sgg

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

VOLUME I

Monday, November 14, 2005

8:30 a.m.

CDER Advisory Committee Conference Room
5630 Fishers Lane
Rockville, Maryland

MILLER REPORTING CO., INC.
735 8th STREET, S.E.
WASHINGTON, D.C. 20003-2802
(202) 546-6666

PARTICIPANTS

Mimi T. Phan, Pharm.D., Executive Secretary

ADVISORY COMMITTEE FOR PHARMACEUTICAL SCIENCE MEMBERS (Voting):

Carol Gloff, Ph.D.
Nozer Singpurwalla, Ph.D.

AD HOC SUBCOMMITTEE MEMBERS (VOTING):

David A. Flockhart, M.D., Ph.D.
(Chair - Topic 1A and 1B)
Jeffrey S. Barrett, Ph.D., FCP (Chair - Topic 2)
Michael D. Caldwell, M.D., Ph.D.
(Presentation only on Topic 1B/Non-voting)
Edmund V. Capparelli, Pharm.D.
David D'Argenio, Ph.D.
Marie Davidian, Ph.D.
Brian F. Gage, M.D., M.Sc.
(Presentation only on Topic 1B/Non-voting)
William J. Jusko, Ph.D. (Recused from Topic 1B)
Gregory L. Kearns, Pharm.D.
Howard L. McLeod, Pharm.D.
(Recused from Topic 1B)
Mary V. Relling, Pharm.D.
Wolfgang Sadee, Ph.D.

FDA STAFFS:

Lawrence Lesko, Ph.D.
Shiew-Mei Huang, Ph.D.
Yaning Wang, Ph.D.
Robert Powell, Pharm.D.

C O N T E N T S

Call to Order
David A. Flockhart, M.D., Ph.D. 5

Conflict of Interest Statement
Mimi Phan, Pharm.D. 7

Update on Previous Meeting Recommendations and Background to the Topics of this Meeting
Lawrence Lesko, Ph.D. 12

Topic 1: Translation of Pharmacogenomics (PGx) Information into Label Updates for Approved Products

Topic 1A: Evidence and Process for Translation of Pharmacogenetic Information (e.g., CYP2C9 Polymorphisms) into Label Updates for Approved Products

How New Insights into Pharmacogenetics Lead to Revision of Product Labels
Shiew-Mei Huang, Ph.D. 57

FDA Pharmacogenetic Labels: A Clinical Perspective
David A. Flockhart, M.D., Ph.D. 70

Topic 1B: Current Evidence Related to the Pharmacogenetics of Warfarin as a Potential Basis for Label Updates

Evidence Supporting Relabeling of Warfarin
Shiew-Mei Huang, Ph.D. 112

New Insights on Warfarin: How CYP2C9 and VKORC1 Information May Improve Benefit-Risk Ratio
Brian F. Gage, M.D., M.Sc. 122

Commentary on Current Status and Next Steps with Integrating PGx Information into Safe and Effective Prescribing of Warfarin
Michael D. Caldwell, M.D., Ph.D. 138

C O N T E N T S (Continued)

Open Public Hearing: (Topic 1A)
Sean X. Hu, Ph.D., MBA 151

Commentary on Current Status and Next Steps with Integrating PGx Information into Safe and Effective Prescribing of Warfarin
Howard L. McLeod, Pharm.D. 153

Committee Discussion and Questions 158

Topic 2: A Critical Path Pilot Project in Pharmacometrics (Quantitative Methods)

Call to Order
Jeffrey S. Barrett, Ph.D., FCP 222

FDA Experience with End of Phase IIa Meetings: An Attempt to Improve Drug Develop Decisions
Robert Powell, Pharm.D. 224

Committee Questions to Speaker 248

Case Study: A Quantitative Approach to Assess a Genomic Design and Biomarker Titration Design for a Phase III Clinical Study
Yaning Wang, Ph.D. 260

Committee Questions to Speaker 274

Commentary on the Case Study,
Jeffrey S. Barrett, Ph.D., FCP 293

Committee Discussion and Questions 299

Wrap-up of Day 1
Lawrence Lesko, Ph.D. 325

P R O C E E D I N G S

Call to Order

DR. FLOCKHART: Good morning. I would like to call the meeting to order. I am Dave Flockhart, from Indiana University. Welcome, everybody, to this Clinical Pharmacology Subcommittee of the Advisory Committee for Pharmaceutical Science. I would like, if we could at first, just for the record, to go around the committee members who are present, if you could just state your name and your institution, folks, and then we will get on with the meeting.

DR. D'ARGENIO: David D'Argenio, University of Southern California.

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

DR. SADEE: Wolfgang Sadee, Ohio State University.

DR. SINGPURWALLA: Nozer Singpurwalla, George Washington University.

DR. JUSKO: William Jusko, New York University at Buffalo.

DR. GAGE: Brian Gage, Washington University in St. Louis.

DR. CALDWELL: Michael Caldwell, Marshfield Clinic.

DR. MCLEOD: Howard McLeod, Washington University, St. Louis.

DR. BARRETT: Jeff Barrett, University of Pennsylvania and the Children's Hospital, Philadelphia.

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

DR. GLOFF: Carol Gloff, Boston University and independent consultant.

DR. DAVIDIAN: Marie Davidian, North Carolina State University.

DR. POWELL: Bob Powell, FDA.

DR. LESKO: Larry Lesko, clinical pharmacology at FDA.

DR. HUANG: Shiew-Mei Huang, clinical pharmacology, CDER, FDA.

DR. FLOCKHART: Thanks, everyone. I would like to have Mimi Phan read the conflict of interest--sorry, I have to state my own name. I am Dave Flockhart, from Indiana University.

DR. PHAN: Mimi Phan, executive secretary.

Conflict of Interest Statement

DR. PHAN: The conflict of interest statement for the meeting of Clinical Pharmacology Subcommittee of the Advisory Committee for Pharmaceutical Science today, November 14, 2005, the Food and Drug Administration have

prepared general matters waivers for the following special government employees who are participating in today's meeting of the Clinical Pharmacology Subcommittee meeting of the Advisory Committee for the Pharmaceutical Science regarding topic 1A, to discuss and provide comments on the evidence and process for translation of pharmacogenetic information, example given, CYP2C9 polymorphisms, into label updates for approved products, and topic 2, to discuss and provide comments on the Critical Path pilot project at the end of Phase IIa meeting which will include a case study. This meeting is being held at the Center for Drug Evaluation and Research. Waivers for Dr. Nozer Singpurwalla, Dr. Jeffrey Barrett, Dr. Edmund Capparelli, Dr. David D'Argenio, Dr. Marie Davidian, Dr. David Flockhart, Dr. William Jusko, Dr. Gregory Kearns, Dr. Howard McLeod, Dr. Mary Relling, Dr. Wolfgang Sadee, Dr. Michael Caldwell, Dr. Brian Gage and Dr. Carol Gloff.

Unlike issues before a committee in which a particular product is discussed, issues of broader applicability, such as the topic of today's meeting, involve many industrial sponsors and academic institutions. The committee members have been screened for their financial interests as they may apply to the general topic at hand. Because general topics impact so many institutions, it is not practical to recite all potential

conflicts of interest as they may apply to each member. 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 an FDA participant has a financial interest, the participants involved in the 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 firms whose products they may wish to comment upon.

Now I am going to read the conflict of interest for topic 1B, which will be current evidence related to the pharmacogenetics of warfarin as a potential basis for label updates: The following announcement addresses the issue of conflicts of interest and is made part of the record to preclude even the appearance of such at this meeting.

Based on the submitted agenda and all financial interests reported by the committee participants, it has been determined that all interests in firms regulated by the Center for Drug Evaluation and Research present no

potential for an appearance of a conflict of interest in this meeting, with the following exceptions.

