I think.

DR. HALE: I think there is a lot of merit to this whole notion. I think, basically, what you are talking about is quantifying our benefit-risk as a function of exposure. I think there is a lot of benefit there, but I think you need to think a little further about the side effects, what are some of the knockons here. For example, this could wind up that when we have a label, we basically have somewhere hidden in

there--if it is not in the label, somewhere behind the scenes, a number which we have quantified as benefit-risk.

In terms of pharmaceutical companies marketing Drug A versus Drug B versus Drug C, they are each going to have this cost-benefit number lurking in the background and that is going to be tied directly to the kind of recommended dose that is allowed.

In other words, everybody is going to be in this game of optimality, what is the dose that gives us that best numeric value which is going to put a lot of pressure on getting your utilities sorted out. I think that is a significant thing that is going to have to be given quite a lot of thought and make sure that all the constituencies impact to get input into the development of those utility functions.

DR. DERENDORF: I think, conceptually, the approach makes a lot of sense. But I think the difficulties are really in the details. For example, it all depends on the PK/PD models that are built into this model. You need two. You need one for the efficacy and one for the safety. There are not that many examples out there that really

have looked at safety PK/PD modeling.

Right now, we are having an effective concentration and a toxic concentration. That is nice and simple, but I don't think it really reflects the real world frequently. So I think there is the challenge because, if the models are wrong, the conclusions will be wrong.

DR. VENITZ: I agree with that wholeheartedly. What I have seen, again, in my limited experience, most of those safety models are empiric. You have seen some of the examples in Peter's presentation. Most of them, you believe you are only at the low end of the dose-response curve because ethically you can't push the dose any higher.

So you are talking about, most of the time, low-probability events. They happen in less than 1 percent of the population even at the highest dose. But they have potentially a very high negative utility.

Those are the ones that are ultimately going to drive you over a therapeutic index; right?

DR. SHEINER: I just say, again, what is the competition. The beauty of talking here is you guys have to make decisions. You have to make them

and you have to make them relatively promptly. So anything that might be a modest improvement, even if it doesn't get all the parts right--but this idea of unintended consequences that Michael is reminding us of is, I think, a very important issue. It happens all the time.

There are probably things we can do about that, but I think that is another reason for testing it out and trying it slowly and seeing where it takes us.

DR. JUSKO: It sounds like there is considerable consensus that this would be a very valuable approach to pursue further looking for more specific examples to apply the methodology to in order to demonstrate the attractiveness of this nice blend of being able to utilize the art and science of what we do.

I think we have concluded our discussion of this topic. Any other comments from the committee regarding this or any other aspects of what we discussed this morning?

DR. LESKO: May I ask just one clarifying question related to the utility function? Dr. Venitz showed us how this can change under different scenarios of variability and I was trying

to, then, leap from there to the need to dose adjust.

Clearly, these utility curves have a peak and a flatness to them or a steepness to them as they go up and down, and I assume that, if the plateau is rather flat or the rise is rather flat, that would kind of suggest that even large changes in exposure would not necessarily require dosing adjustments based on this net utility whereas, if the curve went out and down, as you showed us, that would be a case for a more urgent situation.

If that is the case, it may be worth looking at decisions that have been made on that type of exposure change already and see if there is some consistency in what is currently being done and what is being proposed, and these differences may shed some light on what we should be thinking about in the utility-function area.

But, am I interpreting that correctly?

DR. SHEINER: You have got to watch out for individual versus population. So let's imagine a drug which has essentially no relationship between dose and exposure. You give a dose and you might get any exposure. No such thing exists, but let's just imagine it.

But the exposure-response relationship is reproducible, and so is the dose-exposure relationship, within any individual. What you would see in a dose response, under any utility function, virtually, is it is totally flat because the dose can give rise to any exposure and exposure can give rise to toxicity or efficacy depending on what it is.

And let's say it was one of these things where it was 0.8 and 0.2 for efficacy and toxicity, so it would be positive utility. So you can give any dose you like. You are going to get, on the average, 0.6 or whatever it is. But the reality is that, for some people, they are getting toxic when they don't need to and, for other people, they are failing to get efficacious when they don't need to.

So you have to build in, when you are thinking about these things, what other information you might get; for example, the initial response of the individual or some other test that tells you whether they are going to have this kinetics or that kinetics and so on.

So just going across the population and mixing everybody together, what it does is it gets you a legitimate curve, but it is a kind of a

flattened utility curve because all this variability is mixing in all kinds of folks. So you have to say, what are we talking about? Are we talking about dosing people when we don't know anything about them? Or are we talking about dose people when we know something about them.

You can see, actually, how the special population comes in. You will see that, suddenly, putting in the information that somebody is in a special population changes the utility function for everybody because you have broken them up into groups that have less variability.

DR. VENITZ: But, just to add to that, one of the limitations I didn't point out that the concept of utility functions does, you are trading off probably against utility. So you are saying one person dead out of 10,000 is the same is 10,000 people having a slight headache. You have the same utility value, so you are trading off. You are just doing it explicitly as opposed to right now we are kind of doing it intuitively.

DR. LALONDE: Maybe just a very small last comment is also when I tried to look in the literature, I saw how little information there was in the clinical-pharmacology world so a plea for

people who are doing research in this area to publish their information so that they maybe get at least more in the public domain and people to respond to this with other papers, commentaries, whatever. But there is very little of it, at least in our discipline, that has been published.

DR. JUSKO: That brings up the possibility of a proposal. It seems like, as we went through the discussion of the main topic, the flow chart and all of the specific questions, everything seemed to be too complicated to have any easy answers. What we have come up with is a lot of suggestions of needing to explore these issues further and also the great desire to have many more specific examples to go by to explore what other people have done with more specifics.

So it seems like this would be a very good topic for exploration at a meeting, to have presenters deal with many of these issues and to discuss it more widely. It certainly is one that you will need to develop much more thoroughly as what we have ascertained from our limited discussion of all of this.

DR. HALE: Just a suggestion here, and that is, while this is relatively untested in the

clinical-pharmacology arena, the federal government does have a lot of experience already looking at utility functions in various applications such as the space program, nuclear reactors, et cetera.

So it seems to me that we need to find some appropriate expertise, people with the utility-theory background, to really pursue this. The other is the recommendation to really give some thought to criteria other than just expected utility.

I think one of the graphs you showed on Page 12 actually goes to that, and back to the question that I asked Lewis earlier, because when you pointed out the graph on Page 12, you said this is probably one you wouldn't want to do even though the expected utility approach would tell you to go ahead and administer that dose.

I think, logically, we can all look at it and see that that is probably not a very good idea.

DR. VENITZ: That gets into the issue of how you scale. In other words, is a 0.5 or whatever you come up with, or 0.1, I guess, expected utility at best, is it worthwhile in the big picture. So it really comes down how do you assign utility values? Do you consider other

treatments that are out there?

DR. HALE: That kind of begs the question. In this case, you are saying you didn't get the utilities assigned correctly. I will come back to you; suppose you did get them assigned correctly. Are you going to go ahead and do this even though all of us look at this--I am supposing most of us would say, "That isn't really a very good idea, is it?"

DR. SHEINER: You can't escape that way. The utility, already, in principle, has all the values in it so you can't say, well, a utility of +0.1 isn't worth very much. No; it is worth exactly +0.1 and, if it is positive, it means you ought to do it. If you are not going to do it, then it means you need a more complex analysis of some kind.

But your intuition is good. Pay attention to your intuition. Don't say, oh, well, I guess it says 0.1. I guess my intuition must be wrong. If it doesn't look right, then there is probably more likely something wrong with the way you put the problem together than there is that you are wrong.

DR. JUSKO: Are there any other comments, anything anybody wants to bring up from the

committee members or people from the FDA?

DR. SHEINER: I just wanted to say one thing. This business of other parts of the government having experience and so on, we have just witnessed in the last several months a complete change in public attitude about the value of estrogen replacement for postmenopausal women based on a perception that there is a risk which is something like 5 or 6 per thousand of a not-necessary lethal event that we finally have tied down.

There has been a whole judgement that country has made based on some utility associated with that sort of a risk. People have asked me that because they know I think about this. I say, "I don't know any way to think about, personally, risks of a few per thousand.' I know, as a society, you can work it out and say, how much is it going to cost me, and so on, so that is sensible. But, as an individual to react to risk--and you look around, and most people don't. We all happily get on airplanes or walk around with a sniper shooting at us, and so on.

It is a level of risk at which we simply don't do anything about it because it just doesn't

make any sense to us. So what I am saying is this pervades all of our decisions already and there is nothing the matter with trying to make it a little more explicit in these daily issues that you have to deal with.

DR. JUSKO: On that point that is relevant to many people going to lunch, we will take our lunch break at this time an we will resume at 1:30 to deal with Topic No. 2.

[Whereupon, at 12:25 p.m., the proceedings were recessed to be resumed at 1:30 p.m.]

A F T E R N O O N P R O C E E D I N G S

[1:35 p.m.]

DR. JUSKO: Welcome to the Clinical Pharmaceutical Subcommittee of the Advisory Committee for Pharmaceutical Sciences. We are going to begin the afternoon session with what is scheduled as Topic No. 2, use of exposure-response relationships in the pediatric study decision tree: questions to be asked using the FDA pediatric database.

We have two presenters from the FDA and then we have some additional commentary that Dr. Lesko may wish to discuss further.

We will begin with Dr. Rosemary Roberts.

Topic No. 2

Use of Exposure-Response Relationships in the Pediatric Study Decision Tree: Questions to be Asked using the FDA Pediatric Database

***

Medical and Clinical Pharmaceutical Perspective on the Pediatric Study Decision Tree and Experience to Date

DR. ROBERTS: Good afternoon.

[Slide.]

I am Rosemary Roberts. I am a pediatrician and a mother, as you might surmise from my opening comment. I have been involved with the pediatric initiatives that have been going in with the Agency since the Pediatric Labeling Rule was published in December of 1994. I want to thank Dr. Lesko and his office for inviting me here to participate and to give a presentation at the first meeting of this subcommittee.

I hope that by the time I finish speaking that you will think that we actually do have a rational approach to drug development in pediatrics.

[Slide.]

As you all know, with the incentive program that was legislated with the FDA Modernization Act that was signed late in 1997, the Agency came out with a guidance as to how industry could qualify for this six months of additional marketing exclusivity. There is no doubt that money talks because industry has been very eager to get their six months of marketing exclusivity to the tune that we have issued, to date, 256 written requests to industry and they have sent in over 300 proposals to us requesting to study a drug in the

pediatric population.

When one of these proposals comes in to a regulatory division, there are some questions that they have to ask themselves. The first question is is there a public-health benefit to studying this drug in the pediatric population.

If there is, then that is the first criteria that was mandated in order for us to issue a written request. If there is a potential health benefit to the pediatric population, then we can issue this written request to get the information.

So now we have a drug for an indication that we need information in pediatrics. In what age groups do we need information in the pediatric population. As you all now, pediatrics is not a homogenous population. We have the prematures, the neonates, the infants, children and adolescents. Those are arbitrary names and arbitrary cutoffs. Sometime, we can't use age groups. We have to use Tanner stages or some other physiologic basis for dividing up the age group.

Be that as it may, there are certain things that have to be considered when we ask what age group. There are some conditions, like infections, that occur throughout the pediatric

population as well as in the adult.

But then there are things that do not occur in the entire pediatric population. For instance, let's take Type 2 diabetes. Traditionally, we have thought of Type 2 diabetes or adult-onset diabetes as occurring in adults somewhere in the fourth of fifth decade of life. So when I saw the first written request for Type 2 diabetes with an oral hypoglycemic agent coming to the Pediatric Implementation Team, I thought, "What are we doing here?"

But, unfortunately, in this country, we are seeing a lot of Type 2 diabetes in adolescents, adolescents that are overweight and don't spend much time exercising, at least not physical aerobic exercise. Maybe they exercise their finger in videogames.

So, indeed, we do have a population in this country that has adult-onset or Type 2 diabetes in the adolescent age group. We are currently--metformin was studied in the ten to sixteen-year-old to get information on how to use it, and we were even entertaining going down to age eight, which is sad, but we are now making the diagnosis in the eight-year-old, even.

So we wouldn't study the entire pediatric population. We would request studies in eight to ten years or above because the condition, we don't recognize it below that. So that is one example of a condition that does not occur throughout the entire pediatric population.

Another reason we might not study the entire pediatric population would be a condition such as depression. Although depression, in some form, may occur in the preschool child, right now our studies are asking for seven and above. The reason is we don't have an approved drug in the pediatric population for depression yet.

Until we get some positive studies in this population, using the criteria to diagnose depression in this age group using the valid scales that we have, using the outcomes we have, we don't know how to take the studies into the preschooler.

We do anticipate that, in the preschooler, we may have to have different outcomes. We are going to have to have different diagnostic criteria. And we may have to have different assessments. Remember, it will be in the preschool age, so they can't do some of the stuff the school-age child can.

So there are just two examples of why we might not study the entire pediatric population.

Once we have decided on the ages to study, then what information do we need? In the divisions, what they do is they clearly know what the product is labeled for. They can go into the file of the manufacturer and they can find out what is available in the file. There may be some studies that have been submitted to the IND but they haven't requested it in the labeling. That may be able to be used.

There may be information in the world's literature and some of that may be strong enough to be able to be used. But ultimately they have to determine what is the information that is missing. So, once we have the information that is missing, then what types of studies do we, as an Agency, request in order to fill that information down.

This is the thought process that goes through. And we have gone through it for the 256 written requests that we have issued to date.

[Slide.]

Just briefly, as of September, we have issued a written request requesting 601 studies. Of these, 35 percent were efficacy-safety. Another

30 percent were PK-safety. Another 9 percent were PK/PD.

I am going to talk to you now as we go into the decision tree where some of these products lie.

[Slide.]

This is this decision tree that is in the guidance that is out, the Exposure Response Guidance. Let me just briefly talk about this. There are two assumptions here. Is it reasonable to assume, between the pediatric and adult populations, that there is a similar disease progression and a similar response to intervention.

Why have we used these as the two assumptions because, many times, we don't have actual evidence. Secondly, the 1994 Labeling Rule that we published introduced the idea of the ability to extrapolate adult efficacy into the pediatric population of the condition was sufficiently similar in the pediatric and adult population and if the response of therapy was expected to be the same.

So that is really the basis of where these come from.

Now, our goal, obviously, is to get to the

point where there aren't assumptions but where we actually have the data to know whether the disease progression is the same and whether the response to intervention is similar.

So, looking at this, if you can answer yes to both of these, then that takes you down this side of the decision tree. Now the next box is, is it reasonable to assume similar concentration response in the pediatrics and adults. The best-case scenario is yes, it is reasonable to assume and, therefore, we can extrapolate adult efficacy. We don't have to reprove efficacy in a child through adequate and well-controlled trials, but we can conduct PK studies to achieve levels similar in the adult so we can get the dose right and we can conduct safety studies in the pediatric population so that we know if there is any unique safety concerns in pediatrics.