In accordance with 18 USC Section 208(b)(3), a full waiver has been granted to Dr. Edmund Capparelli for unrelated consulting on the data safety monitoring board for competing firms. He receives less than $10,001 per year.

A limited waiver has been granted to Dr. Howard McLeod for a related grant to his employer which is federally funded for greater than $300,000 per year. Dr. Brian Gage, for a related grant to his employer which is federally funded for greater than $300,000 per year; unrelated consulting for a competing receive he receives less than $100,001 per year; and a related grant to his employer which is federally funded for less than $100,000 per year. Dr. Michael Caldwell, for his employer's negotiations for a federally funded related study in which the drug would be supplied by one of the sponsors of warfarin. The federal funding is proposed at between $100,001 and $300,000 per year.

These limited waivers will allow drs. McLeod, Gage and Caldwell to give their presentations and answer questions directly related to their presentation. They will be excluded from participating in the discussion, deliberation and voting.

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

Dr. William Jusko has been recused from participating in this portion of the meeting. In the event that discussions involve any other products or firms not already on the agenda for which FDA participants have financial interests, the participants are aware of the need to exclude themselves from such 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 firms whose products they wish to comment upon.

DR. FLOCKHART: Thanks, Mimi. I am glad we all survived that! Dr. Lesko, I would like to invite you to come and start us off. The subject of this morning's deliberations is translation of pharmacogenomic information into labels and Larry is going to give us an update on previous meeting recommendations and background to the topics of this meeting. Larry?

Update on Previous Meeting Recommendations and Background to the Topics of this Meeting

DR. LESKO: Thank you, David. Good morning, everyone and welcome to our advisory committee and thank you for coming today.

What I am going to do today is set the stage for the morning, in particular, with an introduction to the advisory committee and then an introduction to the topic.

[Slide]

What you will hear this morning is some discussion of general labeling with regard to pharmacogenomics. You will also hear a specific story about the evidence that underlies the warfarin polymorphism with respect to 2C9 and the evidence related to the labeling, and that is what will be the gist of most of the discussion that we will have this morning.

[Slide]

This is the fifth meeting in a series of meetings for the clinical pharmacology subcommittee. As you can see, we started in November of 2002 and progressed on a regular basis, in some cases having two meetings per year.

[Slide]

The meetings that we have had of the CPSC have basically focused in three broad areas. The first is quantitative methods. We have discussed in front of this committee modeling and simulation approaches that are intended to optimize dosing, in particular dosing

adjustments that reduce the risk of adverse events in patient subgroups.

We have also had a number of discussions on pharmacogenomics and, in particular, we have discussed label revisions of thiopurines and irinotecan to include genomic data that can help a physician guide the dosing of the drugs.

Finally, in past meetings we have talked about drug interactions and evaluation of drug interactions with regard to labeling and the evaluation of enzyme transporter mechanisms in anticipation of a revision of our in vitro and in vivo guidance on drug interactions.

[Slide]

We have also used this committee to discuss critical path initiatives. Critical path is the subject of a document that the FDA released, called "Innovation and Stagnation," and we have, in particular, discussed two broad areas. One as the end-of-phase-2A meetings. In fact, in this meeting you will be hearing an update on those meetings. We also discussed last year a framework for biomarker evaluation and you will be hearing an update as well at this meeting.

In the critical path there is a segment that talks about opportunities looking at biomarkers to target responders, monitor clinical response and measure drug

effectiveness. In many ways, this meeting today and tomorrow is about biomarkers, whether they be genomic or non-genomic, and whether this information can be effectively translated into labels for improving the benefit-risk of drugs and for the information of physicians and patients.

[Slide]

What we have planned for this meeting is a discussion of pharmacogenomic data in product labels. We will be asking for advice on the best way to include this information in product labels to make it as informative as possible, and we will be specifically talking about evidence for including genomic data in the warfarin label. What we will not be talking about today with regard to warfarin is the specific language in the label or where the information might be provided in the label. What we will be speaking about is primarily the evidence and we will be asking you to comment on that evidence as a potential basis for future label revision.

Secondly, we will talk about model-based drug development. This is a critical path initiative in which we will recap experience with our end-of-phase-2A meeting and we have a specific case study that represents the types of things we do in a 2A meetings. It involves

stratification and clinical trial simulation and we will be looking for input on the way we approach that case.

Finally, we will be talking about biomarkers and individualization. We will be updating on the critical path initiative as it relates to biomarkers, and we will be talking in general about the process by which biomarker data, generated during drug development, can be effectively translated into product labels and when that ought to occur.

[Slide]

So, let me begin with this slide which talks about drug labeling since that is the gist of what we are talking about today and in this meeting in general. The drug label, as many realize, is the legal basis of prescribing. This comes from 21 CFR 201.57. You will see it several times today, but it is basically a mandate that if evidence is available outcome support the safe and effective use of a drug only in select subgroups, then the labeling shall describe the evidence and identify specific tests needed for selection and monitoring of patients who need the drug. Those test, in turn, are also derived from biomarkers, whether they be genomic or non-genomic.

[Slide]

Now, label revisions are quite common. Labels, despite what people say about physicians and providers

reading them, are the most frequently consulted information sources and, while FDA has many ways to update the physicians and patients, label updates remain one of the main tools for informing physicians and patients about new risks or new data that allow them to make informed decisions about drug therapy.

Frequently, the original label of a drug product reflects the pre-approval data. To get into the marketplace, efficacy is documented frequently by two randomized clinical trials that have p less than 0.5 but, because of the limited exposure of drugs, safety is often provisional. It is not uncommon then for new insights post-approval to alter benefit-risks and these, in turn, will drive regular label revisions. There may be drug interactions; new information in special populations; of even genomic biomarkers. These revisions are particularly important for individualizing therapy and determining what makes one patient different from the other when it comes to time to dose.

[Slide]

Label revisions also have limitations. We don't always have the extent of information that we need. Many physicians would prefer precise management advice--how do I reduce a dose specifically? But sometimes evidence in labels becomes more descriptive and actions general, for

example, we might say reduce the dose or titrate carefully or monitor more closely. We think this is important information and it stops short of specific dose recommendations.

The reason for that is that frequently in revising labels we lack perfect evidence, for a specific dose reduction for example, but we feel this is not a reason to support inaction when we have a preponderance of evidence that supports safety or efficacy or improved dosing.

[Slide]

As an example of this, last year at this time or just about this time we discussed irinotecan. We know that this was a drug that was effective in first- and second-line therapy for metastatic colon/rectal cancer. But providers and patients faced a clinical predicament. That is, what is the optimal dose of this drug? As we heard last year, it is not well determined.

What we do know from both clinical trials and post-marketing experience is that the incidence of neutropenia, grade 3-4, serious, is high, 35 percent. In fact, nearly 70 percent of patients needed dose reduction. Toxicity of the drug is related to the SN-38 exposure. If you remember the label from last year, it had phrases in it that included "causes severe myelosuppression," "death due

to sepsis following myelosuppression," "adjust doses based on neutropenic count." And, we asked the question can we do better than that?

[Slide]

The problem was accumulation of SN-38, exposure dependent on the metabolism of SN-38 by UGT1A1. We new there was wide inter-patient variability in the activity of this enzyme. Patients who were *28, for example, had a reduced enzyme activity, and homozygous deficient patients had significantly reduced activity and were at the greatest risk of neutropenia. We know that neutropenia matters to patients. It is harmful. It causes hospitalization. It is inconvenient to the family. The original label was silent on UGT information and the approved dose was not optimized.

[Slide]

We presented last year the risk assessment by genotype and asked the question would an adjunct UGT diagnostic test to identify patients who are 7/7 genotype, that is, those most at risk, would lead to a lower risk of neutropenia versus the standard of care without genotype.

As you can see from this table, the prevalent of 7/7 genotype was 10 percent. The risk of neutropenia was 50 and 100. It was not a perfect test. It did not predict 100 percent of patients. Some patients are 7/7 and are not

toxic. But, nevertheless, it was another piece of information to add to the other information that was in the label at that time on age, prior exposure radiation and other cofactors that enabled a physician to make the most informed decision possible in the use of UGT testing and irinotecan dosing specifically.

[Slide]

Shortly after our advisory committee the camptosar label was revised as was recommended by this committee, in conjunction with our oncology counterparts in the Office of New Drugs, and shortly after that the FDA approved the test for UGT that is now widely used in oncology circles.

[Slide]

So, this brings us around to the topic for today. We will talk about another drug, optimizing warfarin benefit-risk with CYP2C9 genotypes. As I mentioned, label revisions are quite common. In this case, there have been over 20 label revisions since the drug was approved back in '54, and we do this almost on an ongoing basis for warfarin. The latest was in September of 2005. That had to do with drug interactions related to proton pump inhibitors and cranberry juice. The evidence for that label revision came from post-marketing surveillance.

[Slide]

To get you thinking about warfarin and genotypes, I thought this quote was appropriate. I am applying it to the success and failure of drug therapy, and it is something Confucius said a long, long time ago: By nature, men are nearly alike--I am not sure he was thinking of genomics at that time--but by practice, they get to be wide apart. I asked Shiew-Mei to read that slide with the Chinese and she can. She will tell you later what that means.

But it points to the issue of adverse drug reactions. Is there a predisposition for adverse events as we might think of with genotype, or is there a susceptibility to adverse events that might be due to the environment? It is probably a combination of both.

[Slide]

This drug doesn't need a lot of introduction. It was discovered 60 years ago and is one of the most widely prescribed drugs in the world. It is also one of the remarkable drugs in medicine in the benefit that it has brought to patient care. It has probably saved endless number of lives over the course of time. Of course, it is intended to treat and prevent thromboembolism in a variety of at risk patients, those with atrial fibrillation, recurrent stroke, deep vein thrombosis, etc.

It also comes as a multi-source anticoagulant in many different strengths, reflecting the difficulty in arriving at the appropriate maintenance dose for patients. Further, since its introduction a long time ago, there has been a significant increase in prescriptions related to the use of the drug.

[Slide]

This gives you a sense of the prescription use of warfarin. It shows a 1.5-fold increase or 45 percent increase just in the last 6 years. You can see the trend which kind of tracks the increase in the elderly population in the country, a typical population for receiving warfarin along with many other drugs.

[Slide]

The efficacy of warfarin is not a debate. There are many prospective clinical trials that unequivocally demonstrate effectiveness. When one looks at mortality risk in untreated patients for example with atrial fibrillation, there is a 2.5-fold greater risk than in warfarin-treated patients. The risk of ischemic stroke in these patients with warfarin is reduced by 65 percent. If you like quantitation, the number need to treat versus placebo to prevent one stroke is 32. So, the drug undoubtedly is effective.

[Slide]

There has been a continuous debate, however, of a global problem of adverse events with the drug. Two million people in this country receive warfarin. If you look at surveys that are published in the literature from academic centers and hospitals, warfarin is at or near the top of those surveys in the amount of adverse events.

It is not only in the United States. In Sweden 70,000 patients, one percent of the population, receive warfarin and there it tops the list of drug-induced adverse events. In the U.K. 600,000 patients take this drug, many of them over the age of 80, and there are 18-20 episodes of hemorrhage per 100 patients.

[Slide]

Recently there was an article on a ten-year survey of adverse events and warfarin accounted for 3.6 percent of all those drug adverse events. It was the fourth ranked drug after drugs like digoxin and some anti-infectives. But 15 percent of all the severe adverse reactions in that survey over the ten-year particular were due to warfarin, and that was second only to digoxin. So, its efficacy well established; so is the toxicity of the drug.

[Slide]

The safety of warfarin then becomes an issue. The major risk is bleeding. It is frequent and severe on a

relative basis. There are 1.2 to 7 major bleeding episodes per 100 patients. It is reported to be responsible for 1/10 hospital admissions. The relative risk of fatal extracranial bleeds is open to debate. The range is anywhere from zero to 4.8 percent depending on the study and the study, in turn, is dependent on the patient population and co-morbidities. So, naturally, it is going to be variable. The number needed to harm, in contrast to the number needed to treat, is 333. That is the number of patients to be treated in order to elicit one adverse event.

[Slide]

The problem with warfarin, of course, is doing with the drug. The dosing is complex because of the following: It has a narrow therapeutic index which we define as a small separation between dose-response curves for preventing emboli and excess coagulation. When dose adjustments are necessary it has a nonlinear dose response, response here being defined as the INR, the main way we measure warfarin dosing. Small changes in dose may cause large changes in INR and there is often a time lag between the change in dose and the change in INR.

Lastly, there is a wide range of doses, 50-fold range of doses, 2-112 mg per week to achieve a target INR in most patients of 2-3. Usually, in most patients that

range is somewhere between 2-10 and 2-15 mg. And, we know it is variable. We know there are many intrinsic and extrinsic factors that account for that variability in dose requirements.

[Slide]

This is an information-rich slide that I would take a moment on because it makes the point that mechanistically we know what we are dealing with warfarin on the pharmacokinetic side. We know there is large inter-individual variability that is related to S-warfarin metabolism, warfarin being a racemic drug there is S and R. S is the active metabolite and virtually all of that, roughly around 80 percent, is metabolized by 2C9. Genetics is the predominant determinant of activity of 2C9. There are three alleles basically in all the populations that have been studied. There haven't been many more than these three identified.

What this table shows is the genotype from one study, which I have indicated on the bottom of the slide, which shows the distribution that is typical for the different genotypes of warfarin, and *X is either *2 or *3. So, this is a homozygous wild, heterozygous with an allele variant, and the homozygous with the allele variant. You can see the distribution and there is a fairly substantial

portion of the population, 40 percent, that have one or two variant alleles.

When you look at enzyme activity, whether it is in vitro or in vivo, this is the remaining enzyme activity compared to the wild type and, not surprisingly, this is reflected in the S-R warfarin ratios in the blood stream where they change as you go down in genotype with the SX having the highest blood levels from equivalent doses. That, in turn, is due to the clearance of the drug. You can see that the clearance of the drug goes down as you go through the genotypes and, as a result, the weekly doses of the drug go down. So, from these type of data, from the pharmacokinetic data, it looks like the genotype is the major driver of dose.

[Slide]

These two graphs basically compare the relationship between the clearance of the drug on this side and the typical dose that would be appropriate based on 100 percent for the wild type and the dose reductions based on the reduction in clearance. So, it shows us a way to think about reducing dose using the typical exposure differences that we have in patient subsets with changes in clearance and changes in area under the curve. And, we know that for this drug exposure matters and area under the curve matters.

[Slide]

With regard to the pharmacodynamics, we have a mechanistic basis for variability in response. We all realize INR is the measure of intensity of anticoagulation. When we look at dose-plasma level INR relationships we find that plasma warfarin, in turn dependent on clearance and in turn dependent on genotype, is a strong predictor of changes in INR measurements.

The INR itself accounts for 15.3 percent of variance in warfarin and there is wide inter-individual variability in terms of the INR predicting clinical outcome, with stronger correlations as the INR value goes up. The point of this slide is that response with a given INR is also variable, as is the response with specific dose.

So, the difficulty in achieving a target INR and the frequency of adverse events illustrates the limitations of INR, with due respects to its primary benefits which is the main way we monitor therapy with this drug.

[Slide]

The importance of the INR is illustrated on this slide. What this shows is the benefit of INR as it relates to stroke prevention. These are odds ratios and these are the INR. As I mentioned, the typical range for most indications is 2 to 3.5; sometimes it is higher on some

specific applications. But you can see that the odds ratio of preventing a stroke gets better and better as INR gets into this range. So, there is no doubt that being in this range is beneficial in terms of stroke prevention.