Now, the Rule of '94 is very clear. It says, extrapolate adult efficacy because we don't feel you can extrapolate safety. Now we have forty-three products that have been labeled since this initiative started. We have several examples where there have been some safety concerns that have come out through studying the pediatric

population.

Now, I will just give you a couple of examples quickly. For gabapentin, which is an anticonvulsant that is approved now in children down to age three for adjunctive therapy for partial seizures. The labeling now contains, in the warning sections, neuropsychiatric adverse events that were found in the pediatric population three to twelve as a result of the studies. Such things as hostility and aggression are now in the labeling.

If we can say yes to both of these, but it is not reasonable to assume a similar concentration response in the two populations, then we move over here; is there a PD measurement that we can use to predict efficacy. That takes us down to this box here. I will show you on a later slide several examples of where we have actually been able to conduct PK/PD studies and then get an idea of what dose we need to use, conduct the PK studies to a targeted concentration, conduct safety studies and label the product.

I think I will move on so that I can actually show you some examples here.

[Slide.]

Here are some examples of where we have actually defined PD measurements. We have used these measurements. They are in written requests that we have issued to date for various indication and for various drug classes.

Here for HIV and for all the drug classes that we are currently studying in the pediatric population, the pharmacodynamic endpoint that we have used is the assessment of changes in the plasma HIV RNA levels as well as the CD4 cell count. So we don't take and reprove efficacy. We have them study the child and to target to the HIV RNA plasma levels and, thereby then, get the dose that is appropriate for children as well as getting some safety information.

Another example would be gastroesophageal reflux where we look at changes in the intragastric pH. That is for both the H2 receptor blockers as well as the proton-pump inhibitors.

I must say that we have had a change in thinking here with the products for gastroesophageal reflux disease and that is basically in the age group of the infant, one-year-old and less. The clinical manifestations of gastroesophageal reflux are very different than

in the older child or in the adult who experiences more of a heartburn and all the accompanying symptoms of that.

These children have problems, respiratory problems. They have problems with regurgitation and aspiration, apnea, et cetera. So we now have a new template out, and it is up on our website, that indicates that we really need to look at clinical outcomes in this population.

Then we also have, for juvenile rheumatoid arthritis, for the NSAIDs, if we are looking at the signs and symptoms of arthritis and their resolution, we have a guidance out now that says we can actually extrapolate that from the adult. So what we do is, for the pharmacodynamic parameter, we look at clinical responses such as joint evaluation and a SED rate as well as a global evaluation and we have used that now in labeling two NSAIDs to date for juvenile rheumatoid arthritis, etodolac and oxaprozin.

DR. SHEINER: Excuse me. I'm sorry; how do those differ from what you would use for an efficacy endpoint?

DR. ROBERTS: Well, we did not do adequate and well-controlled trials. We didn't reprove they

were efficacious. What we did was we studied, and there were less than 100 patients that were studied for both of these drugs, and we actually had them use a dose to see if you could get the appropriate clinical response as you would in the adult, and look at pharmacokinetics and thereby determine what would be the appropriate dose to get an appropriate response. e response.

But we didn't reprove efficacy all over again. As it turns out, for etodolac, the information we got was that actually they handle the drug differently in the pediatric population and we really need to double the dose in order to get an efficacious dose in the pediatric population.

[Slide.]

Here I have put in some examples of classes of drugs or indications for which we have used this decision tree and we are currently getting information. I would like to point out that the one path I showed you where we get a PD and then we do these PK/PD studies as well as safety, we have used this now for the H2-receptor blockers and proton-pump inhibitors, as I talked to you about, with the caveat that we have changed for

the less-than-one-year-old for the HIV drugs.

We also have a group of drugs for conditions where you have to reprove efficacy in the pediatric population. That would be for the antidepressants and for the antihypertensives, the anticonvulsants and migraines. Why for the antihypertensives? If a drug can treat blood pressure in the adult, why do we not think it will treat blood pressure in the child?

The Cardiorenal Division is concerned, and we are assuming now that it won't work the same as in adults because the etiology of hypertension in the child is very different from the typical etiology in the adult. So, until we get some experience in the various classes of antihypertensives to show that, indeed, if you treat blood pressure in the adult, you are going to be able to treat blood pressure in the child, even though they have very different etiologies, we are asking for efficacy studies.

So, hopefully, down the line when we have got some of these products well studied and labeled, we will be able to not have to worry about assuming that the response to intervention is going to be the same.

The same with the anticonvulsants.

for the last part of my talk, I am going to talk about a condition and some of the factors that you need to consider as you approach using this particular decision tree. This decision tree is a way to start thinking about how to develop drugs in the pediatric population. It is not going to address every situation.

As a matter of fact, this particular group of drugs that I am going to talk about right now, the asthma drugs, they don't fit on this. Arzu pointed that out to me. She says, "You haven't got that coming off the right box." I said, "There is really no box to have this come off from here."

But I want to use this as a case in point.

[Slide.]

Okay; asthma. This is a condition of reactive airways and inflammation. We do know that the progression in the pediatric population really is the same as in the adult in the sense that it is airways that are reactive leading to bronchoconstriction, leading to a lot of mucous formation and going on to a full-fledged asthma attack in the child as well as in the adult.

So if you look back at that tree, which

you should have in your handout, we do know that the progression is the same. The question is is the response to therapy going to be the same. For beta-adrenergic agonists, or bronchodilators, we know that the response to therapy is going to be the same.

Therefore, we should be able to follow--let's go back here--we should be able to say yes to both. It is reasonable to assume a similar concentration response in pediatrics and adults.

You know for many drugs that work as a bronchodilator, if you think of aminophylline, which isn't really used a lot now, fortunately, because it has a lot of side effects people don't like, but we used to actually look at target dose levels because we knew what dose level usually gave an effect and we also knew what dose levels caused side effects.

So we should be able to go down here and conduct PK studies and safety studies. And yet, I have put these people clear over here, these drugs. The reason is these are inhaled products. As an inhaled product, we want them to act locally in the pulmonary tree. So PK isn't going to help us.

Yes; we have a PD parameter that we use in our studies and, in the older child, six and above, the PD parameter that we use is the same as we use for adults and that would be to look at the forced expiratory volume in one second using a hand-held spirometer.

However, we can't use PK because we are not looking at PK at the level where the inhaled product is working. So, one of the factors that we have to consider, then, is the route of administration. I have that up here in this particular box.

So, although we know that the beta-adrenergics are going to act the same in children and adults and the progression is the same, if we use this particular mode of administration, then what we have to do is we have to go back and we have to do full-fledged efficacy studies because we don't know what dose in the child is going to lead to the effect.

It is going to be the same thing for the corticosteroids, although they act in a different manner and they act mainly on the inflammation, if it is inhaled, we are going to have to do those studies again.

If we look up here at Montelukast, it was the first of the leukotriene-receptor antagonist products. It was approved in adults and it was originally studied in children. It was studied in children in the older age groups of six and above because the PD parameter we could use and the question was was the response to therapy the same.

Nobody knew if children had these leukotriene receptors, if they had them, were they activated. So we had to do full-fledged efficacy studies in the child. It turns out that they responded just like the adult. So, as a result, we now know that children have them and we feel that the response to therapy is the same. Again, the progression of the disease is the same.

So that puts us, for Montelukast, which we had up here for the older age group, we now know they are reacting the same and the studies that were requested in the written request said, do population PK to get the dose right and do safety studies.

Here is Montelukast now. So, for oral drugs where PK can be used, we can actually take and get them to follow down here.

Just a couple of other points I want to

make about asthma and these factors. There is even more concern here for these inhalation products. For asthma, if the child is less than six, many of them can't actually do the hand-held spirometer, so you can't use that PD endpoint in the younger child so we have to go back to signs and symptoms of asthma. So that is one of the other changes that we have to make.

The other thing is the device has to be considered in these inhalation products. So we may know how to use a device, or the child can actually use a device similar to the adult, but when it comes to the devices for the younger-page child, they have got spacers in different things. Different manufacturers have different spacers and products.

So we have to study, using efficacy trials, because there is no way to take any kind of PK or PD or any way to know if it is going to be efficacious other than to do the study with the product that is investigational in this age group and with the spacers and with the devices that are available to the pediatric population in the United States.

So I hope that I have tried to show you

how we use this tree and that it does provide a way for us to think about studying children. This is not a perfect decision tree. We have talked about making some modifications to it. As information comes back, based upon the studies that we have, we are going to be able to make some of those assumptions and turn them into actually evidence and feel much more confident that we can go one way or the other along that decision tree.

Thank you very much.

DR. JUSKO: Does anybody wish to clarify any questions?

DR. SHEINER: Just one question. For that class, it was some fairly large number, where you did decide that it was adequate to simply find out what the right dose was by looking at the PK, have you had enough subsequent experience with those drugs or prior experience when they are used off-label to indicate that, in fact, that decision tree for those drugs actually your judgments were more or less right and you did get the dose right and nothing turned up that you were giving too low or too high in general doses or anything like that.

DR. ROBERTS: Are you talking, Dr. Sheiner, about going down the right-hand side?

DR. SHEINER: The right-hand side; right. The ones where you are willing to believe those assumptions. And then you said, I think, in one of your first slides, you showed about thirty or so where you had done that. I just wondered if you had any follow-up experience and whether you were satisfied with the results.

DR. ROBERTS: We certainly have used it for the antihistamines, for like allergic rhinitis, because to try to study--first of all, we know that the disease progression is similar. We have assumed, and we now know from studies of these products, that there response to intervention is going to be the same. There is a great difficulty, especially in the child that is in the age group of twelve months to four or five years of age, that you can't really get a good assessment of whether they are responding to these products using the scales that we typically use for the older child or the adult because it is things like, "Are your eyes watering less?" "Does your nose itch less?" "Do you have less discharge?" Those kids can't answer those kinds of things.

So there we have successfully used information based upon PK and safety. We have

found, with loradatine that, in the population of the two- to five-year-old, they actually need less drug than the older population. They don't seem to be clearing it as well.

We have seen, in other instances, where we really would have gotten the dose wrong if we had just treated children as little adults. With etodolac, I mentioned, that was using a PK/PD. We need to use about twice as much as we would have anticipated.

With fluvoxamine, which is approved for obsessive-compulsive disorder in children eight and above, the original studies whereby we got labeling, actually, for fluvoxamine for this condition, when it was analyzed, there was an effect but it seemed to be that all of the effect was in the eight to eleven as opposed to the twelve to sixteen-year-olds. So we asked them to go back and analyze why that was.

In that study, when they went back, they found out that we were actually underdosing the adolescent and that you really needed to titrate them up to the adult dose whereas the eight to eleven-year-old boys, you could use the labeling that we had in the product, and the eight to

eleven-year old girls appeared to be being overdosed, so you had to be very careful about titrating them up too far.

So we have had examples of where we really had missed the dose. Of the twelve out of the forty--we just had three new approvals and we haven't had a chance to look at those labels yet--but twelve out of the first forty products that we labeled had either significant dose or safety information. So that is about one-third of those products to date.

DR. JUSKO: I think we will go on to Dr. Selen's presentation now.

Efforts to Optimize

Pediatric Clinical Pharmaceutical Studies

DR. SELEN: Good afternoon.

[Slide.]

As Dr. Rosemary Roberts said and Dr. Lesko said what you are hearing today is we are at the right place at the right time. We are having a lot of pediatric studies coming in. There is a lot of information coming in and there is a lot of intelligence going behind all of these things.

So what we are trying to do, really, is optimize and learn from these studies. Clearly, we

have certain facts that we know. We know that the pediatrics are not small adults and, in fact, Dr. Capparelli was reminding me, we also know that the pediatrics and adults are not so different from each other. Adults are not the Martians. So we can also extrapolate. But we can't really go by the weight-normalized parameters as well. We have some issues with that.

What are the other things that we know? We know that the pediatric studies are clearly complex. There are many issues and many study-design aspects and so I think we will have to be more careful in looking at the pediatric data and looking for studies.

So, knowing all of these things, then, the next question is can we optimize pediatric studies. To do this, in our Office of Clinical Pharmacology and Biopharmaceutics, jointly with other members from the Office of Clinical--actually, this is a big group of individuals. I don't want to, perhaps, go into all the individuals that are involved, but I would like to say that, with the joint effort of many individuals in the Center, we are trying to look at the ways that we can optimize clinical pharmacology studies.

For these studies, we know that now we are at the very beginning but we hope that these studies will continue to be optimized, provide information so that we will really have the public health benefits.

[Slide.]

I mentioned acknowledgments. There are many individuals involved and I am going to refer to Knowledge Database which is really starting from a research project including individuals as Dr. Roberts, Bill Rodriguez, Dr. Tandon and other individuals, Dr. Lesko and others. So this is an effort, really, to look at the incoming information and to make the most of this information.

[Slide.]

So what I would like to do is this afternoon, I have a few slides. I want to talk about this knowledge base, give you some background on this, and also get your input on this because this is, again, like Dr. Lesko was saying at this point--this has such a huge potential and we want to have a right questions asked. We want to sort of start at the right places and get the most of this information base.

[Slide.]

There are two primary approaches in here, two levels. One of those is more specific to the drug. We are looking at the factors that are unique to the study drug. Are they race effects, age-related effects or gender effects? As a result, can we optimize the dose for the pediatric patient so they will be treated--they will have the maximum benefit.

So the first level is drug-specific. And the second level, or the second objective of this information base, is how can we learn across studies because we are going to have many drugs coming in, like from the same particular class. Also, if you look at the way the metabolite is cleared from the body, they will also have some commonalities and maybe there is a way of looking at the similarities and looking at the study designs using this information and optimize them.

So there is a huge list of questions that can be posed. The whole sort of objective is, that I hope we will achieve at least some of it this afternoon, is to have your input on some of those aspects.

[Slide.]

As I said, we started working on this

knowledge base some time ago, on and off. It started as research project and it is sort of rapidly blossoming and I hope it continues to grow.

The main source of information currently is the studies that are coming in as pediatric submissions. This is our starting point. These are the studies that have been conducted as part of the written request lectures and also other studies that come in to the centers, pediatric studies, that have pediatric pharmacokinetic data are part of this knowledge base.

What we also like to include is also to have something to compare with that information, which is the literature data, if available, dosing information, and any other information such as the metabolism. That will be very critical how it is in adults and we will look for the similar characteristics or similar patterns in the pediatrics.

So we are trying to incorporate all of these things.

[Slide.]

As it stands, there are several different types of files in this knowledge base. There is a section that specifically deals with information

with data that includes such as specific information, the drug, the dose, the dosage form and patient characteristics, the demographics. If we have pharmacokinetic data on the parent drug, fine. If we have also the metabolite, even better. And it includes information such as individual data, obviously, and mean data.

Of course, again, the pediatric decision tree is also captured in here and how this drug was fitting or not fitting into any one of those boxes, how does this sort of fit into the whole picture of things. Again, this will also eventually help us sort these out as we improve on the decision tree as sort of the thinking behind it that will lead us and give us information.

[Slide.]

There are two questions that I will be posing at the end. One of those is essentially what will be things that we can be collecting in this database, what other information.

[Slide.]