[Slide]

But if you imagine laying this slide on top of that other one, you can also see that there is also an upper limit of INR where the odds ratio goes up with respect to the risk of extracranial bleeding. So, putting the two curves on one another, you can see very easily that it is critical to maintain a patient in this therapeutic range of INR of 2 to 3 or 2 to 3.5.

[Slide]

Now, there is an unequivocal association, in my opinion, between 2C alleles and warfarin-induced bleeding. This comes from three different studies that were conducted really around the world in the global sense and in each case the odds ratio for intracranial bleeding, warfarin-induced bleeding, was 2, or really approximately 2 in each of the studies that were presented. The references, again, are on the bottom of the slide. So, we have some linkages between the various mechanistic aspects of warfarin response.

[Slide]

Another point that I want to mention is the quality of anticoagulation, which is generally poor despite INR monitoring. There are several studies that point to this. I have picked out a few. We measure the quality of anticoagulation by the time that a patient spends within the target range, and 62 percent is probably the upper limit and as good as it gets. There is more time that a patient spends below than above the therapeutic INR range, perhaps reflecting the conservative nature of the dosing because of the fear of over-dosing and intracranial bleeds.

Another study to target INR range in 100 patients was achieved on 44 percent of the patients. Again, sub-therapeutic levels predominated over super-therapeutic levels.

Finally, in another study only 14 percent of patients met the criteria for quality anticoagulation control, which was defined as time in the INR range.

I guess this points to words that we can do better. There are many reasons for the results for these types of studies. It is reported in each of them that concern about adverse events is one of the reasons for conservative dosing and slow titration up to a stable steady-state dose.

[Slide]

Now, as you think about warfarin in our discussion today, think about the two phases of dosing for this drug. The first is the induction phase. This is a naive patient going on the drug for the first time. There is an intended therapeutic INR range of 2-3. Typically in this scenario there are daily, biweekly or weekly INRs depending on how fast the patient is stabilized. It is characterized by frequent dose adjustment and response to INR so it is a reactive period of time. And, generally one reached an INR target in 4-5 days on average but it may take 7 or 30 days to reach a steady state both in INR and the dosing.

This is in contrast to what I will call the maintenance phase. The maintenance phase is when target INRs are achieved. First following the induction phase INRs are done less frequently. Doses are changed less frequently but, of course, dose adjustments are needed based on changed in the patient's situation, whether it be that a drug is added or deleted from the regimen, dietary changes occur and things of that sort. What I am going to focus on primarily is the induction phase and not say much about the maintenance phase.

[Slide]

I appreciate there are many ways of approaching the induction phase of dosing with warfarin, but I think

most of the paradigms that I have seen in clinical scenarios--and we will hear more about this later this morning--is that the initial dose is selected as the estimated maintenance dose based on patient cofactors. A typical starting dose of this drug is 5 mg per day. Those cofactors which predict higher doses and, by converse, loser doses or warfarin, include things like the indication; things like co-morbidities; the patient's age; male or female; ethnicity; intake of vitamin K; weight and concomitant drugs. So, all of these things, along with the physician experience, are taken into account and the predicted maintenance dose is then initiated and the patient has the initial dose.

[Slide]

INR monitoring during the induction phase is tricky. The label says for warfarin to individualize dose based on the rise in INR. The INR rise, in turn, can be deceiving in that the initial dose suppresses only one of four factors and there are actually more that are responsible for clotting, and these are the vitamin K dependent factors. So, the initial effects of warfarin are on Factor VII so within 3-5 days INR appears to increase. It starts out at 1 in an untreated patient; begins to rise; reaches a stable INR in 3-5 days.

Because the half-life of these clotting factors varies up to 60-plus hours, continued dosing of the drug inhibits the other factors, resulting oftentimes in an overshooting of the INR since there is this delayed effect between dose change and INR measurement. INR in the first 4 days of therapy has a 65 percent rate in predicting dose--not bad, but can it be made better is the question.

[Slide]

To show this schematically, I would say that warfarin dosing in practice translates into one size fits few. Basically, we have a population of patients. The initial dose of 5 mg per day based on patient cofactors of age, gender, etc. is intended to adjust that initial dose up or down depending on the nomogram that people are using.

Typical INR range at that point in time in the first 4, 5, 7 days--one is looking for a range of 2-3 but frequently INRs are 30, 35 percent below or 20, 25 percent above so there is a need for a dose adjustment. Based on that INR, the dose is changed. It is increased or decreased and actually there is a time lag until we see what the next INR is. So, that is repeated and the dose is adjusted again and again.

Generally, in most anticoagulation clinics this takes upwards of 30 days of INR measurement and dose adjustment until finally a stable maintenance dose, with an

INR range of 2-3, is achieved. Notice that that range of variability in INR is compressed by virtue of having the feedback of the INR.

But what happens at the end of the day is that instead of 35 mg the distribution of doses within the patient population may be significantly less and in some cases significantly more. So, the question is if I have taken the cofactors into account in the initial dose but it fails to predict the final dose--and these are actual doses taken from population studies of warfarin--what is missing and what could be added to enhance the predictability of the initial dose?

[Slide]

This is a clinical example that is typical of the problem with initial anticoagulation rate. It is one of several that appears in the literature and it shows the difficulty in the early dosing of this drug. This is a patient that was an elderly patient in a nursing home who received a prescription for warfarin because of a diagnosis of femoral deep vein thrombosis. After about 7 days of dosing, as we tack this patient, INR was 2.5. The patient was thought to be okay and was advised to continue for 12 weeks. Unexpectedly, the INR shoots up, very high, 66. Investigation showed it wasn't related to over-dosing in the nursing home. It was kind of confusing what that was.

The patient was treated in the hospital, discharged, and 4 days later the INR bounces up again. Again, this was unexpected at the dose the patient was receiving. Investigation revealed no changes in drug, diet, medication, but the half-life of this drug is 10 days so steady state was still being achieved on the dose. Eventually the drug was discontinued. The half-life of the drug was fairly long. Ten days later it was 1.1 in the blood so that the patient became sort of stabilized and warfarin was discontinued.

The point of this case study is that the spiking, the variability in the INR is not typical in the case of an induction regimen. Further investigation of the genomics of this patient found that the patient was a homozygous with compound allele variant *2, *3 which, in retrospect, explained the problem of adverse events with a normal dose of warfarin.

[Slide]

So, the message that I would like to emphasize then is that implications of the difficult induction phase for patients with 2C9 alleles result in more frequent changes in daily dose; delayed stabilization and hospital discharge; multiple visits; additional investigations to figure out what is going on in the increased risk of bleeding. In fact, the *2, *3 are consistent risk factors

across many studies, many of which I had on the slide, and the magnitude of that risk all point in the same direction but the magnitude does vary from study to study.

[Slide]

Which brings us around to what we want to think about during the course of the morning. If the risks of warfarin are greatest in the induction phase related to incorrect dosing, with the risk of either bleeding if we have too high a dose or thromboembolism with too low a dose, can we do something about that? We know that the majority of warfarin-related adverse events occur during the first 30 days of therapy along with the induction phase of dosing.

[Slide]

This illustrates that. This is the frequency of major bleeds following the initiation of warfarin dosing. It shows major bleeds as a function of time. The first bar is up to 4 weeks of therapy, what I have defined as the induction phase, 3 percent in this study. As you can see, over time, as the dose stabilized, INR stabilized, adverse events go down. Not shown on this slide is the cumulative increase in adverse events that do occur with continued dosing of the drug, but since we are only looking at the induction phase and perhaps ways of improving that, we will focus on those first 4 weeks.

[Slide]

So, prospective genotyping of 2C9 is part of the theme of today's meeting. Would knowledge of a patient's genotype improve the dosing during the induction phase and reduce the incidence of adverse events which I have already defined? I want to emphasize that whenever we talk about genotyping we are not talking about it as a replacement for other cofactors that are used in clinical practice, but it is an additional piece of information that would be added to the normal standard of care where age, drugs and other things are taken into account in predicting the maintenance dose during the induction phase.

[Slide]

Sometimes people like to think about the incremental value of genotyping. We discussed this in prior discussions of thiopurines and irinotecan. Looking at incremental value, the question could be asked how much of the inter-patient variability in dose is accounted for by genotyping?

What I have done on this slide is tried to compare and answer that question. There are three references, all of them are recent. What this shows is the relative percent of variability in the dose explained by 2C9 alleles alone and by all other factors. Now, all other factors includes the usual cofactors that are associated

with maintenance dose prediction--age, body weight, body surface area, indication, gender, etc. When we compare the two, the alleles alone predict 27 percent, 12 percent and 20 percent, depending on the study, of variability in the dose comparing the maintenance dose to the induction dose. On the other hand, all other factors combined predict 10 percent, 18 percent and 27 percent respectively of the variability. It would seem to suggest that in multiple studies genotyping is at least as good, if not better at predicting dose-response variability than the other cofactors.

Later on today you will hear about a new gene, VKORC1. I have not included that information in this slide but I think what we will hear from the published literature, from Dr. Gage, is that the addition of VKORC1 continues to add to the prediction of variability in dose necessary for patients in the population.

[Slide]

Finally, we get around to how genotyping might help in the label of warfarin, particularly with regard to anticoagulation in the induction phase. Some ways one can think about it, using what we have learned in our discussions about thiopurines and irinotecan, is that genotyping information, when used in addition to the other cofactor information that we have, can identify high risk

patients for adverse events. So, somebody with one allele or two alleles of 2C9 would be at greater risk during the induction phase.

The technology of genotyping is such that there is no need to delay dosing. A patient could start on the usual 5 mg dose, or whatever dose is determined by the obvious demographics and clinical scenario, and a genotype could be obtained in the first few days of therapy. The turnaround time on this test can be less than half a day.

Knowing the genotype can result in more conservative dose increases than one might normally use. It might also lead one to more frequent INR measurements. It also might lead one to think about lower target maintenance doses for that patient. Conversely, genotyping can be thought of as identifying patients likely to require higher maintenance doses. If one were a wild type genotype, coupled with other factors such as younger age and ethnicity, one could estimate a relative risk of toxicity by proceeding with maintenance dosing. One can also identify low risk patients, low risk patients who don't need anticoagulation as badly that they need warfarin and for somebody at risk, based on the genotype, a physician might think about alternatives for that patient, such as aspirin that works better than placebo as an anticoagulant, not as good as warfarin. Finally, when

faced with a patient as I profiled in that case study, we think genotyping would be a useful tool to investigate unexpected toxicity of resistance to help guide what to do next in terms of drug dosing.

[Slide]

These are my colleagues that helped put together the evidence database that I presented today. I want to thank them and acknowledge them.

[Slide]

Finally, after laying this ground work for you, what we are going to hear is a review of the way we have genotype data in labels currently and look at that generally, and we are going to further look at some of the evidence--not what I have covered today but what Dr. Shiew-Mei Huang will cover later for warfarin specifically in terms of labeling. Thanks.

DR. FLOCKHART: Thank you, Larry. We have a bit of time since you have been very responsible. So, I think it would be appropriate if members of the committee had a couple of questions of clarification before we go on to Shiew-Mei's presentation. Does anybody have a specific question? Yes?

DR. SINGPURWALLA: How do you define odds ratio and how do you estimate it?

DR. LESKO: If I had my evidence-based medicine book I could remember what it is. It is a standard formula for calculating odds ratio. The numbers I took were from the published literature that I had on the slide.

DR. SINGPURWALLA: They are not from the empirical data?

DR. LESKO: No, it was exactly from the literature and the methodology that was used in that paper.

DR. SINGPURWALLA: Well, it is the ratio of two probabilities so I was curious how you get those.

DR. LESKO: I don't know the specifics of the formula. Do you think--

DR. SINGPURWALLA: It is standard. I just wanted to know how you get it.

DR. LESKO: Yes, I think it is a standard way of calculating. The data that I showed cam from a meta-analysis of several studies of warfarin toxicity looking at odds ratios.

DR. SINGPURWALLA: So, what you should really do is start with the odds ratio given in the published literature, update it in the light of your empirical evidence, come back with the posterior odds ratio and then put that up.

DR. LESKO: Yes, that is a little more sophisticated than I prepared for today really, but the

odds ratio from the publication, yes, we can certainly do that and figure out initial odds ratio was.

DR. SINGPURWALLA: I have a second question. Sometimes it is true that an additional piece of information increases your uncertainty because of surprise. So, there is always a tradeoff between the cost of getting additional information and the disadvantage of increased variability. In this particular case, what is the cost of getting this genotype information or is it free?

DR. LESKO: I don't know of any introductory free trial period of genotyping! There are at least two published studies on the cost-benefit of 2C9 genotyping that have been done. Both of them conclude the same thing, that it is cost effective. The prices that were estimated are those prices that would probably be typical in an academic medical center. As far as I know, the test itself is not commercially available via an FDA-approved test so I don't know what the commercial market would be as far as charges go. But the cost of hospitalization with a major bleed is fairly substantial. I have seen ranges anywhere from $16,000 to $30,000 depending on the stay in the hospital. The cost of genotyping, just as a ballpark estimate, going out on a limb, might be $200. I am just basing that on an estimate based on some other genotyping but others who do these tests might be able to comment.

DR. FLOCKHART: Wait, you were referring in part to the cost of finding increased variability in these kinds of tests. So, I think part of the thing you didn't address, Larry, was the idea of being able to reduce the amount of medical interactions in a patient because of a more confident idea of the dosing schedule in advance. Have you considered that? I haven't seen any studies that have directly addressed that.

DR. LESKO: I am not aware of any studies that actually looked at that. By and large, the metrics of intervention with 2C9 are related to clinical outcomes. I think it follows that different clinical outcomes, whether it is reduced hospitalization, that type of data is in the literature--reduced adverse events. I think one can extrapolate from that, although I am not aware of any specific study that has looked at those kind of metrics at this point in time.

DR. FLOCKHART: I think it is an important thing. Over the years I have talked to anticoagulation clinics around the country and, in general, the physicians and prescribers are certainly concerned about intracranial bleeds and over-dosing, but the actual interaction time that they have with patients, the patients that are the most difficult, most complicated, are the people on high doses who are very difficult to manage. So, a reduction in

them, from their perspective--a reduction in the interaction time with them would be valuable. It is something we haven't really thought about a lot but I think it is important to have as part of the discussion whether we could better at predicting that too.

DR. LESKO: Yes, there is a fair amount of data in the literature that compares, I would say, the success or lack of success of anticoagulation clinics. People have estimated the cost of running anticoagulation clinics and the benefit that has been accrued from that in terms of the types of things you are describing. I don't have that data at hand but I know there is published literature on it. But that might be a source of that kind of information that we could possibly gather and take a look at.

DR. FLOCKHART: Dr. Sadee?

DR. SADEE: Larry, you stated that 80 percent of the metabolism is accounted by 2C9 but that is patient dependent, isn't it? In those patients where you have low 2C9 activity it probably goes down to 20 percent, in which case drug-drug interactions would be targeting another cytochrome isoform. So, how do we deal with that?

DR. LESKO: You know, that is a good point. Seventy-five percent via the 2C9 for the S-warfarin is a population estimate, all patients. I think what you are raising as a question is what is the relative risk of other

cofactors in different genotypes. So, if I happen to be a poor metabolizer but am also taking a drug that inhibits the metabolism, or if I am a fast metabolizer and taking the same drug, what is the relative effect of genotyping? It is something that I think of as a gene by cofactor interaction and I am not sure we have the information.