The second question is going to be what will be the more appropriate questions. I am going to ask for your input on that as well, and how can we go about this. What are the best questions to

ask?

[Slide.]

Just to sort of give you a feel for the type of information in the database, I will select something from the literature, just as an example. I don't want to mask as a drug from one of our drugs in the knowledge base, but I thought, I will just pick a drug. It is adefovir dipivoxil. It is published. We can call it Drug A. I can just point out a couple of things that are unique to this because it will help in the discussion as this drug is primarily eliminated by the kidneys so there is no metabolism involved.

This is also one of the considerations in our pediatric pharmacokinetic studies. We talk about the ages. We talk about the maturation. So if we say that the kidney is mature at a certain rate, maybe after two years old, we don't know to have data from pediatric patients perhaps, we have to focus on. So this is why I selected this example and we can talk about that.

They have looked at two doses, 1.5 milligrams per kilogram and the other dose is 3 milligrams per kilogram which is, again, similar to what we have seen in our pediatric studies. We see

sometimes two or three doses and it is used for selection of a better dose.

The sample size is fourteen pediatric patients which isn't really very many. As a kineticist, I would like to see more because we know there is more availability in data. But, in this case, they have fourteen patients and the age range is six months to eighteen years. So it is a reasonable size, on the small side, but it is okay.

[Slide.]

One of their observations is the first block, is the charts that they are looking at, the area-under-the-curve values. Essentially, what they have observed is, after this twofold difference in dose, 1.5 milligrams or 3 milligrams per kilogram dose, when they look at the blood concentration time profiles, they could not see a difference. They all looked similar and they couldn't really tell which one had--if you were just going to look at the blood-concentration profiles.

The doses were twofold different but they couldn't tell the difference by just looking at it. They compared the area-under-the-curve values and they looked fairly similar, although there was a

twofold change in the dose.

Now, they are saying, okay, they have reported the dose as by body-surface area, milligram per meter square. When they do that, they could see a correlation between the dose and the area-under-the-curve value. So this is just becoming--it is kind of hard to read this but it is just axes, the Y axis is the area under the curve and the X is the dose.

In one case, it is by body weight and, in the other case, by body-surface area. So, depending on how you report this information, you have a different observation. This is kind of like the comment you made, Dr. Sheiner, earlier on quantum mechanics. Your observation is, perhaps, influencing the outcome.

Or decision, which parameter to report. If it is reported in one way, if it is milligram per kilogram, that is part of the knowledge base, are we going to calculate the body-surface-area-corrected parameters. Now, that poses another question because not every study, not every submission would include this information done both ways. And it may not be necessary to do it both ways, but it is a point to consider.

In this series of graphs, what we are looking at is now the correlations. On the Y axis, the parameter is the area under the curve. The first is the area under the curve. Then it is Cmax and they were able to measure concentrations eight hours after dosing, the last collected sample.

On the X axis, in each and every one of them, it is the age of the patients in the study. Since this is cleared by the kidneys, one would say, okay, after two years old, the kidneys will function as an adult and there will not be such a change in the area-under-the-curve values because it should be comfortable.

But what is happening here that, as the children are getting older, the area under the curve is increasing. So there is a change, age-dependent change, in their clearance. Now, you could point out and say, well, this is an oral dose. Maybe it is not just the clearance changing. It could be the fraction of dose absorption is changing. It is an apparent oral clearance, the F value that we don't know. So maybe the F value is influencing this observation. That is where we are seeing this age-dependent change and as the children are getting older, now the area under the

curve is increasing so the CL over F really smaller there than it is at the other end. So we are seeing a difference here.

So, given that now, which one is changing? Is it the clearance that is changing? Is it the fraction of dose that is changing or is it the combination of both? Now, we don't know that. But, at least, it illustrates one point that if it was just only a clearance-related issue or if it was the assumption that the clearance did not change after two years old, there is something that is not right.

There is something that doesn't exactly fit in.

Yes; you have a point?

DR. CAPPARELLI: Are these normalized at all to size and, if so, in what fashion? In other words, some of the patients were on different milligrams per kilo doses and you would expect, if clearance is flat based on body-surface-area allometric scaling that you would see this sort of phenomena.

DR. SELEN: You are saying that this is--no.

DR. CAPPARELLI: In other words, this is

raw data. Is this all 3 milligrams per kilo? Is this all 1.5 milligrams per kilo or has it been--

DR. SELEN: It is a normalization in dose, I believe. That is what I understand. That is why I isolated the example. So the normalization will take away the effect of the body weight, which is your question.

DR. CAPPARELLI: Right.

DR. SELEN: You are saying if the body weight is influencing this observation. If the publication didn't do that, let's just work with the premise, that the body weight is normalized so it is not the influence of the body weight because there are cases like that if you need to take into account the change in the body weight, you still see the age relationship. So that answers your question.

DR. CAPPARELLI: Right. If it is from the publication, then it would be by weight but it won't be by body-surface area.

DR. SELEN: Yes.

DR. CAPPARELLI: Okay.

DR. SELEN: Let's just work with the concept here because the example is not the specific publication. But let's just work with it

that they have taken into account the changes in body weight. They have normalized it appropriately and the change we are seeing can be attributed to the oral clearance change which will include either the change in the clearance or the fraction of dose absorbed, or both, which we don't know.

But we do see this and we do see this even when you normalize for body weight. So this is just an example that the type of information you see--but, sometimes, the type of information you see also is the area-under-the-curve values tend to get extrapolated more than our routine 20 or 30 percent extrapolations.

So then it becomes a problem. Then you have to look at the individual values, how accurate they are or how correct they are. So we have to also have an understanding of the parameters that are involved in this and sort of leading to the decision, going down the decision path.

But, nevertheless, there are examples like this that show that there is a good correlation between age and pharmacokinetic parameters. The reasons for that could be many of the things, including the metabolism, the maturation of the metabolizing enzymes or just an absorption event as

it might be in this case.

[Slide.]

What the authors have done, again this is not example-specific. This is just something to illustrate the point is they are comparing area-under-the-curve values, first of all, the comparison of the parameters are Cmax, C8 and they are just looking at the doses, 1.5 and 3 milligrams per kilogram and they don't see a difference in these two parameters. They are seeing, even with the twofold change, they can't detect a difference.

[Slide.]

Now, this could be for many reasons. Again, it could be the sample size. It is just to illustrate the point that--or maybe if there were more individuals in a certain group, they could have made differences. Or it could be just the pharmacogenetics. It could be individuals that have certain different metabolizing capacity.

One thing they have also looked at is the second bar, Graphs B and C. In this case, in these two slides, in these two charts, they are looking at the three parameters, Cmax and the concentration at eight hours and the area under the curve. In these three charts, or two charts and three

parameters, they have grouped the data by the ages, the age groups, the under-five-years-old and over-five-years-old. Again, they see a significant difference.

The point I would like to illustrate in here is not the significance for this drug but the relevance of breaking by age groups, and where do you decide it should be, at five groups, what break point, or based on the physiology, if this is really unrelated, that we are seeing, well, after three years, it should be similar to adults, so it should have been broken zero to two and two and older.

So there are many different combinations. Or one could say, perhaps, it should not be handled in this manner at all. This is arbitrary or artificial because we don't have all the supporting facts.

But, in any case, even with the small sample size, they are able to see significant age-related differences in the three parameters, Cmax, C8 and area under the curve.

So, technically, as in this example and other things that we are looking at, there are many components and many parts of the puzzle. While we

are looking at this information, knowledge base, we are trying to collect data from pediatric studies, we are trying to incorporate information from literature and we are trying to extend it to the point that we can really look at it and learn from it and use it as information for designing other studies, for looking at dosing recommendations.

So there is a major emphasis here. Of course, this is a beginning. I certainly hope it will continue and develop into a product that will benefit for the pediatrics.

This is an old article, journal, that says pediatrics is for children. I guess it is needless to say that is all, I guess, the reason for doing all these efforts and activities.

[Slide.]

So the two questions to the committee, and, at this point, I can turn it to you, Dr. Jusko, and we can go with those.

Committee Discussion

DR. JUSKO: As we discuss the two questions that are posed, perhaps there could be some further clarification of the pediatric database.

DR. SELEN: Certainly.

DR. JUSKO: Am I correct in assuming that most of these studies are small studies like you have described, fourteen to twenty children, various drugs.

DR. SELEN: I think the point you are making is an excellent one because depending on the type of the study, if it is a traditional study design, the sample sizes are smaller. So we have sometimes twenty children, or twenty-four or thirty. But if the study design is a population pharmacokinetic design, then we have more datasets and more patients.

So it varies across. They range. There are not more than a hundred patients in a study. I have not seen a number exceeding that. But they range from, I guess, twenty, twenty-four, in that ballpark.

DR. JUSKO: Typically, are the children those in whom the drug is indicated as opposed to, say, normal volunteers?

DR. SELEN: They are patients. They are patients. The only exceptions to this might be the very, very early studies before the ethics rule that we may have had some gabapentin data that might have been conducted in healthy volunteers,

some pharmacokinetic studies. But I could easily say 99 percent or more would be patients because this is an effort of emphasis that has been on patients for the last three--Rosemary, you can answer that.

DR. ROBERTS: Actually, this is a very good question. Unlike the adults, where phase I studies, for certain product areas, are done in the healthy adult who is informed of the potential risks and signs an informed consent, in children, because they do not sign their own conformed consent--we actually had a meeting of our Pediatric Advisory Subcommittee of the Anti-infectives Advisory Committee that was formed early in 1999, and one of the first ethical questions we took to them was is it appropriate to do nontherapeutic studies in the normal child versus the patient.

The advice we were given, and the advice we adhere to, is that children should benefit from being a participant in a clinical trial so they either have the condition or are susceptible to the condition.

Actually, the reason we took this was because we were amazed at the number of traditional PK studies that were being done in the pediatric

population or had been done. So we took this issue and, from that point on--this was actually in November of '99. Subsequent to that, we only asked for patients in the pediatric trials and we also, at the recommendation of that subcommittee along with a mandate from the Children's Health Act of October of 2000 have incorporated the Subpart D, the additional protections for children that were part of the departmental regulations but not a part of our own regulations, we now have incorporated those additional protections for children into the FDA regulations.

DR. SELEN: Thank you.

DR. JUSKO: And then one more question on the database. Typically, these studies are studies purely in the particular pediatric-patient group and there are typically no comparison studies with adults, unless it is from the literature or previous studies done by the company.

DR. SELEN: The studies and the written-request letters are always for the pediatric patients. So our source is coming from pediatric studies. We try to sort of have historical data or adult data as a comparator. But, at this stage of the game, it is fairly

limited. But we would like to have that for everyone so we have a good comparison.

DR. JUSKO: Richard?

DR. LALONDE: In response to what other information should be collected to pick up on Edmund's comment, I would encourage you to relook at how some of the pharmacokinetic parameter-scales with body size. If you are going to have a rich database, that would be interesting because, as you pointed out, the differences you saw because of age there are most likely due to how the doses were normalized per kilogram and clearances don't change as a linear function of weight.

So it is really kind of an exponential function. So it would be interesting to see, maybe across compounds that are eliminated by different mechanisms across different age groups, as you look at body size, to see the allometric approach, for example, there is a tendency to predict very well, body surface area, weight, all those things, because I really think it is actually--sometimes people are misled by information. They say, it looks as if the disposition of the drug is changing as a function of age when, really, it is not.

DR. SELEN: That is a very valid point. I

can't say for each and every one of the things that applies, but there are some cases, even after you correct for body weight, you still see the age effect. It is just the case that I guess the maturation is an event in terms of the enzymes that are responsible for metabolizing the drug.

DR. LALONDE: I think that the question is how do you correct for weight. I think that is a key thing to see if you are going to take away all these body-size effects or not.

DR. JUSKO: In that particular case, and in many cases, I would go further and say it is simple and straightforward enough to obtain information on creatine clearance. That drug is one you stated was primarily cleared by the kidneys. Having a relationship to creatinine clearance that, in turn, are related to body size might have considerably clarified what was going on.

DR. SELEN: You have a good measure of the--

DR. CAPPARELLI: That is not that easy to do. In looking at drugs, especially in these populations, serum creatinine based in adult laboratories, the precision with which you get

back, you are dealing with creatinines of 0.2 versus a creatinine of 0.3.

Getting urine collections, which I think is an important consideration in study design, maybe not for this aspect, but we are always trying to maximize information when we are collecting it in kids. But you really need to have--looking at serum creatinine, I have been surprised at how poorly it predicts, in a sort of relatively healthy kid population, the clearance of renal drugs.

I think part of it has to do with the precision issue and the equations that we are forced to use to sort of estimate creatinine clearance. There becomes the other issue, if you actually want to measure creatinine clearance, which probably would help, but I think one of the issues there is that you are getting full urine collections becomes difficult.

One of the things that I would add, in terms of additional information and it was maybe alluded to earlier is, besides the age, is looking at Tanner staging in that sort of window where that becomes important and also looking at the pharmacogenomics for the drugs that are metabolized because one of the things you see with a lot of

these curves is you will have one or two outlined points which confound your whole conclusion.

So if there is an explanation for that that is something that is easily measurable, I think that that should be included.

Then, lastly, just getting to the point that was I think brought up by Richard as well, we really need to be thinking about presenting the data in a unified fashion. In terms of the sizing function, weight is probably the best way to dose but it is definitely not the best way to describe PK parameters.

Going with allometric scaling which doesn't account for all the age effects, and it certainly doesn't count for some of the bioavailability effects is important. But I think it is one measurement that can be done accurately; i.e., weight. You don't have to get a height and a weight. There is at least a scientific basis for utilizing that sort of an approach and presenting the data in that fashion and maybe looking across several renally eliminated drugs and looking at the fractional excretion of the drugs may provide some very powerful information as long as we scale it appropriately.

DR. SELEN: Thank you. I also wanted to go back to the creatinine clearance because what is your experience with systatin C. We are looking for different ways of getting that information about the kidney function. There are some publications on systatin C as being a potentially useful measure, more precise and more accurate.

DR. CAPPARELLI: I haven't seen it used in pediatrics at all. I think, clearly, we need more information. But, again, say you are looking at your antibiotic where you don't have a life-threatening infection, kids are relatively healthy. I think that, in the relatively healthy population where they don't have hypertension, they don't have a lot of comorbidities, you may not see the variability in renal function that you do, say, in an adult population that isn't accounted for by size once you get out of the initial maturation phase.

DR. SHEINER: Did I understand you to say that the database consists of the raw data as well as the analyses?

DR. SELEN: Currently, it is just the pharmacokinetic parameters, individual ones and--yeah; I mean, it can--

DR. SHEINER: That is the biggest thing; get the original data.

DR. SELEN: Get the raw data.

DR. SHEINER: Doing "meta-analysis" when you have essentially transformations of data by different models, different folks, some of them have standard errors, some of them don't have standard errors, some of them have taken out outliers and some of them haven't, for all kinds of reasons. I am not impugning anybody, but trying to put that together and draw a conclusion from that is--you have got to work three times as hard as if you just have the original raw data.