One question that might be asked is do these factors operate independently or are they co-dependent on each other. I think the answer probably lies in looking at individual patients within some of the population data studies that we have. Maybe we will hear a little bit about that today from others that work with this in a clinical setting, but I think it is a fair question that would be worth looking, relative value of genotyping in subsets, although subsets defined by different cofactors are going to be quite expensive. So, it is going to be a tough question to answer I think.

DR. FLOCKHART: I think which patients are most vulnerable to drug interactions is something that we need a little work on. Dr. Barrett, one last question?

DR. BARRETT: Larry, when I look at the data in this capsule, it appears that the 2C9 is actually more predictive than just on the pharmacokinetic side. You showed a nice connection of the INR to plasma concentration, but we also know that within a patient, as

far as the pharmacokinetics go, there is a little intra-subject variability. Yet, when I look at that profile in terms of the INR accomplished in the patients you showed, I mean it is very striking to see, you know, what can occur within an individual. So, not so much on the issue of other cofactors that may be predictive, but as far as the relevance of 2C9, do you feel that there is an additional benefit to the genotyping beyond just explaining sources of variation on the kinetics side?

DR. LESKO: Yes, I think you made an important distinction between inter- and intra-subject variability. In fact, the intra-subject variability is quite low which explains why we can have a drug like this with a narrow therapeutic index both effective and relatively safe. As we talked about, there are some problems.

The question is what is the major driver between the dose and the exposure to the drug? Is it age? Is it genotype, or what? The data that I tried to show is the relative contribution of these factors to the clearance of the drug which, in turn, determines the exposure. The patient that I showed in the clinical scenario was, to me, representative of the lag time that you have between a change in dose and a measurement of the INR. The rate of change of the INR and the rate of change of dose are not in sync with one another so that being reactive in adjusting

dose to an INR causes problems with regard to that lag time. Reacting to a high INR and reducing the dose and reacting to a low INR and increasing the dose--it takes some time to see that based both on the half-life of warfarin and the half-life of clotting factors that it is inhibiting. So, I see some problems in the estimation of the time to consider the observed INR a steady state as well as the dose itself.

DR. FLOCKHART: Last, Dr. Relling?

DR. RELLING: Just very briefly, the statement that intra-patient variability is low obviously depends on the patient population. For example, in the pediatric cancer population the intra-patient variability is so high that we can't even use warfarin. For example, those patients are on a couple of days of Septra. So, I am sure that there are some important patient populations where that is not true and we should at least acknowledge that.

DR. LESKO: Thank you. That is a good point. I was thinking of a stable patient and what you have pointed out is that a lot of things can change.

DR. FLOCKHART: Thank you. Let's move on to Shiew-Mei Huang's presentation. While Larry has laid the scientific groundwork nicely, Shiew-Mei is going to specifically talk about translational pharmacogenomics

information into label updates. So, this really gets to the real substance of the meeting.

I might just give the committee a heads up that later on we are going to consider some questions. Shiew-Mei is going to outline what these questions are. In general, the FDA is asking for our advice in a qualitative way but there is one decision point on which we may have to take a vote and I would encourage you to look at that carefully. It relates specifically to what information we include in the label about warfarin. Shiew-Mei?

Topic 1: Translation of Pharmacogenomics (PGx) Information into Label Updates for Approved Products Topic 1A: Evidence and Process for Translation of Pharmacogenetic Information (e.g., CYP2C9 Polymorphisms) into Label Updates for Approved Products

How New Insights into Pharmacogenetics Lead to Revisions of Product Labels

DR. HUANG: Thanks, Dave.

[Slide]

In my discussion on how new insights into pharmacogenomics lead to revisions of product labels I will be presenting in two parts. First, I would like to talk about pharmacogenomic information in the labeling. After the break I will talk about evidence supporting relabeling

of warfarin to include pharmacogenomic information in the label.

[Slide]

Larry has mentioned 21 Code of Federal Regulations that includes a statement on the following: If evidence is available to support the safe and effective use of drug only in selected subgroups, then the labeling shall describe the evidence and also identify specific tests. So, this regulation will apply to pharmacogenomic subgroups.

[Slide]

The FDA pharmacogenomics working group, which is chaired by Dr. Lesko and includes members from the Center for Drugs, Center for Biologics, Center for Devices and the Office of Combination Products has formed a subgroup to make recommendations on how to incorporate pharmacogenomic information in the label as to where the information should be, which section, and what information should be included in the label. Our general recommendations included in the background information.

[Slide]

We recommend that all clinically relevant information on the effect of polymorphic variation in either drug metabolizing enzymes, transporters, receptors and proteins, and their effects on pharmacokinetics,

dynamics, clinical effects, both safety and efficacy, be described either in the clinical study section or in the clinical pharmacology section. So, this is where you describe the evidence and a description of studies that led to the correlation of the effects of the genotypes and the clinical observation.

[Slide]

And, if this information has important implications for safe and effective use and that result in recommendations that could be placed in various sections of the drug label, such as indications and usage, dosage administration, precautions/warnings, contraindications, boxed warning, additional information may be placed in clinical studies after the description of the correlation data and/or adverse reactions, and when there is a test available we will put it in the laboratory testing. If the information resulted in placement of the genomic information in the first five sections, then this information will also be placed in the highlights section. The highlights section is not in the present labeling but this is being proposed from a proposed rule on physician labeling, which was published in 2000. The final labeling and the associated guidance document should be published soon.

[Slide]

So, when will we put information in indications and usage? If a drug is indicated only for a population with a certain genetic makeup, and a genotypic or phenotypic test is to be conducted prior to the prescription and administration, then you can put it in the section. An example is Herceptin.

[Slide]

In the current labeling in the indications and usage for Herceptin it says that Herceptin is indicated for metastatic breast cancer whose tumor over-express HER2 protein. Additionally, it also tells about the test. Patients whose tumor evaluated with an assay validated to predict HER2. So, this is a required test before prescription and administration so it is in the indications and usage section.

[Slide]

When will we put it in the dosage administration section? If dose recommendations are different for pharmacogenomic subgroups, then it will be in this section. An example would be irinotecan.

[Slide]

First I want to go to the clinical pharmacology section. This is where we put the evidence and summary of studies supporting the other recommendations. Here it discussed the active metabolite of irinotecan; how it is

metabolized and indicated the specific allele that would reduce enzyme activity, and also give the distribution of the prevalence of this population that is homozygous, and further description of clinical studies. So, this is descriptive information showing the evidence.

[Slide]

When we have the recommendation, for example with dosage changes, we put it in dosage and administration. Here, under the heading of dosage in patients with reduced UGT1A1 activity it discussed patients who are homozygous for UGT1A1*28, and there is a recommendation for reduction in the starting dose by at least one level of camptosar.

[Slide]

So, when would we put it in precautions and warnings? If individuals with certain genetic makeup are more sensitive to one of the severe adverse events--and again I will use irinotecan as an example--under the warnings section it actually describes UGT1A1*28 as increased risk for neutropenia. It further discusses reduced initial dose and refers back to the dosage and administration section, and also describes what to do with heterozygous patients in some of the information related to that population.

[Slide]

What about contraindications? If individuals with certain genetic makeup are more sensitive to one of the life-threatening adverse events that cannot be managed via dose reduction, for irinotecan we have recommended for homozygous *28 a change in dose. But in cases where this cannot be managed we put it in contraindications. The example that we have so far is thioridazine.

[Slide]

Under the contraindications section it indicated elevated levels of thioridazine can increase the risk of torsade de pointes and so it mentions that it should be contraindicated in patients with the genetic defect resulting in reduced level of CYP2D6. Again, this is an action of recommendations.