So I would really encourage you to have a standard PK data form. It can't work for everything, but PK is pretty reasonable and with information on when the sample was drawn, when the things were taken, so you can get the raw data in there. Then you can really pool data and get the power from it.

Do you have any information in there--in the population PK studies, what information do you generally have about dosage?

DR. SELEN: Whatever is provided.

DR. SHEINER: Okay; there, again, trying

to know something about what actually happened within the last couple of half-lives would be useful. There are forms, at least, where you can inquire. I am not saying that they are accurate, but they are better than saying that, if somebody is on a BID drug, then they took it every 8:00 a.m. and 8:00 p.m.

So I would say that the quality of data could really be improved by attention to getting the details.

DR. SELEN: I agree wholeheartedly. Thank you.

DR. DERENDORF: Is there any pharmacodynamic data in the database?

DR. SELEN: This is just the beginning. We have a few studies, some pharmacodynamic information. But I think, as these studies come in, obviously, we will be incorporating it into the database, so there will be some.

DR. DERENDORF: In the first presentation, I think an example was mentioned about that you needed twice as much than you thought?

DR. SELEN: With the drug clearance being--I think was it--

DR. DERENDORF: Was it twice the dose or

twice the concentration that you needed?

DR. ROBERTS: We had to go twice the recommended lower dose in the adult.

DR. DERENDORF: But the concentration that you produced was the same?

DR. SELEN: The target usually is the concentration exposure profiles, isn't it, that we try to match?

DR. ROBERTS: Yes.

DR. SELEN: So if the dose wasn't really providing that concentration, then we had to double, like the example I had, the clearance was much higher in the younger group so the area under the curves were very small, or whatever it was, the clearance. So we tend to see the same trend that the drug level are lower in the pediatric--

DR. DERENDORF: I am saying don't take that for granted because, just as enzymes mature, so do receptors and the sensitivity may change and the EC50s may be different. In the adult, that is well documented. In the kids, there is not much data out there that I know. I would look out for it.

DR. JUSKO: I think there was the implication that, with this additional should be as

much pathophysiological information about chemical parameters, the disease states. It sounds like there is a potpourri of different conditions. It is going to be difficult if you have the complications of a particular drug, of a particular patient group and different pathophysiology that may exist.

DR. SELEN: I think that is sort of, with certain drug--I don't want to go into the details of this, but it becomes very important what stage they are at. It can sort of give us a handle on how much of the drug is being absorbed, so it becomes very important, the point you are making, that we know exactly if they are really at a place where they can absorb more or less. It is the underlying condition.

DR. JUSKO: To what degree can you examine these current studies for their possible faults and thereby provide recommendations for improved protocols for future studies? This last one, the one you had from the literature, had they given an IV dose, along with an oral dose, it might have clarified a lot what was going on.

DR. SELEN: Sometimes I wonder if stable-isotope studies--there are so few

publications in pediatrics with those. I have seen a few, but there are very, very few. So would they have helped, for example, to look at the metabolite patterns profiles? Or have, like you said, one of them labeled and then you have a true assessment.

But, again, these studies could be complicated and you have to wonder if the end was going to be justified maybe for a selected compound. But it is clear we are going to learn a lot from these studies and, hopefully, we will be able to make knowledge out of the information.

DR. JUSKO: If, as has been brought up, there are problems in measuring creatinine in pediatric patients, then it should be a fairly straightforward task for the companies doing these projects to enact a more specific and sensitive assay to get such measurements more accurately because changes in renal function clearly are important to document.

DR. SELEN: It seems one of the things we were looking at with systatin C, for example, it looks like there is a range of companies that do the analysis and there is a huge range of prices for the assays. But, perhaps, if there was a lot of interest, if the method was developed further,

it could be reasonable, perhaps not very expensive, and maybe a preferred route to go.

We kind of looked into that a little bit. But it is a good point.

DR. HALE: I have a question here. Is there an effort made to coordinate this database with adult data? Is that a conscious decision you have made?

DR. SELEN: That was one of Dr. Lesko's points.

DR. LESKO: It seems we have to sort of get a handle around all these data. Part of the problem is trying to figure out what we have and what would be useful. For example, if we were to look at this database, it seems to me something that would be helpful would be to able to move drugs or drug classes from one box on that decision tree to another.

For example, we have, from Rosemary's data, 35 percent of written requests require efficacy-safety. Let's put safety aside because that is going to be required in any case. But now we have efficacy. If we were to go into that efficacy database and, in fact, look at PD information, that might be clinical outcome, it

might be biomarkers, it might be surrogates, and look at the exposure-response relationship for that in the pediatrics, then pull out corresponding data from the adult database, what would be the criteria to say that that is similar enough so that, in future studies, those drugs or drug classes would require only the PK study; in other words, reduce the requirements for studies in pediatric patients through a statistical exposure-response type of approach.

So, one of the questions would be what would be an approach to deem two exposure-response relationships similar. That is one of the questions of research, I think.

Lewis asked the other question. On those drugs for which we have deemed pharmacokinetics and safety to be the way into the marketplace, what has happened in the post-approval? That is sort of testing that box as well and I think we can do that over time when we have more experience. Right now, there are not a lot of drugs that have been approved in that box.

There is another part there that says conduct PK/PD studies in kids when it is not reasonable to assume a concentration response

relationship is the same. What if those studies were looked at again with that PK/PD study compared to a PK/PD study in adults; could that comparison be made to sort of change our thinking on that?

So I think there is a methodology question here in terms of comparing these exposure-response relationships and setting up some system of decision-making that we say they are similar or not.

Let me throw my second part that I think we need some input on. We have encouraged sponsors to do sparse-sample strategies when possible given the nature of the pediatric populations. There seems to be an uneven record with these studies in terms of them providing answers that we would like to know.

My impression--I don't have numbers, but others that look at this data all the time can probably say is that we reject quite a few of those for a variety of reasons. I guess one of things I would like to see us get to is some sort of standardized approach to doing these sparse-sample strategies in kids that we can all agree would be a reliable method to do that. That might be--again, given the time we have, we can't talk about it all

today--something in future. We might want to come forth with a proposal of template, if you will, or something like that for sparse-sample strategies and use that routinely in kids.

So those are some thoughts, if anybody has any comments on either one of those two things.

DR. HALE: That sounds really reasonable to me. I think one of the things that--this strikes me very much as a bridging kind of situation to a special population. It just happens that these are pediatrics rather than a different race, et cetera.

This probably isn't what you want to hear but it strikes me that, in a lot of cases, it is going to be a little bit idiosyncratic. When you talk about your database, it seems like it is going to be so specific to the therapeutic area--once you get outside things like dosing regimen, body weight, age, things like that, it seems like there are going to be enough therapeutic singularities that I am not sure that things are even going to match up.

DR. SELEN: You have a good point there. We have discussed this because, again, it comes back to having things standard so it is earlier to

put them all together and pull them and look at them at the same time. But, even for the same therapeutic area, depending on the age of the child, the end measures are different.

So there will be differences. It is not going to be avoidable. We have to accept that because this is the pediatric data and this is a unique feature of these studies, that is it not similar to adults that we can have one standard form.

But if we have an underlying common form and some small variations on this, that will have gone a long way. That will work tremendously because, you are right, that, for each therapeutic area, we will not be able to have the same identical format, the same template. It is not going to happen. We won't see all the age groups. We won't see the same--that is a given.

But, if you were going to look at, for example, in terms of how drugs are cleared, if they are P453A drugs, or if they are more the renally eliminated drugs, perhaps we can go from those angles and have some uniform aspects for those elements.

So there is a lot of interest that perhaps

we can sort of strive and make a standard form, a standard platform that will apply given that it is not going to fit in each case. So it will be some certain parameters that will perhaps work.

DR. HALE: One other follow-up question here or suggestion, both. I guess I am presuming, in many cases, the people doing studies in pediatrics will be the same sponsor that has done adult trials and will already have a pretty sizable experience base in terms of what is going on with that drug, that therapeutic indication and will confer with key opinion leaders, et cetera, to figure out what should be the same, what should be different, and actually have already answered these kinds of questions when they propose doing pediatric studies.

So how much are you looking to sponsors to input into this on a case-by-case basis as opposed to up-front putting some guidelines in place.

DR. SELEN: We always welcome the interactions. I think the divisions really work very closely with the sponsors when the studies are being designed. So I think that information, that link, is there. So this is just sort of getting over towards here as to what can be done better,

what other things we should be thinking of.

But this is not replace interactions that sponsors have with the divisions. I think there is a very good dialogue between the sponsors and the Agency.

DR. JUSKO: To follow up on that, I think it eminently reasonable that the sponsor incorporate these data into whatever population, PK or PK/PD analysis that they may have developed for the drug in the normal and special-population groups that they have studied.

DR. SELEN: Ideally, I would say I hope that happens. But I think, perhaps, sometimes the realistic flow of things is that there are time lines and there are certain things that have to be meeting a certain question. So maybe some of the questions that are on the broader scale, can we look at this in a global view, can we learn more from this, may not be the objective for a drug-development program.

So I think there are some sort of similarities but I think it will probably have a lot of different perspectives as well.

DR. SHEINER: I would like to say something to Larry's points. That flow chart is

useful in putting them into boxes. Maybe one of the things you could ask of the people who use it is that when, for example, you put them in the box of meeting efficacy as well as safety, there are two possible reasons for that.

One is that you do not yet have the information that will allow you to accept the assumptions that would allow you to go down the right-hand side and the other one is you actually know something that says it is not going to be the same.

It seems to me it is the first group, the unknown ones, that the data gathering wants to focus on and the analysis wants to focus on so that they can be moved or drugs of that class can be moved subsequently, as Larry suggested, into the other boxes if it turns out that you suspected some problem but, in fact, it didn't arise.

Let me just make one quick comment as one of the guilty parties here on the sparse-sampling design. I really do believe that I always did say that you would only do that if you couldn't do something better. I am sure it wasn't heard that way, but I would repeat that. It is not a good design. It is sometimes the best you can do and I

still believe in not making the best be the enemy of the good. So, sometimes it is good but I have come to the point of view that an observed dose, if it is oral drug and it has a half life of more than a half an hour, is almost necessary and more than one sample on the occasion after that dose is also very important.

So I would be very interested in working with the committee and others on a template that says, don't waste your time. If you don't know what dose they took, you don't exactly know when the sample was drawn and you have only got one of them, you are fooling yourself.

DR. LALONDE: If I could just add a comment to what Lewis was just mentioning there in terms of these boxes in the decision tree, it would interesting to see the top two assumptions, again, the one especially about similar disease, I think, progression, to see if ever that assumption was not satisfied, or the second one was satisfied, that you had a similar response based on the experience that you have to see if you might still be able to put these drugs down the right-hand side of your decision tree.

In other words, you might say, well, we

are not quite sure about the disease etiology between children and adults, but the drug--say, it is blood pressure, for example, that the drug does lower blood pressure and when we have tested this across a bunch of different compounds, so far we have seen that it seems to work out fine.

So, just a thought.

DR. ROBERTS: Let me make one comment there. Actually, we do have an example where the disease progression is different. That would be HIV. HIV presents, in children, much differently and the course is much different in children than it is in the adult. However, we do know that we are targeting the same virus.

Using the pharmacodynamic marker of the HIV RNA levels and targeting so that we can bring those levels down, there we have been able to check that just to lower the similar response to intervention and go down the right-hand side. So that is one example where we have been able to do that.

The other area where we could probably get away with that is in the area of the antimicrobial agents because, again, you are targeting the agent. We know that, for some of these agents, you need

to--for instance, with the beta lactams, you need to target to get above the MIC for a certain period of time in your dosing interval in order to be efficacious. So we have some where we can do that.

DR. DERENDORF: Are there any plans to expand this approach to the elderly as well, because I think all the things that we have said, we can apply just as well to the old and very old patient.

DR. SELEN: I will pass it on to Dr. Lesko to respond for elderly plans.

DR. LESKO: The question was with plans, and I would say no. Plans haven't been talked about. That is not to say the suggestion isn't good. I think there is some urgency with this database because so much has been done, so much has come in. I think there is an expectation we need to do something with it whereas with the elderly, we have had other ways of dealing with that.

It is not unimportant but I think it is not in the plans right now. But I think what we can learn here may be transferrable to the elderly and other special populations.

DR. CAPPARELLI: Getting back a little bit to the HIV example and disease-state progression, I

am a little confused by the terminology in the sense of this is a slightly different change in wording as to what had been, I think, in the '94 Pediatric Rule where there were issues of disease-state similarity or similar effects.

If you start extrapolating down to the newborn where HIV, as you say, is much different but you start looking across other disease states, the progression, and I see progression as sort of the longer term, is much different for almost every disease in newborns than it is for adults.

So even though some of the other drugs move into those categories, maybe I am misinterpreting progression or I am overextending the definition because, it seems to me that you are going to end up with cutting across pediatrics into maybe separate age categories that end up going one path and down another because you have got some information.

But, clearly, in the very youngest infants, I see almost everything going down to the left.

DR. ROBERTS: I won't disagree. We have had very few studies in the neonate as a result. They are so different. I think, with respect

to--there were lots of comments on what we should use for sufficiently similar conditions in the pediatric and adult population.

This is what we have come up with. I won't say it is the best but, clearly, the onset of the disease and the characteristics for HIV are different in the pediatric population versus the adult, especially as you get younger. When it comes to the neonate, they tend to be in a category in and of themselves. As a result, we have very few studies that have gone down into the neonatal age group because we don't really feel we can extrapolate.

DR. SELEN: Even the neonate, one week old versus two weeks old are different, as you know.

DR. CAPPARELLI: Right. But I think some of the thoughts in terms of if we are trying to achieve an effect, and getting away from efficacy, and we know the mechanism of action, there are certain things that we can look at to assess similarities. I know, at least our group had proposed looking at effects of catecholamines on vascular tone, for instance.

While it may or may not be different, the disease state certainly is going to be much

different. Some of the effects that we are shooting for clinically are the same and I think the utility of some of that information is the greatest in this population because they are the group that has the most difficult-to-predict pharmacokinetics.

Clearly, they are a difficult group. Even within the group, it is difficult to know what the appropriate dose might be between just a couple of weeks of age or different degrees of gestational age at birth.

DR. ROBERTS: We actually have a Neonatal Working Group. It is with the NIH where they are trying to actually lay out some of these issues that are peculiar to the neonatal population and trying to decide the best ways to move forward with studies in that population.

DR. JUSKO: I think ours scheduled time frame leaves us five minutes to conclude this topic area. Perhaps we could finish with any burning indications for Question No. 2, what research questions and priorities would best serve pediatric healthcare.

Would that be okay, Larry? We have sort of been discussing these in the context of all that

we have talked about so far. In my view, and as Hartmut has expressed, a very high priority would be further evaluation of pharmacologic or pharmacodynamic differences in the younger age group compared to adults.

I believe you are posing this question in terms of the available database but probably in the context of looking forward in the future as well and advising companies.

DR. SELEN: Exactly. This is the beginning. This database is the beginning. We have just started and there is a lot more room to make this grow and I certainly hope it will continue to grow because there is a lot more to learn from this. So we are looking for all the ideas, input, that you have that we can really optimize the information from these pediatric studies.

DR. CAPPARELLI: Along those lines, and along the lines of moving drugs from one box to another, I don't know if much has been done in terms of surrogate markers that one could use. It would be similar between the adult and pediatric populations that could be integrated into these PK studies easily. I would be thinking about maybe

first approaches in terms of the classes of categories of looking at those things and getting a handle on some of those biomarker relationships, if not a true surrogate marker, but at least to give, I think, more validity to our exposure targets that we are shooting for.

DR. SELEN: I think you also said about genotyping earlier on, so, to have an understanding of the extreme values. Thank you.

DR. JUSKO: Any other further major comments? I think that will be sufficient, then, to conclude this topic area. We have identified that this is an extremely fascinating database and there are all sorts of opportunities to mine it for interesting observations and important factors affecting drugs in young children.

We will resume in fifteen minutes.

[Break.]

DR. JUSKO: Topic No. 3 is entitled Scientific and Practical Considerations in the Use of Pharmacogenetic Tests to Determine Drug Dosage and Administration. Joining us for this session is Dr. Richard Weinshilboum who will be speaking shortly.

Also, by telephone communication is Dr.

Wolfgang Sadee from Ohio State. Wolfgang, can you hear us? [No response.] I am told he can hear us but we can't hear him. Also, Dr. Mary Relling may in phone contact as well. Mary, are you there? [No response.] No Mary.

Beginning this session is a presentation by Dr. Lesko.

Topic No. 3

Scientific and Practical Considerations in the Use of Pharmacogenetic Tests to Determine Drug Dosage and Administration

***

Current Experience and Clinical Pharmacology Perspective

DR. LESKO: Thank you. I just wanted to clarify something before I get into this because the agenda that has been circulating has a few errors and I don't want to offend anybody. Dr. Sheiner is an M.D. Dr. Weinshilboum is an M.D. Dr. Mary Relling is not in Ft. Lauderdale, Florida. She is actually at St. Jude's in Memphis, so there is a little glitch on our schedule here and I just wanted to make sure I said we are sorry and clarified it.

[Slide.]

Now, to get down to the business of genetic tests. I think this is a very exciting topic for us to be talking about in this subcommittee. In bringing this to the committee, I wanted to let you know that I am wearing a different hat right now because I am Chair of an FDA Working Group on Pharmacogenetics and Pharmacogenomics. In this working group are representatives of all our centers, the Center for Devices, Center for Drugs, Center for Biologics, NCTR and all disciplines, clinical, clinical pharmacology and preclinical.

This group was organized over one year ago by the Center Director in CDER and it reflected, I think, her enthusiasm for us to explore the applicability of this scientific in drug development and regulatory decision-making and, in particular, can the science of pharmacogenomics impact risk assessment and risk management.

So we have been discussing this for some time. We had a public workshop in May of this year sponsored by PhRMA and FDA and DRUSAFE. It was a very successful workshop in identifying issues. Amongst the issues we discussed at that workshop were issues surrounding the use of genetic tests to

determine drug dosage.

So this meeting is the first step and the first public discussion of this for us. There are going to be some subsequent discussions of this topic, perhaps at the Oncology Drug Advisory Committee meeting in February. That is a possibility, and then, certainly, discussions before this committee in future.

[Slide.]

So this is the introduction to really our keynote presentation by Dick Weinshilboum. But I wanted to set the stage.

We are using as a model compound for discussion here 6-mercaptopurine which, as I said earlier today, is given chronically to maintain remission in children with ALL and it is also widely used in other populations.

[Slide.]

I presented this all earlier so I am just going to fast-forward and just clarify terminology which is always brings confusion to a discussion of genetics and genomics. I am on the right-hand side, focussing on pharmacogenetics, the study of genetic variations amongst individuals affecting liver enzymes that metabolize drugs. That is the

narrow world in which we are focusing today.

That is not to say there isn't a broad world of pharmacogenomics on the right which I will sort of describe as the study of genetic variations affecting the rest of the genome that affect drug response, and that covers receptors and transporters and a whole bunch of other things.

But, for simplicity, we will be on the right.

[Slide.]

I would also like to make a distinction for the purposes of discussing this between two types of genetic tests. The first is the genetic test for diseases. This would be using these tests to identify a potential patient's risk, prognoses, diagnoses. I like dividing this because there is a big difference, I think, in the level of public concern about confidentiality, equity and privacy when we are talking about these types of tests, tests for disease, as opposed to genetic tests for dose dosing.

We are in the latter category for the purposes of this advisory committee. These tests, in contrast to the other ones, are intended to be used to optimize dose and frequency. This is

consistent with the public's expectation of the agency which is to facilitate safer and more effective drugs.

[Slide.]

If we take a look at the current 6MP label language, one could argue that this is not necessarily optimal language based upon what we know about this drug today. I don't know exactly when this label was updated last. It is an old drug. This is from the current PDR. What it says in the Warnings Section of the label; "There are rare individuals with an inherited deficiency who may be sensitive to the myelosuppressive effects of the drug developing rapid bone-marrow depression."

It goes on to say that, "Substantial dose reductions may be required to avoid the development of life-threatening bone-marrow suppression." And then it goes on to describe it a little more.

It does not say anything in great detail about the frequency of these rare individuals in the target patient population. It does not go on to say what magnitude of a deficiency patients have and what the dose ought to be reduced to. These are all possible improvements in the label if the evidence is there to support to inclusion of the

information.

[Slide.]

This is just a suggestion. It is one that came from some of our discussions in our working group. There is nothing official about it. It is a proposal to say how can genetic tests improve a label, and this is an example.

The first step is where does this information go on a label. One could imagine this information in the clinical Pharmacology Section of the label where we talk about wide interpatient variability and the inactivation of 6MP by a specific enzyme to an inactive metabolite and then talk about the prevalence of the different genotypes in the population with 10 percent of the population having intermediate activity, 0.3 percent are virtually deficient.

One could also argue that this information could take a more prominent role in the label. Under the Dosing and Administration, for example, some information could be provided about the availability of genetic tests, commercially available, and that prescribers might consider using this test in patients with regard to their TPMT status.

There is also a suggestion here about a possible reduction in dose. So that is an example of how genetics tests might be incorporated into the label. It is only an example for discussion purposes.

[Slide.]

When we have discussed this internally, some of the discussion revolves around, for a genetic test, for this one specifically as a model, who would be the patients most likely to benefit. In this case, one might argue, that the patients in whom signs of toxicity, for example, based on CBC counts or neutrophils, those in whom these signs of toxicity occur early in therapy might be tested to determine their genotype. This is different than every patient being tested for their genotype.

Another target population might be those patients receiving combination chemotherapy where the combination drugs, each of which has their own similar toxicity or overlapping toxicity and it may be unclear which of the drugs in the regimen may, in fact, be causing this problem; for example, neutropenia.

Those might be two situations where testing might be facilitating better drug therapy.

[Slide.]

In addition to those, I wanted to share other issues that come up in the context of 6MP but I would ask you to sort of think about genetic tests in general. What if I was talking about a 2D6 test, for example, and incorporating that information into a label of a product that is a 2D6 substrate.

With this drug, specifically, why hasn't this testing been incorporated into pediatric-oncology standards of care? There may be other ways to get by with this drug, as we know. Would this add something to the standard ways of monitoring therapy.

Another issue that has been discussed is does the prevalence of low TPMT activity, which is 1 in 300--the intermediate is 1 in 10--justify routine testing of TPMT status? Does it justify optional testing? Does it warrant getting this information into the product label?

A third issue that is of concern would be how reliable and available do commercial genotype and phenotype tests for TPMT status need to be? Again, this is true of any genetic test. In the absence of overt toxicity, what evidence supports

the efficacy of a lower dose of 6MP in those patients with poor TPMT activity. One would lower the dose for safety issues. What do we know about efficacy under those circumstances?

Now, when I say issues, the issues are those issues that would prevail in the discussion of standards of evidence, issues that would come into play in getting information into a product label for a genetic test. I don't think they would be that much different in cases of other genetic tests.

[Slide.]

Some of the questions for the committee, recognizing, again, we have limited time today. We don't expect full answers to these but we would like bring them back at the right point in time; what major findings would support inclusion of a genetically tailored dosing regimen in a package insert? What is the evidence? Where in the label would this information best go to be most effective in optimizing drug therapy and under what conditions, what evidence, would testing be best be put in the label as optional or mandatory?

They are unanswered questions but they are questions we are going to have to struggle with as

these tests become more mainstream and widely available.

So, with that, I am going to leave the remaining time to our guest, Dick Weinshilboum.

I will turn it back to Bill.

DR. JUSKO: Thank you, Larry.

We will go on to Dick. Before we proceed, we wanted to see if the people listening on the telephone are able to communicate with us. Wolfgang Sadee? [No response.] Mary Relling? [No response.]

Assessment of TPMT Testing and Impact on Risk Management

DR. WEINSHILBOUM: First, let me say thank you for having me and let me thank Larry. Secondly, let me say the only reason I would possibly be here today is because of TPMT because I flew here from North Carolina where, as of last night, I was meeting my newest granddaughter, the only granddaughter and the newest grandchild. Today, Larry, by some sheer random chance, is the birthday of the mother of that granddaughter, so I am in serious trouble with my wife and there is no other topic in the world that would get me here other than TPMT.

[Slide.]

So, with that introduction, let's--I look upon what you are doing here--first all, I am delighted to be here because I remember Carl Peck inviting me to the FDA about ten years ago and I was saying things like pharmacogenetics and pharmacogenomics and TPMT and it was clear the time was not ripe.

[Slide.]

Let's begin what I think is basically going to be a step in a process. That is what Larry said. So the drugs we are talking about here are the thiopurine drugs, 6-mercaptopurine, 6-thioguanine and, of course, azathioprine which has an M and azol up here through both and through both nonenzymatic and glutathione-dependent processes is a prodrug that is converted to 6-mercaptopurine in vivo.

[Slide.]

What we are really talking about is a twenty-year history, and I think you are going to hear this recapitulated with 2D6 with regard to trying to understand--and this is my definition of pharmacogenetics which is a little different than Larry's because, from my perspective, it is the

study of the role of inheritance in variation among individuals and their response to xenobiotics including those that are regulated by the FDA; that is, drugs.

So I define pharmacogenetics fairly broadly. I will tell you what I define pharmacogenomics as, and, not taking a Taliban-like approach to the theological underpinnings of the definition, I will let anyone else believe anything they want to about this. But I know we have got Howard here. He will keep me honest and correct anything I say that is wrong.

[Slide.]

So the targets have been traditionally, as Larry said, drug metabolism, genetic variations of drug metabolism. This is really where the field has come from and, as a clinical pharmacologist, I am delighted to say it, in general, has begun with clinical observations so it has been bedside to bench and back to the bedside.

What we know, as Larry was pointing out, is that the same genetic variations will apply equally well to drug transport, to receptor interaction. I noticed one of your questions related to haplotype and I will use that word again

later because what I view we are going to do here is just raise a series of issues.

There aren't any answers. You will eventually have to come up with some pragmatic approaches, but we need to at least highlight the questions. In many ways, TPMT and 2D6, if they didn't exist, you would have to invent them because they have served as demonstration projects to highlight issues.

Then we have to say what are the practical ways of dealing with these issues.

[Slide.]

This is where it all started. This shows you the biotransformation of 6-mercaptopurine. Even the Mayo medical students, to whom I have been teaching pharmacology for thirty years, know that xanthine oxidase is involved in this process some way or another and there are rare patients who have hereditary xanthine oxidase deficiencies who are at severe risk for toxicity with these drugs but they are extremely rare.

George Hitchings and Gertrude Ellion, God love them, knew when these drugs were developed that S-methyl metabolites were found in the urine. The enzyme was first described by a man named Remy

who is retired from the Department of Biochemistry at Bowman Gray, or I guess, Wake Forest University Medical School.

I was in Winston Salem this morning. That is where I started my tour here because that is where my daughter did her residency in pediatrics and where she practices pediatrics. So this enzyme had never been explored in humans until 1978 when we published a paper and said, is it possible that--this was an assay for this enzyme--that there might be differences among individuals in this pathway and, if so, that they might be inherited and, if so, that they might play a role in individual differences in therapeutic efficacy and toxicity of these drugs.

Obviously, the reason Larry invited me to fly up here from North Carolina was the answers are yes, yes and yes. So, if that is the case, then what are data and what lessons--because that is really the important thing, not the specifics but the lessons that might come out of it.

[Slide.]

So what we did was develop an assay for the enzyme. We weren't thinking this way then but, Howard, these were phenotypes that we were going to

be looking at and a radiochemical assay and we were looking at it in the red blood cell because I am just a poor old clinical pharmacologist and I wanted something that might actually be useful in a patient where we could draw a blood sample and determine what might be going in.

[Slide.]

What we found, and this is a Northern European population sample of blood donors at the Mayo Clinic, was, among 300 randomly selected subjects, about 90 percent of them had high enzyme activity in the red cell--and, in case I forget to tell you, the NIH study sections, and I am on the Council for NIGMS and they have been funding my grants for these thirty years, but study sections kept saying, "This guy is so crazy in Minnesota, he thinks that red cells are the liver."

No, no, no; we never thought that. That was always a hypothesis but, as a matter of fact, I will tell you that the level of TPMT measured in the easily accessible tissue, the red cell, reflects the level of activity in the liver, in the kidney and in every tissue that has been examined to this point and, when we get to the molecular data, it will become clear why that is the case,

not always the case, but for this polymorphism is it.

So 90 percent of the population from a Northern European population, and Larry hinted at this, and the language in that labeling, I think, I think is interesting. It says, "population." Whose population? A Northern European population, because the population--and I know, you have to get my words down and you are going to have a devil of a time--a Northern European population has the trait of high-enzyme activity.

About 10 percent, or actually 12 percent, are heterozygous and have intermediate activity and this one lady down here had zero enzyme activity. That is exactly what the Hardy Weinberg theorem would predict for a single locus with alleles for high and low enzyme activity, allele frequencies of 94 and 6 percent.

Using very sophisticated techniques developed by a monk in a monastery in the Czech Republic using segregation analysis, we confirm that this is an inherited trait. We hadn't cloned anything. This was a time before anyone had cloned much of anything.

[Slide.]

This is a little more accurate picture of the way these drugs work and I think it comes back to the complexities that Larry was hinting at; that is, azathioprine is a prodrug that is converted to 6-mercaptopurine in vivo. It can be oxidized or methylated and 6-mercaptopurine is, itself, a prodrug that undergoes a series of metabolic activation steps to form 6-thioguanine nucleotides. Clearly, this activated metabolite is correlated, when measured in the red cell, once again, and this is mainly work that came from Sheffield, England and Lynn Leonard and John Lilliman using the UKAL, the United Kingdom Acute Leukemia trials, that this appears to correlate with toxicity but the question is why.

When I met Lynn Leonard, I suggested to her that maybe the kids who have--these were kids with ALL who have this pathway partially blocked pump more of the drug down here and they will have higher 6-thioguanine nucleotide levels and they may be the ones at risk for toxicity.