[Slide]

What about adverse reactions? If individuals with certain genetic makeup had a higher rate of adverse reactions, that information would be included in the section.

[Slide]

Here I included examples from atomoxetine. The information was derived from the labeling and we made a table to contrast the different rates of various adverse events in both poor metabolizers or extensive metabolizers

of CYP2D6. This is descriptive information included in this section.

[Slide]

What about the laboratory testing section of the drug label? When a specific laboratory test is available we include it in this section. The examples I will give are related to atomoxetine and azathioprine.

[Slide]

Under the laboratory test section of atomoxetine it indicated that laboratory tests are available to identify CYP2D6 poor metabolizers. It further indicated that higher levels in poor metabolizers lead to higher rates of some adverse effects of strattera.

[Slide]

With azathioprine, this is a recent labeling revision and under the laboratory test it discussed thiopurine methyltransferase testing. Here it not only discussed that genotyping or phenotyping patients is recommended, it also discussed what are the specific alleles that may be related to reduced level of TPMT. It also discussed further heterozygous patient and homozygous patient information.

[Slide]

So, what I have discussed are some examples of labeling language in the drug label. What about device

labeling? The subcommittee of the FDA from the genomic working group has discussed what information to put into various sections of the label of a genetic test. For example, some information will be in the intended use section, summary and explanation of the test section, text procedure, limitations, summary of expected results or performance characteristics of the test. These are just examples.

[Slide]

What information do with put in the intended use? I will use the recently approved UGT1A1 assay as an example, where in the labeling it describes that this is an in vitro diagnostic test for detection and genotyping and it indicates specific alleles, *1, *28, of the genes. Then, it says it can identify patients with greater risk for decreased UGT1A1 activity. This current labeling does not identify drugs to be used. So, this could be used in general for drugs that are metabolized by UGT1A1.

[Slide]

What about a summary explanation of the test? Again the example used is UGT1A1 molecular assay. Here it says it can be used to identify patients that may require dose modification for drugs that are metabolized by UGT1A1. It did include information on irinotecan. It did mention that patients with *28 genotype are at greater risk of

irinotecan-induced toxicity. So, this section includes clinical studies that provide information on the genotype and the clinical observation.

[Slide]

This is just a very brief description of the highlights of our recommendation. I have also given some examples of current labeling where the information may not have been consistent, and we would like to ask the committee later for recommendations in general and also what is the best way to present genetic information in the labeling so that it is useful for both providers and patients.

I have shown some progression of labeling language, such as in thioridazine. It didn't mention poor metabolizers or alleles; it just mentioned patients with genetic defect with reduced activity of 2D6.

Then, in the atomoxetine label we mention CYP2D6 poor metabolizers. We didn't mention alleles or other information. A more recent labeling revision on 6-MP and azathioprine we indicated thiopurine methyltransferase alleles specifically in the labeling, and also in the most recent revision of irinotecan we included alleles again.

So, what is the best information that will be useful? Is it the phenotypic information like PMs, EMs, the alleles or nucleotide information? And, when is it

proper to include information regarding ethnic distribution on the prevalence of these alleles and nucleotides in the label?

[Slide]

We will also like to ask how should results of a genotype test be reported? We can alleles or information that we know there are clinical data to support a correlation. What about other genotypes where clinical significance is not certain or is incomplete but we may have some in vitro data?

For example, for irinotecan currently we have *1, *28 so the report could mention the *1 wild type, homozygous for the TA6 repeat in a promoter region, or it could say *28 is homozygous for 7 repeats. It could say *28 heterozygous with 6 repeats and 7 repeats. Or, it could report as "other" and our question could be in this "other" section should we further report out if we can detect TA repeats 5 or 8 if we only have in vitro information?

So, this is what I would like to present before the break in this section on general labeling recommendations.

DR. FLOCKHART: Thanks, Shiew-Mei. I think again we are fabulously ahead of time--actually we are about five minutes ahead of time--so I think it would be appropriate

to have a couple of questions from the committee, if possible focused on, if you like, points of information. We are going to actually have a much more substantive discussion of the questions later but I think a couple of questions to Shiew-Mei right now would be germane if members of the committee have questions. Dr. Sadee?

DR. SADEE: I think giving different levels of detail in various labeling is a dangerous thing to do. I think there should be a hierarchy and you start off with a simple statement that there is evidence for reduced or enhanced activity based genotype. Then you go into the next step and say what might be the nature of this, and then what might be the specific genotype, and then what might be the clinical relevance, and that scheme should be followed every single time.

DR. FLOCKHART: Wolfgang, you are getting to actually debating the point. We are just trying to point out more information. We will get this stuff later. Dr. Relling?

DR. RELLING: Mine is in the same category.

DR. SINGPURWALLA: Mine is not. Do I understand you correctly that the labeling is for the provider as well as the patient?

DR. HUANG: Yes.

DR. SINGPURWALLA: That is what you have said in one of your slides.

FDA Pharmacogenetic Labels: A Clinical Perspective

DR. FLOCKHART: Shiew-Mei and Larry asked me to provide a clinical perspective on this.

What I did in preparation for this were two things. First, I in general tried to go, if you like, to 10,000 or 20,000 ft. to really try and think about practicality of pharmacogenetics in the clinic. In that sense, I think I have to say that I believe this is a very, very important meeting from that point of view. We starting really on the edge of a lot more academic input into pharmacogenetic tests, so there are a lot more tests being considered both at the device section within the FDA right now and many others that the academic community is generating.

So, I think the decisions that we suggest here--we are not actually going to make decisions, of course, we are going to suggest things to the FDA--are critical to the practical use of pharmacogenomics in the clinic and that gets to important issues which are germane to the public health of citizens in the United States, but particularly to this nasty thing unique to our culture, reimbursement. I think what we are about to discuss is tremendously

relevant to practical use and to reimbursement for tests in the clinic.

[Slide]

I have a slide here which is rather self-explanatory but i wanted to expand on it. It is on the purposes of pharmacogenomics. There are three. I think it is used for many, many different things but three centrally. First, to predict response and thereby improve prescribing and the public health. That is the principal reason we are here today. Second, it is tremendously useful from a research point of view in elucidating a drug's mechanism of action, in part because if one finds a genetic association in an association study one presumes that that is the first evidence for a hypothesis that that gene or that particular pathway is involved in the drug's mechanism of action. We have had interesting surprises in that respect over the last 25 years. Because of that, it is a value in the industry to identify targets--choke points, genes, receptors, enzymes and transporters that are involved in a specific drug's action.

[Slide]

We have talked a little bit about the labeling progress that we have had, and I would submit to you that this is very, very early days and that we have to seriously think about how we have done this to this point, and I

would submit to you that we need to look at this critically. We have made a series of very, very delicate, tentative steps to this point--very delicate, very tentative, in some situations very and carefully considered. And, these have implications for how many, many tests will be developed here on out.

[Slide]

We have, of course, TPMT labeling. I didn't put up here and I should have put up thioridazine labeling, which I think was a very important thing because of where the information about thioridazine is in the label. It is in the adverse reactions section, not in the indications and usage section related to the warnings section. As a result, the test is largely not reimbursable for that particular indication.

Irinotecan, I am going to talk about the Amplichip approval by the device section of the FDA, which is the first FDA approved pharmacogenetic test and has led the way for--I don't want to overuse the way--tsunami but there is a large number of other companies and tests that are apparently under review at the FDA for similar indications of cytochrome P450 enzymes and others.

Of course, the second of these, and I think in many ways the most carefully thought through, is the UGT1A1

test approval not just for irinotecan but for UGT1A1 in general.