[Slide.]

Here is a very early paper. I think these are data we published in Lancet in 1999 showing the predicted inverse relationship between the

genetically determined level of the enzyme activity in the red cell which reflects the activity in other tissues and the 6-thioguanine nucleotide levels measured in the red cell, and these are the heterozygous kids having these higher levels.

[Slide.]

Much more striking were four patients, and these were data published, I think, in 1989 in Clinical Pharmacology and Therapeutics. These were patients who had profound myelosuppression. Others were up in the thousands of picamoles per 108 red cells that Lynn Leonard had and a group of controls. These are dermatologic patients treated with azathioprine.

Much of the toxicity, and this is going to interesting, has been reported in patients treated with azathioprine by dermatologists and gastroenterologists because, in preparation for this meeting, I think I went through every clinical report of toxicity that has come out. They are interesting and I will mention those to you in just a moment.

These people had life-threatening myelosuppression. They were hospitalized for weeks and some of them for months. Many of the cases of

fatality were, in general, in these people who had zero enzyme activity. Now, that is interested because Larry asked the question, gee; is one in 300 important. The answer is it depends. It depends. It depends on how severe the toxicity is. It depends on the therapeutic index of the drug. It depends on the risk-benefit ratio which I think is what we were supposed to talk--so the answer will be different for different drugs and for different indications.

There won't be one answer and the Taliban would be disappointed but I am afraid there is no easy path to truth.

[Slide.]

Having said that, here is a publication that appeared in The Lancet in the early 1990s after we had published these data. This is a heart-transplant patient being treated with azathioprine. Here is the dose of the drug. Here is the white count. It goes down. The drug is stopped.

This is a German patient. The white count goes up. The drug is started again. The white count goes down to zero. The drug is stopped, started again here. The patient expired here with

massive sepsis. I have met this transplant surgeon. He won't transplant anyone, and won't treat with azathioprine, without measuring TPMT first after this rather devastating experience.

So this is, once again, azathioprine. When I go back and I look through all those clinical reports, what I find are two kinds. Number one, anecdotal case reports that are like this. They are dramatic and they are striking and the endpoint is such that when the physicians have been involved, I will tell you what their answers to the question is. That is not scientific. That is anecdotal.

The other is because of tie-ins with the fact that there are large-scale clinical trials of 6-mercaptopurine in the treatment of acute lymphoblastic leukemia and the results have been pretty much the same.

It is to the point, now, these kinds of cases are not reported. If you go back, when did they peak, and you plot them, it was in the early '90's. Then they went down. For two reasons. Number one, because they had been reported already. Number two, because of fear of litigation.

No one will publish these cases because

what if they were asked, "Could you have sent a blood sample to," fill in the blank, "and determined ahead of time that this might have been exquisitely sensitive to the drug?"

I have talked to the physicians. It still happens. I get the calls. Dr. McCleod gets the calls. I hope Mary Relling is there. She gets the calls. But nobody--and we need to be realistic here, so part of, I hope, what we are doing is facing the realities. This is such a dramatic example that the reality is that nobody will report this kind of case anymore.

They are built into the ALL trials, the NOFO trials and Howard can tell me about what goes in the United States because, as I said, I am just a poor old internist. I am not an oncologist. I am just a clinical pharmacologist.

[Slide.]

So what are the data? If you review all of those cases, what do they really say? If you have genetically very low--that is the 1 in 300 among Caucasians from Northern Europe--TPMT, you are at greatly increased risk of thiopurine toxicity. If Mary is not involved, I am really sorry because a lot of those data really came out

of the St. Jude studies.

It was, I think, 1991 that Bill Evans reported a case report of a child with ALL. I think that was the first of those kinds of cases that was reported. It is the St. Jude's group who has demonstrated that about one-tenth to one-fifteenth the standard dose will give you therapeutic efficacy without a dramatic increase in toxicity in these kids.

Mary, I think, was the first to report increased risk for secondary neoplasm in these kids. That is, we now cure this disease in 80-plus percent of these children but that means that they can develop a secondary neoplasm. She found that low or intermediate TPMT is a risk factor for secondary neoplasm. The Nordic Leukemia trials with Dr. Schmiegelo as the primary principal investigator in the big trials appears to confirm that.

We have reported, with Lynn Leonard and there are a lot of other reports, less compelling evidence for decreased therapeutic efficacy at high TPMT, but there are data out there less compelling than this. These are pretty compelling data.

[Slide.]

Having said that, what made a lot of this possible. It was having what I have called an intermediate phenotype, or you can use the term surrogate or what have you; that is, the 6-thioguanine nucleotide levels and the collaboration with Lynn Leonard that made--because there are a lot of reasons why people with these diseases develop myelosuppression. TPMT deficiency is only one of them, but it is now one that we now potentially are in a position to understand, to predict and to prevent.

So no one has ever claimed that low TPMT is the only cause for myelosuppression in children with leukemia treated with this cocktail of cytotoxic drugs. Number two, the ability to associate these kinds of studies with ongoing, very expensive but well-organized clinical trials. There is virtually not a child with ALL in the United States who is not on some sort of a protocol, and having the ability to connect with those trials.

The area with narrow therapeutic indices are within the area of cardiovascular drugs and the area of antineoplastic drugs, among others. AIDS is going to be another area. Being able to

associate these kinds of studies with ongoing clinical trials has clearly helped to develop the evidence base that enables us to be having this discussion today.

[Slide.]

Here is my definition of pharmacogenomics. As someone who has been doing pharmacogenetics for thirty years and using techniques at first that Mendel would have recognized, it is the convergence of those kids of pharmacogenetic advances irrespective of whether they deal with drug metabolizing enzymes or transporters or receptors, with the dramatic changes that have occurred in human genomics which have speeded the process up and have developed technologies which mean that the issue of genotype or phenotype, it is going to be much cheaper, the genotype, than the phenotype.

But there are going to be some problems and we need to talk about those before we are done and so we will.

[Slide.]

Here is the gene. It is easy for me to put the up now. Now you just type NCBI into your web browser and you go look at it. It was about a year and a half out of the life of Diane Otterness

and Carol Szernlansky in my lab in 1996, we published this gene structure. I won't bore you with the CDNA which took a year and a half out of a guy named Ron Honchell's life--Ron is at the FDA now--to get the CDNA. That is so old-fashioned, paleolithic; right? It was five or six years ago.

So the gene is 34,000 nucleotides long. It is on the short arm of chromosome 6. There is a process pseudogene in humans which really screwed things up but we won't worry about that right now.

[Slide.]

So, with that information available, Bill Evans' lab and our lab, within six months of each other, published the underlying genetic basis for the common polymorphism in Caucasians.

It is called Star 3A. That is because Bill had published a Star 2 variant that is less common. It has two non-synonymous c-snips which, translated into English, means changes in single nucleotides that change the encoded amino acid. I see Roberto Guercelini laughing. When Roberto was a post-doc in my lab, he used to bring a tape recorder in and record our conversations and he said he was going to play them back at half speed to try and figure what the heck I had said.

I think I got that right, didn't I, Roberto? So here we have two non-synonymous c-snips, one in exon 7 and one in exon 10. This variant has an allele frequency of about 5 percent in Caucasians. It is common. One out of every 20 copies of this gene in Caucasians is this variant. That allele has never been seen in anyone from Han Chinese, Korean or Japanese.

You can get the exon 10 variant and allele facility, Howard, of 1 to 2 percent. Would you agree with that--which is a little higher than what you find that variant in Caucasians. But this one, I don't think, has ever really been reported in anyone who, like my wife, would say that they are truly a Han Chinese. We collaborate with some people in China. They are confirming data that Howard published several years ago when he was in Scotland.

So this is the underlying basis for high, low or intermediate. But let's kind of bear that in mind because what I am going to tell you is that there are a whole bunch of other variants that are much less frequent. If you are doing a DNA-based test, then they also are associated with low enzyme activity and at what level do you feel comfortable,

Larry, with accepting that.

[Slide.]

I also bring up the nasty word "haplotype" because TPMT is a great example for haplotype meaning all of the variants that are found up and down an allele--that is, this is the most common variant in Caucasians. This is the most common variant in Asians and it is found in Caucasians, not quite at the allele frequency found in Asians.

Bill and I used to argue about whether this one, the Star 3B existed. I think he now accepts that it does but at a very low frequency.

If we have a kid who is a compound heterozygote for a Star 3B and a Star 3C, they are going to have low levels of enzyme activity. That is very, very unusual among Caucasians. It actually may be more frequent among other populations. Howard, I have seen some data that indicate that.

That is quite different than the therapeutic implications of what would give you most commonly this snip and this snip in heterozygous which would be one wild-type allele and one allele like this.

Oh, my gosh; DNA is not the answer to

everything, says the fellow who has been using DNA for twenty--that is, it is going to get more complicated unless our friends from biotech can come up with absolute ways to get us haplotype down approximately to the 10 kb that separate these two snips. If you want to talk about that in detail, we can. That is a much more practical issue of haplotype than the kind of issues that Howard and I sat in another windowless room in Montgomery County not long ago watching multiple haplotypes as a way to actually get at function.

This is a real practical issue and we are going to have to think about it.

[Slide.]

This is just to make the point about ethnic differences. This is data from a Korean hematologist-oncologist, Dr. Parkash. She published this is Clinical Pharmacology and Therapeutics about ten years ago, 300 Korean kids. She got this nice Gaussian distribution without anybody here and without anybody down here. That is, in general, the kind of data that you were seeing, I think, too, and that has been reported repetitively and that our Chinese collaborators are seeing in Canton when they look at a series of

ethnic groups in China.

So the labeling is going to be an interesting issue, and how you approach the labeling, how all of us jointly approach the labeling--I use the royal "You" is going to be interesting.

This is just to remind you what that Caucasian frequency distribution looks like, but there is another point here. From here to here, within this homozygous high, these are people who, within the open reading frame, have the same sequence, you have got just as much range of activity as you do from here to here.

Does that make any difference and why is that? One of the reasons has to do--so we are used to allelic heterogeneity and ethnic variation in allele frequency, but there is a variable number tandem repeat that is GC-rich repeats. This gene, like most of the methyl- and sulfo-tranferases that we study doesn't have a top box, but it has got this GC-rich area with 17 to 18 base pairs repeated from three to nine times. The higher the number of these repeats, the lower the level of enzyme activity.

So not everything is a nonsynonymous

c-snip, so you can modulate activity and, yes, when we can afford to look at the entrons, then we are going to find that there will be some really interesting stuff there, too.

So the current level of technology will probably tell us, most of the time, who is going to be high and low or intermediate. It will miss some of them. Howard may have a different opinion on that, but it will miss some of them. The percentage is fairly low. And there will be no right answer to that question. It depends. It depends on how important it is to them.

[Slide.]

This is just to show you that, in a population study we did--this is 1100 samples from Mayo Clinical Laboratory. We phenotype and do about 5,000 to 6,000 of those a year, about half on our own patients, half that come in from outside. There are commercial labs that do the genotyping. The higher the number of repeats, the lower level of enzyme activity. A French group first reported this and deserves credit for it.

[Slide.]

So, to sort of finish--we will finish kind of where Larry left us; that is, the drug

metabolizing enzymes and probably TPMT and D26 are the oldest, best-developed, examples, have served to demonstrate the basic principles. TPMT is dramatic because the therapeutic index is so narrow and the consequences, and there are many examples like that example I showed you from the heart-transplant patient, of death when this hasn't been recognized in patients because the consequences are dramatic.

So it helps to illustrate a series of points and they are good demonstration projects that will help to develop principles that, hopefully, will apply more widely.

[Slide.]

These drugs--I mean, it is fascinating. It is too bad George Hitchings and Gertrude Ellion are now gone. They were wonderful people and I think it is wonderful that they were recognized with Jim Black for their contributions in drug development and how important that is.

[Slide.]

I don't think that Dr. Remy, who, as I say, is retired from the Department of Biochemistry at Bowman Gray--I sat in his living room a couple of years ago because I go down there fairly often

having a two-year-old grandchild there--so I get down there often.

I sat in his living room having a cup of coffee, and I said, "Why did you look at this enzyme in rats and mice?" He said, "Because George Hitchings told me it might be interesting." He said, "Does anybody really care?" So it is nice to be able to tell him that what he did in 1963 people are still quoting and paying attention to.

I would be happy to answer any questions, have clarification or corrections with this august group, and I know a lot of the people around the table. I am used to corrections, not quite as many as I get from the Mayo medical students, but I would be happy to deal with any questions or corrections.

Thank you for having me.

DR. JUSKO: Are there any questions for Dr. Weinshilboum?

DR. LESKO: The comment about the number of tests being done at Mayo, 5,000 or 6,000 per year, let's say, over the course of years, is there any way that data could be looked at to answer the question of clinical impact that the testing has had prior to and after--I know there is a common

denominator of how much drug is being used, but it is possible to look into the data to say that it has had or hasn't had a clinical impact and what the level of evidence to address that might be?

DR. WEINSHILBOUM: As long as the committee understands that what they are hearing is anecdotal, idiosyncratic and one person's impression, I will be happy--the test has been available as a standard clinical test for phenotype. I was trying to make the point, this is the case where you have got both phenotype and genotype tests available and I notice that the proposed labeling said one or the other, think about this.

The tests have been available since 1991 as a standard clinical test. By the way, I have no personal financial interest in that test in any way, shape or form. I own not a single share of any pharmaceutical or biotech testimony. The Mayo Clinic is a highly socialist organization, Scandinavian Americans, so that when I do consult for drug companies and biotech companies, the consulting fee goes back to help us achieve our institutional missions and research and education.

Having said that, then--I mean, I think is

important to say those sorts of things. Having said that, then, the test has grown from a few years ago, I would said, 1,000, 1,500 tests. It has grown dramatically. The greatest single growth has not been in the ALL area. That, thank god, although it is the most common neoplasm of childhood in the United States, is a relatively small part of the use of these drugs.

Gastroenterology is the biggest part. The growth has been in gastroenterology, dermatology and in a variety of autoimmune diseases, in our practice, the gastroenterologist being the biggest.

We see something like, I think, 1,500 new cases of Crohn's disease, new cases, per year, so these are kids who are being started--and they are generally teenagers who are being started on these drugs. These drugs are at the mainstay.

The impact, in that area as opposed to the relatively small and stable group of ALL patients--and I don't mean to downplay that. I just think we need to put this in context--is that our gastroenterologists in the Crohn's disease clinic in one academic referral center are generally doing the testing at the front end because they are so concerned about the relatively

rapid development of profound myelosuppression in the 1 in 300.

If you are seeing something like 1,200 of these kids a year, then it become a few patients each year. We do see referrals, and I don't want to violate any patient confidentiality issues, referrals from outside who require prolonged hospitalizations because of profound myelosuppression, not having recognized this problem.

I realize that, in general, the resistance, and I speak as a clinician now, the idea is, gee, are you saying that we are not taking good care of our patients or watching them. Of course, no one is saying that. It is just that this new information has come along. We now understand this variation in response to the drugs and the question is at what point does the cost-benefit ratio become acceptable.

I firmly believe the answer is it differs, it varies, for saying at this point we will test everyone. Our gastroenterologists, and once again, I am speaking for someone else, it is my impression that they test everyone at the front end.

The other issue is the issue of following

the course of therapy. One could have prolonged discussions and they relate to clinical practice rather than what the labeling will be, about following the 6-thioguanine nucleotide levels with regard to how is the patient responding.

I think that is a different issue but I think we need to put it on the table. Finally, we need to realize that there are going to be practical clinical issues that arise if you wait because many of the patients we see where folks have waited, they are profoundly myelosuppressed. They have now been multiply transfused. We can't do the phenotypic tests.

Even the DNA tests get confounded by what they have received in order to treat the problem and there the genotypic test using buckle smears is one of the things we commonly are called on to deal with.

Now, I hope Mary is there. Is Mary there? If not, I will turn to Howard because Howard was at St. Jude when I first met him. He has been involved right from the beginning with story and I certainly want to give Howard a chance to amplify or correct any misconceptions I might have conveyed. I look upon this as a dialogue where we

are all trying to learn together in this brave new world.

Howard, any comments or corrections?

DR. McCLEOD: I think there are more than Norwegians in Mayo Clinic. I should say that from the start. You talk about Norwegians. There are also quite a lot of other ethnic groups up there now.

DR. WEINSHILBOUM: There are. Weinshilboum, for one.

DR. McCLEOD: One of the things that has become very clear is that this is not an ALL boutique. The data that is most solid, from Mary Relling and others at St. Jude, for what you would actually do with the genotype comes from the ALL literature. But the most common use, overwhelmingly, is the rheumatologist, the dermatologist and the gastroenterologist.

Unfortunately, those are three clinical groups that are not as good as others at managing acute toxicity. I say that as a general observation rather than a personal implication to anyone. The hematologists-oncologists are used to people crashing and salvaging them. So when they hear about this sort of thing, if it is not part of

their practice, they often say, oh, well; we are doing okay now.

Talking to a lot of patients, things that we don't really worry about like anemia and neutropenia, do affect the quality of life quite a lot. But, as per this morning's discussion, how do you put a number on a decreased quality of life in terms of Jrgen's analyses and these other approaches.

A lot of things that are affected by, for example, the 10 percent of the patients, the heterozygotes, that get toxic but don't die, a lot of the things that affect them are hard to put a number on. So, how do you go and make these analyses to make firmer studies.

The other component that you mentioned is that there is not the infrastructure in this nation to go out and do pharmacovigilence in a way you could in some other nations. So the quantitative longitudinal data for the implications of this testing is very hard to come by.

Some of the Scandinavian groups are starting to think about this and, hopefully, we will get data from them about how you take an entire nation's population and apply this in terms

of the context of this drug use.

So we are left with less than adequate data on the efficacy side of TPMT genotyping and extremely convincing data on the toxicity side for TPMT. So, the number of diagnoses that have been made at autopsy is far too high and, from a safety standpoint, the drugs that have been recently pulled off the market from toxicity, the frequency of toxicities were much rarer than as seen with TPMT.

So, if you look at it as an example, compared to the more recent drugs, this drug would be long. So I think those are just kind of some scattered thoughts to follow up some of the things you have already said.

DR. WEINSHILBOUM: While Howard was speaking, I would like to follow up on one other thing that Larry said. The implication was that pharmacogenetics and pharmacogenomics is "easier" than disease diagnosis from a confidentiality, sensitivity-of-the-patient, issue. And, of course, that is true.

The problem is that, in this example, TPMT is ubiquitously expressed in human tissue. It goes back through evolution to bacteria. That is where

Remy, one of the places, he first described it. We don't have any idea what the natural substrate or substrates is or are, if they exist, other than xenobiotics.

But most of the drug metabolizing enzymes, so that I could talk about catechol-O-methyl-transferase, which has common genetic polymorphism and, of course, it metabolizes L-dopa and methyl-dopa, but it is rumored that it metabolized--my old mentor, Julius Axelrod received the Nobel prize, in part, because he showed that it metabolizes endogenous catecholamines and there are data that it is a risk factor for a variety of diseases.

The genetic polymorphism, which we described twenty-five years ago, is a risk factor for breast cancer and it is a risk factor, according to recent data from the NIH, for schizophrenia. The fact of the matter is, the enzymes, the proteins, will not sit still for artificial definitions, that they just deal with chemicals that are manufactured by the pharmaceutical industry or come in from the environment.

TPMT, we eventually figure out what it is,

what it "does," and maybe we won't. But, as a matter of fact, that is probably going to be the exception, that the vast majority of xenobiotic biotransforming enzymes will also biotransform endogenous compounds and we cannot assume that, because we have a test for, fill in the blank with your favorite phase I or, in my case, phase II, enzymes, that they will not represent risk factors for human disease.

So I think that these nice boxes that we arbitrarily, because of the way we organize things, put things into, biology will refuse to sit still for that. You may have a different view, Howard, once again.

DR. JUSKO: We have the opportunity for comments from our people listening on the telephone.

DR. McCLEOD: Oh; wonderful.

DR. RELLING: Larry, hi. Can you hear me? I don't know that I have anything to add. I have been looking over the product labeling for the mercaptopurine, and it is surprising for me that there are things listed, potential warnings, as to having at this age--for example, renal (inaudible), which actually seems to have very little data

whatsoever to support it whereas we now have probably something like thirty or fifty high-quality applications indicating that TPMT status is definitely associated with toxicity, and there is no information in the prescribing as to how to handle that for assessing patients.

So I am having trouble understanding why pharmacogenetics is being treated so different than others for risk factors and variability (inaudible).

DR. JUSKO: Thank you, Mary. Your conversation was broken up slightly but I think we got the gist of it. Wolfgang? [No response.] This is no Wolfgang.

Are there any other comments on this TPMT, in particular, before we move to the general questions?

DR. WEINSHILBOUM: I want to apologize. I will have copies for the committee of all of my slides and they will be made available to you electronically. But I was building a doll house for a newborn as of last night.

DR. VENITZ: Can I ask you a question before you leave? You mentioned some discrepancies between the phenotype and the genotype. Can you

elaborate on that? What is the frequency?

DR. WEINSHILBOUM: Actually, the only point I was trying to make was that if we just genotype for what we know today, we will--and Howard, I think, has published as good data as are out there on a population basis, we still are left with a certain number of individuals and we probably could debate on that for a prolonged period of time where the phenotype, which will be lower intermediate activity, won't match the genotypes that we know today.

Howard, I think your estimates are about 95 percent and I will let you speak for yourself of the phenotypic low-activity samples that would be picked up that way. I will have to say that, in a study we did, of 2,609 consecutive clinical samples from individuals, it was closer to 10 percent that the phenotype, by which we mean intermediate or low activity, we could find no currently understood genetic polymorphism or other DNA-based sequence information to explain that.

Howard, you do have very good data.

DR. McCLEOD: In the review articles, we have tried to put 85 to 95 percent. Sometimes, the 85 falls off, but the real answer is that it is

somewhere around 95 percent of the variants that are out there can be detected by these three main polymorphism.

Some of the additional ones--there are at least eight, or nine, excuse me, published and there will be additional ones that will be found over the years, very rare singleton type variants.

Another important point on that is, if you looked at the right side of Dick's histogram for the population there, the 90 percent of the population that were wild type had a lot of variability. Some of that variability will be explained by other variants that are found, or the NTR in the promoter region or whatever you might at the DNA level, and there will be some variability that will not have a genomic explanation. It will be dietary influences or whatever you want to come up with.

Dick made this point already, but DNA will not be everything for any aspect of pharmacology much less TPMT.

DR. JUSKO: Maybe I could pose a question that Larry brought up as one of his issues. Dick, you indicated that it has been found that one-tenth to one-fifteenth of the standard dose works well in

children with ALL. Is that also the case, also the experience, of rheumatologists and dermatologists, GI people, in the use of these drugs in patients with the other indications?

DR. WEINSHILBOUM: That is a fascinating question. I think, when I said that, I said that the best data with regard to ALL were the data that Mary Relling and Bill Evans have developed at St. Jude. They were the ones who really, I think, were in a position to develop those data.

Our gastroenterologists at Mayo, because they are big-time users, feel that the drug is frequently used with aminosalicylates which inhibit TPMT and that complicates life, so we are going to have all the complications. I am just reiterating what Howard said.

He implied that there is some evidence "of induction." I am not using that in the NIH study-section terms but of increase in level of enzyme activity in patients who are treated chronically with these and other drugs. There is evidence of drug-drug interactions at the level of inhibition of TPMT and then, on top of that--so life is not going to be simple here--but, on top of that, then we have the issue of what is the

appropriate dose in other diseases.

I think Howard, in his comments, was, perhaps, a bit harsher than I might be in dealing with our gastroenterologic and dermatologic colleagues in that I don't believe that the data are out there which are as compelling as the data from St. Jude with regard to ALL about how to approach the balance between efficacy and toxicity in these other disease states.

Howard, once again, you may have a different point of view.

DR. McCLEOD: I agree with you. I think that there are some people who go to the one-tenth of the dose and titrate up based on toxicity. There are some people that just stop using thiopurines and go to a second-line agent. There are some people that do a combination, depending on the day of the week.

But, what there isn't, is good cohort data of the type that Mary Relling has published from St. Jude. That is what is missing, is these large cohorts where people were uniformly treated and managed so that we can actually have more definitive answers outside of childhood ALL.

DR. LESKO: Actually, I had two questions.

The first question is is the one-tenth of dose based upon exposure to 6-thioguanine or is it based upon a proportional reduction in TPMT activity? What is the basis for the one-tenth of dose recommendation.

Secondly, if you were to think about patients that are referred because of toxicity, or at least suspected toxicity, to 6MP, what percent of those patients are, in fact, poor TPMT genotypes? Do we know that?

DR. McCLEOD: Mary, do you want to take that one because you have the most recent breadth of experience?

DR. RELLING: Can you hear me okay? I hear a crazy echo.

DR. JUSKO: Yes, Mary. We can hear you.

MS. REEDY: If you are on speaker phone, if you will turn that off and use the hand-set, you will get less echo.

DR. RELLING: I am not on a speaker phone. What was the first part of the question? I'm sorry?

DR. LESKO: The first part of the question, Mary, was is the one-tenth of dose based on blood levels of 6 thioguanine?

DR. RELLING: Yes.

DR. LESKO: Or is it based upon something else?

DR. RELLING: The one-tenth of the dose was based on clinical tolerance. Our policy was to use the TPMT status to determine whether 6-mercaptopurine was the culprit drug or not. Once we determined that 6-mercaptopurine was likely the culprit drug based on low TPMT activity.

Then we titrated that dose to the peripheral white-blood-cell count as we would do in any other childhood leukemia. So, actually, the thioguanine nucleotide level still is extremely high in those patients. So I can't say that what we did was the correct thing to do because we do have some concerns that there may be secondary cancers in patients with those high thioguanine nucleotide levels even if they don't experience a lot of neutropenia from that.

So, we sort of disagree with the concept of a target thioguanine-nucleotide level because we don't believe that that has been established in ALL and I don't know if it has been established in any other diseases.

DR. WEINSHILBOUM: Mary, this is Dick

Weinshilboum. Dealing with our gastroenterologists, they would feel exactly--they would second what you just said with regard to the treatment of Crohn's disease. They are not certain that the same range of 6-thioguanine-nucleotide levels are appropriate for treating Crohn's disease as are appropriate in ALL. After all, the targets may be somewhat different and what is the appropriate surrogate marker or markers remains open to serious question and the best data, probably, that are out there are for ALL.

So I think that the questions that are being asked are exactly the right questions.

DR. RELLING: Right. To me, the best rationale in leukemia treatment is the fact the every drug we use is myelosuppressive. What TPMT does is help us focus in on the correct drug to adjust as the culprit for myelosuppression. That can't really be said in noncancer diseases, in general.

Then, I'm sorry; I don't know about the second part of your question.

DR. LESKO: The second part of the question had to do with patients that are referred because of suspected 6-MP toxicity. How many of

those, in fact, are confirmed to be poor TPMT-activity genotypes?

DR. RELLING: About two thirds, in that preselected group.

DR. LESKO: About two-thirds?

DR. RELLING: Yes; that is published in the Journal of Clinical Oncology last year. So those are very motivated clinicians. Those are clinicians who were suspicious of thiopurine methyl-transferase insufficiency and who were following their patients closely and who were motivated to enroll their patients on a protocol and send us samples.

Out of those samples that came, two thirds of them that had (inaudible) also had at least one mutant allele for TPMT. If we look the converse way, if we look at all (inaudible) of heterozygotes, which make up 10 percent of the population, only about 38 percent of them had toxicity that was severe enough to make us decrease their doses.

DR. LESKO: Mary, that last figure, was that--I was trying to get the patient population there. Is that patients in whom you didn't know the genotype in advance, but 38 percent of those

eventually required a lower dose? I wasn't clear on that last thing you said.

DR. RELLING: That's correct. So of the patients turned out to be TPMT heterozygotes about 35 percent of them required a dose decrease in order to keep their ANC in the target range. Now, that doesn't mean they perhaps would have benefitted from a dose that is decreased if only they lower their PGN level because what happens in that group, a huge percentage of them develop secondary tumors.

So our policy is to decrease the dose of TPMT moderately in all TPMT heterozygotes no matter what their tolerance. That, for us, means we give them 60 milligrams per meter squared instead of 75, or lower if they are having acute hematopoietic toxicity.

DR. JUSKO: Another general question that was posed earlier by Larry is how reliable and how available are the commercial tests to TPMT, for the several people that are using them.

DR. McCLEOD: I think that there are three different types of tests that are out there. There is this genotype test. There is the phenotype test measuring TPMT activity in red cells. And then

there is the endpoint test measuring the thioguanine nucleotides. There are commercially available tests for all three of those endpoints that are out there that are robust and that perform a CLIA-certified environment.

So, in terms of availability, they are available and they are robust. They are not widely available. One of the most common phenomenon that I find in this is people calling up wanting me to test in the research setting not realizing that there is a CLIA-certified laboratory that would perform the test.

Also, there are only a few one-stop shops for this, so there is at least one company that, I believe, does all three of the components. There are other institutions that just do the phenotyping, for example. A number of institutions have a home brew where they will do testing for their institution by not commercially outside the institution. So a lot of the larger academically minded institutions will do that sort of approach.

Mayo Clinical Laboratories, which is separate from Mayo Clinic, I understand, but the same place, offers the phenotyping test. Then there is a company in San Diego that offers the

genotyping and the thioguanine-nucleotide levels. Dick or Mary could elaborate on that if there are additional resources.

So it is available. It is not as well publicized as it could be.

DR. JUSKO: So, if a pediatric oncologist in Buffalo, New York wanted to test a patient, the test could be done in a relatively--with a fast turnaround someplace?

DR. McCLEOD: Yes.

DR. HALE: Could I get a little clarification on the test performance? Do we know about the false-positive and false-negative rates?

DR. WEINSHILBOUM: I can comment on the fact that our clinical lab, obviously, has those data. What we are really talking about with the genotype-phenotype correlation was an attempt to get at, with regard to genotyping, the potential for false-negatives; that is, we would miss patients whose phenotype--and it is an advantage, actually, to be able to compare those, at least at this stage in the development of the assays.

I quoted a figure, Howard quoted a figure, from one of the studies that he did which is an appropriately highly cited study. With regard to

the false positives, I think there are less data available because, in general, what we will do in our setting, and I use the royal "We" because I don't do this, I don't run a clinical lab and I am not CLIA approved for anything, is to go back and retest anyone who shows up as potentially being either heterozygous or homozygous low.

Mary may know a good deal more about what is done with the genotyping tests Of course, there are broad issues that relate to the technology platforms and the way in which the snip detection--right now, I think, Howard, we are talking just about snip detection. We are not talking about haplotype. Larry raised the issue. I think it is going to be an interesting one.

Committee Discussion

DR. JUSKO: I think it would be appropriate, at this point, to return to Larry's last slide, the general questions for the committee.

DR. LESKO: From the handout or from the computer.

DR. JUSKO: It is on another screen, so let's start with the handout.

DR. LESKO: It is on Page 16.

DR. JUSKO: The first question posed is what major findings would support the inclusion of a genetically tailored dosing regimen in a package insert.

DR. McCLEOD: I will kick it off, I guess. I think that there is already pretty clear evidence for the relationship between a homozygous variant genotype and toxicity. So, to me, for the toxicity evidence is just a robust correlation between a phenotype, such as toxicity, and a genotype or a measure of the enzyme variant.

So, to me, that data is already there. The data for the relationship between a heterozygote genotype or phenotype and toxicity is less well-developed. We did one study, a cohort study, a relatively small study of 67 rheumatoid-arthritis patients, and found that the heterozygous patients came off therapy quite acutely because of toxicity.

But that study has not really been duplicated outside of a single Japanese study that I am aware of that did evaluate that and, thankfully, did find the same types of results. So there is still more evidence needed to really define what the implication is for a heterozygous

genotype in the types of patients that commonly get thiopurine drugs.

So, Mary's study in the Journal of the National Cancer Institute in 1999 for childhood-leukemia patients was able to show, as she mentioned just a few minutes ago, that somewhere around 35 percent of patients with a heterozygous genotype required a significant dosage reduction. So we do have that evidence.

We don't know what the case is for gastroenterology patients, for rheumatic-disease patients or for the dermatologic diseases. One, one piece of missing evidence is for these other groups, which are the more common numerically, patients that are getting thiopurine drugs.

So one initial bit is the clear evidence that this genotype will give you severe toxicity 100 percent of the time, or the majority of time.

DR. WEINSHILBOUM: I guess I would agree with what Howard just said. For the homozygous-low individuals, the data are so compelling that no longer will those studies be published nor, as I think I implied, no longer will anyone even attempt to publish them for a variety of reasons that go beyond the scientific.

For the heterozygous individuals with ALL, I believe that Mary and the St. Jude experience have developed data which indicate that this is also an issue, toxicity. On the therapeutic-efficacy side, I hope I made this point, the data are less compelling. There are data out there and it may well be that as this august group deals with pharmacogenomics, that the more challenging issues and the broader area where pharmacogenomics potentially has implications is not necessarily this kind of demonstration project where we are looking at the toxicity end, but issues of individual variants and therapeutic efficacy.

I think those will be challenging times and I am looking forward to what you are going to recommend as you begin to move into those area because I think that is where the broadest application will apply.

Howard implied that these drugs probably, in today's world, might not stay on the market. But they certainly have proven useful in a variety of settings and thank god that they were placed on the market.

But, Howard, don't let me put words in

your mouth.

DR. McCLEOD: I think that is exactly right. If you look at, at least what I am aware of, of some of the drugs that have been hauled off the market fairly recently because of their toxicity profile, the number of patients with toxicity were much fewer than the number of patients that get toxicity from azathioprine or mercaptopurine.

It is a situation where if this had been a new drug introduced a few years ago, it may have come off for that very reason. There have been as many or more deaths from thiopurines that have been published, in addition to the unpublished ones, than the drugs that have come off the market recently.

So I think, if we look at that context--it is too bad that Lew Sheiner had to fly back because he had a mantra he was chanting throughout the morning of trying to look at what we are comparing this against.

If we are trying to look at an ideal world, we do not have enough data to say that TPMT genotyping, or any other genotyping for the most part, will let you tailor the exact dose for each

individual patient on both and efficacy and a toxicity basis.

But, in trying to make a drug safer, there is enough evidence that this genotyping will make drugs safer. One in 300 is not common unless, as Rick said, you are that one. If you are that one, then it is a little bit too common. As mentioned already, autopsy is terrible place to make the diagnosis.

DR. HALE: I would like to make a few comments about Larry's general question there. We have already hinted at the first one about the false-positive and false-negative rates and coming at this kind of from a statistical and utility approach that those can actually be very important when you look at because a false-positive rate, when you have got a rare event, even one in 300, you can wind up finding--in this case, even if you have a 1 percent false-positive rate, you can wind up three of your four positives turning out to be false positives which could deny therapy to people, or force them to alternate therapy.

We need to look at the cost, not only to people who get the drug that shouldn't get it, but also the cost of withholding the drug from people

who would benefit from it. So we are talking about utility.

Things like the speed, convenience, cost and reliability of the test all impact on its use and the fact that it is too cumbersome or too costly, it won't be used at all. On of the other things is actually the proportions. When one does the utility, you have to have the numbers--you have got the one in 300 here, the 10 percent. Those can impact broadly on whether it is a good risk-benefit thing or not from a population point of view and not just do we have a test. It is more or less from the population point of view, does it make sense for the population. So you really have to think about the population risk-benefit.

The other consideration that has occurred to me here is the difference between a demonstrated clinical benefit where you prospectively do this versus the post hoc analysis where you look at the people who have had these events and then you say, "Well, this was particular genotype." So have we prospectively done a study using this kind of screening.

DR. McCLEOD: Mary, if you can hear us, I wonder if you could comment on your data for

false-positive rate because you are in a situation where not only you are genotying but you are also phenotyping, so you would actually have that information, and also the last comment about whether--I am not aware of any prospectively randomized trials where people looked at genotype versus no genotype, either at the toxicity or efficacy area, but Mary Relling may have that data.

DR. RELLING: We have never (inaudible) and, as far as I know, no one else has of a false-positive phenotype. As Dr. Weinshilboum mentioned, there is a theoretical possibility for a heterozygote in some racial groups (inaudible) to distinguish from homozygous, but there are ways to get around that.

If we use phenotype only, we do see putative false positives so we see occasionally low red-cell TPMT activity which does not have mutation. So, in the absence of toxicity, then we generally retest phenotype, an independent sample, and usually activity is then normalized. There might be very rare cases where the activity remains low and we don't see much toxicity.

DR. WEINSHILBOUM: That comes back to the issue that I was raising earlier. You only know

what you know and there was a time we didn't know about Star 3A. Once you know about Star 2 and Star 3A, then you find Star 4 which is a spice-junction variant and Star 5 and Star 6 and Star 7.

DR. RELLING: Right.

DR. WEINSHILBOUM: So you learn to look further and further. The gene, itself, is 34,000 nucleotides in length. I don't think anyone sequences through the whole gene. So what is the definition of a false positive? I think you would have to go back and say, does the phenotype remain constant and, until we understand the functional implications of every change in the DNA, we aren't in a position to really answer the question.

So you have to define practically what you are doing. These are real-life issues that we are all going to be entering into as we begin to use DNA-based testing. But there is a difference, and the difference is--you raise an interesting question when you asked about the question of how difficult is the test.

Pharmacogenomics, rumors to the contrary, has been around for decades. It has been resolutely ignored for decades but it has been around for--the concepts have been there. The

major problem with 2D6 was that, prior to the time that we understood the DNA base-sequence variations, you had to use a test drug and my colleagues, in internal medicine and in psychiatry, would not do that, so that the practical reality--and I am just repeating what you said just a minute ago--was such that, unless you had a rapid turnaround, reasonably robust test, our clinical staffs, understandably, were dubious that the cost-benefit ratio was acceptable.

What has changed with the genotyping is that we now can, with a variety of technology platforms and so who cares which one it happens to be, it will be different tomorrow anyway, with some of the people sitting out in the audience, I hope, being responsible for that.

As the technology platforms mature, the DNA base testing gives you rapid turnaround and the ability to get the information back to the clinician quickly, hopefully validated in such a way that we can feel confident about what we do know.

I think that we need to be practically minded. Some of us, who have been using the word "pharmacogenetics," I will tell you when I came to

the FDA ten years ago and pharmacogenetics, everyone's palms got sweaty, their pupils dilated and they weren't very interested because it wasn't really a practical reality.

What the genomic revolution has done has been to make that a practical reality. That is where the technology changes have been different. You don't have to give debrisoqin and collect a twenty-four-hour urine or look at a plasma sample or even use caffeine as a probe. Now, once, again, Howard and Mary have a different take. That is part of the reason we are sitting around talking about this today. There is absolutely no doubt in my mind about that.

DR. JUSKO: On that note, perhaps we have resolved Question 1, stating what major findings would support the inclusion of a genetically tailored dosing regimen in the Package Insert. It sounds like, for TPMT, 6-mercaptopurine, there is considerable enthusiasm and considerable use of having these genetic tests available, although there are some scientific and clinical issues remaining to be resolved particularly what does one do with that information in terms of patients who might need to have far smaller doses than the rest

of the population.

In terms of trying to generalize this type of consideration, it seems very likely that it would need to be done on a case-by-case basis, much like Dr. Weinshilboum proposed, that one must do this with making what we discussed earlier today, risk-benefit considerations will depend on the drug and the types of toxicity and efficacy that is being considered.

Easier questions to deal with is the second one, where in the label should such information be placed? In the interest of time, I will concur with what Larry proposed for TPMT. The proposed labeling in that case seems to be very logical positioning of the information as well as the type of information.

Maybe in the last couple of minutes that we have left this afternoon, we can, perhaps, address briefly the third point, under what conditions should testing be optional or mandatory prior to dosing. Maybe we have addressed a lot of this already but perhaps someone with more expertise could comment on that.

DR. McCLEOD: The conditions for optional testing are obviously a lot easier to define than

mandatory testing. The problem with mandatory testing, even an example like thiopurine methyl transferase is that we have gotten by without it. When you talk to pediatric oncologists that want to bother getting TPMT testing, they just say, well, we just salvage the patients that crash.

While that is not a very user-friendly way forward, it is the reality in a lot of situations. So, making something mandatory has to have much clearer evidence that it is cost-effective in the true pharmacoeconomic sense of the word and a beneficial way to go forward.

There has only been one analysis of pharmacoeconomics in the TPMT example from Mayo Clinic and there needs to be a lot more. So, in terms of mandatory, I think, in the general sense, there needs to be evidence that you can either benefit from testing everyone or that you can select the best patients to test.

One of the things, I believe it was Larry, mentioned was that the patients that start having a fall in their white count then go forward to mandatory testing. That, I think, is a good idea. There is no information that I am aware of to select the trigger for that to be initiated, and so

that is something that would need to be worked out.

But that context of having patients declare themselves, at least in part, rather early while it is still--I hate to use the word "safe," but safe, would be one way forward to that.

Mandatory testing for TPMT in the absence of clear pharmacoeconomic analysis, I think is too early. We need the information about how much this would really cost. I know it is $300 an assay but we don't know how much we are saving by catching the 1 in 300. So that sort of information is needed before you can make that mandatory in my opinion.

DR. RELLING: I agree. I think that there would be tremendous skepticism and hesitation on the part, even of pediatric oncologists, to mandatory testing. I guess that emphasizes that the other therapy has a huge effect on one's ability to diagnose the myelosuppression but it also impacts on how 6MP is in the context of all the other therapies. I think it would be very difficult to write guidelines that would be a sufficient rationale for mandatory testing before treatment.

DR. WEINSHILBOUM: Mary, I would agree

with that. I do think that this group--I sit on the Council for one of the NIH Institutes. It is always amusing to me to hear them say, well, this isn't a mandatory policy. Of course, that is like an 800-pound gorilla crawling in bed with you and saying, "Don't worry; this isn't mandatory," or, "I am from the government; I am here to help you."

So, let's be realistic. If the labeling changes, even if it is not mandatory, the implications are significant and they will ripple through the clinical community. So, as long as we all understand that, I couldn't agree more with what you and Howard have said. I think it is premature to talk about mandatory testing, but there are practical implications to any labeling change which this group is more sensitive to than a basic clinical pharmacologist like myself.

DR. McCLEOD: The language that has been mentioned, that Larry presented, and a lot of it, I believe, had been--Larry, you included a lot of Mary's stuff in there as well?

DR. LESKO: There was some of Mary's stuff and some stuff from our internal discussions combined.

DR. McCLEOD: The nice thing about that

language is that, if nothing else, it increases awareness that it is a problem and that something can be done about. That, I don't think, is too much to ask. I think there is enough data to support that sort of thing.

The language, at least the way it was read today, was not gorilla-ish in terms of the way it was present. So, if nothing else, making people aware of this sort of issue in the labeling is necessary. There are people who, for some reason, haven't heard Mary or Dick speak on this topic. There aren't very many of them, but there are a few.

So that is necessary and there will be, I think, from a safety standpoint, although this is hard to document, there will be lives saved through this sort of inclusion in the labeling.

DR. WEINSHILBOUM: I couldn't agree more and I am enthusiastically supportive of the kind of mild informative language that Larry suggested. I just wanted to be certain that we were all aware of the implications of even moving that far which I think is probably timely for this particular example.

DR. JUSKO: I think we have had a very

enlightening discussion of this topic as well as the others. This point in our schedule calls for Larry to make some concluding remarks.

Concluding Remarks

DR. LESKO: That is always hard after about eight or nine hours of intellectual discussion, but let me conclude by simply saying thank you to everybody for their contributions today and, again, for accepting the challenge of being on this committee.

I would say the quality of today's discussion and the intellectual level met or far exceeded my expectations. I have been through about a hundred advisory committee meetings so far and this one was very enlightening and very helpful.

I think, for you, the members, as we act in information coming out of the committee, I am sure you will feel a sense of satisfaction that you have contributed to the advancement of drug development and regulatory decision-making. Our commitment is to move forward on these issues and to take the input you have given us and begin to organize ourselves to move forward.

When we see you all again in six or twelve

months, hopefully, six months, we hope to present new information on these topics and also we have this backlog of other topics we hope to bring to the committee along similar lines of what we talked about today.

So it was helpful for us. I hope it was fun for you and I think we are all hoping to move science forward for the betterment of patient care. So thank you, everybody for coming. And also to our guests who came up on birthdays and, some by some defective technology, we really appreciate all of that. Thanks a lot.

DR. JUSKO: On behalf of the committee, thank you for inviting us and thank you for being so well-prepared with useful information and bringing in outside experts that considerably enhance the ability to assess and discuss these topics.

[Whereupon, at 4:42 p.m., the meeting was adjourned.]

- - -