[Slide]

And just to point out some things that relate to translating pharmacogenetic data into labels, these are some of Dr. McLeod's data on the association between UGTA repeats and neutropenia and activity. The left-hand slide simply shows grade 4 or 5 neutropenia. This is life-threatening, nasty, expensive neutropenia in an ICU, and it is incidence across three genotypes, 6/6, 6/7 or 7/7 genotype. For the 6/6 genotype the incidence is 9.6 percent in this fairly large study. But in the same genotype the overall response rate was 41.9 percent. So, this would seem a valuable genotype in which to use irinotecan relative to the *7/7 genotype, also identified by the FDA-approved test which has a much higher, 4-fold higher incidence of grade 4-5 neutropenia and a significantly lower, 14.3 percent, rate of response.

So, clinicians should be able to take this kind of information and use it in a clinic but, in fact, one thing I did in preparation for this, I presented this information to two groups of people. One is the group of oncologists, which is about 50 practicing oncologists at Indiana University School of Medicine, and the second is to a private group of oncologists at Methodist Hospital, which

is part of our system in Indiana. Just put yourself in their situation for a moment. They never heard of UGT1A1, including the academic oncologists. They never heard of *28. They consider irinotecan a third- or fifth-line drug. And, if you actually went to the label, the original label, they found the original label quite helpful. It generally gave them doses to use, and so forth. This kind of information is interesting to them academically. They like to think about it when they are waking up in the morning, i am sure, and dreaming at night but in terms of practicality, how would you get this test; what do you do with it; very specifically what dosage changes do I make based on these data? This is important information but it doesn't give them the specific information they need to move forward with it.

[Slide]

A similar argument can be made with warfarin and vitamin K carboxylase and 2C9 dose. This is some of the nicest data from Allan Rettie's group. Here it shows the mean warfarin dose, and this is warfarin dose, if you like, at steady state or once it has been reached. You see here patients of three 2C9 genotypes and variant 2C9 genotypes, the three different genotypes of vitamin K carboxylase. These data are going to be described in much more detail later on, but you can see nice and important changes here.

They are not only statistically but probably clinically relevant. And, this is a little bit closer to what you want so you can imagine a group of people here whom you might start at 1 mg or 2 mg and a group of people whom you might feel more comfortable starting at a higher dose but this, again, is a retrospective study.

Two important things about this are not obvious from these data when you show it to community groups or even academic groups. What is the incidence of each individual genotype in a population? That doesn't just jump out at you. You don't know that from looking at that. That is the first thing they ask in this setting--is this a ton of people or is this one or two people? They want to know that. It is not really obvious from this way of presenting, and I am really talking here about communication within labels. We like to use the academic data that we have spent so much time, money, energy and pain generating in lots of contexts, but we really need to be thinking here about communicating exactly what we want them to do in order to keep people safely and effectively treated.

The other, of course, important point here that really Howard McLeod and his group made possible was a replication population, in other words, making sure that these data hold up not just in one population but in a

second population. That ought to be, I think wherever possible, an absolute criterion for information that we include in labels.

[Slide]

Let me just consider specifically these people's reaction to the irinotecan label itself, which was the last one discussed at this committee. Now, importantly, in the dosage and administration section--and I am going to make a lot of points here about which section the information goes in. Shiew-Mei spent some time on this already but we have to recall that 90 percent of the time, more than that in most of the studies, where people go to the label, go to the PDR is for the dose. So, this would be the most often considered section.

Let me must read it through with you: The section is headed dosage and administration with reduced UGT1A1 activity--remember, these guys don't know what that is. When administered in combination with other agents, or as a singe-agent, a reduction in the starting dose by at lease one level--in both settings, neither of these groups could interpret what one level meant. What does one level mean? Does that mean one milligram? What is a level? Is it some predefined thing that you guys you have or is it something in widespread community practice? I asked them. They don't know. So, widespread practice in private

oncology, they can't interpret this. What does this really mean by "a level?"--should be considered for patients known to be homozygous--well, they know what this is anyway. Then you go to "see the clinical pharmacology and warnings sections." These are, if you like, "Seep Space 9" to these guys.

[Laughter]

So, going into that section is something that they have not only never done before but it is something that is couched in language that could be essentially in ancient Greek from their point of view.

Now, here is the real topper from their point of view: The precise dose reduction in this patient population is not known and subsequent dose modifications should be considered based on individual patients. So, they go, oh, it doesn't really matter anyway. I can do what I want. Right? That is exactly what they conclude.

[Slide]

So, some barriers to effective labels, and I am really into this because I think it represents the most important series of barriers to the future not only of pharmacogenetics, which I have to declare vested interest in, but also to its ability to really improve patient care.

This is just a true thing, time in clinical settings is limited; information overload is absolutely the norm; very few prescribers read labels; lawyers do.

[Laughter]

It is funny but it is actually a barrier. It gets in the way. It overcrowds labels with information and it is like it is not on the mind of a prescriber; it is on the mind of a prescriber and that might be good or might not be good. So, this is the mantra: simplicity, accuracy of presentation is not negotiable. I could have maybe added to this a couple of other points about communication but this is one of them: A picture is worth a thousand words to these people.

[Slide]

Now, this is one of Shiew-Mei's slides. All clinically relevant information on effect of polymorphic variation in drug metabolizing enzymes, transporters, receptors and/or other proteins on pharmacokinetics, pharmacodynamics, clinical response, both safety and efficacy, goes in these sections, the clinical studies section and the clinical pharmacology section.

I am worried abut this because I think the subjective interpretation of what is clinically relevant is in huge danger of being over-interpreted. So, we could put all kinds of information in here and, at the moment, if we

a look at all labels you can't find a general rule. I mean, there is all kinds of density of information for some labels and very little for others. The more dense information tends to be more recent labels. We have actually more information, more accurate, high quality information in labels than we have ever had before, but it makes for information overload. I am not saying that we shouldn't put a lot of valuable information in there but I think we need to be really careful about what goes in here because we are about to hit a large amount of data coming, particularly from the genome studies and others, which would on this basis go in these sections.

[Slide]

This is an old slide from an old chapter I wrote many years ago. At that point I titled it a clinical perspective on the hierarchy of pharmacogenetic information. This was just a histogram, if you like, of all the SNPs. There are two or three million in the human genome, and some of these were based in exons and, therefore, coded for protein. Some were non-synonymous and, therefore, might matter for that reason. In other words, they change an amino acid. Their code is degenerate and so it is possible to change base pairs and not change an amino acid. Some of them are non-conservative. These are all assumptions, that this would be less and this would

be more clinically relevant. A huge amount of information is now stuck here; perhaps change of activity in vitro. Some do change pharmacokinetics but there are relatively few, though increasingly this is a big number.

Really, what I want to point out here is in this area, here. I am a clinical pharmacologist. We spend our time measuring responses with p values and the GCRC. So, I can tell you about hot flushes. I can measure blood pressure. I can measure pulses. I can measure changes in serum clotting factors, and so forth--all with p values. But in terms of the people who measure clinical outcomes, who publish those journals, who follow the cost of care, who follow reimbursement, who follow deaths, progression of cancer, and so forth, this stuff isn't really in general very meaningful. What is valuable are those parameters, things that really change life expectancy, quality of life as measured by validated measures that they use. So, there is actually a huge gap between what I measure as a clinical pharmacologist in a clinical research center in a normal volunteer and what these people would consider really important outcomes influenced by these SNPs or genetic variants down here.

[Slide]

So, the first problem is that clinical relevance is over-interpreted. It can't be just a p value; it has to be more than that.

[Slide]

The second I think is that the iterative value of tests is not presented. Now, this was the second large point that came out of these interactions with the community groups, and Larry referred to it a couple of times. But really the idea is I have a lot of predictors already; why do I need this extra thing?

[Slide]

So, let's consider this carefully, and this is just a fundamental truth of any test used by the pathologists in clinics so that people who run clinical labs, generally pathologists but also the Association of Clinical Chemistry, is concerned about this, as is the Association of Molecular Pathologists and a number of organizations who have some representatives here. So, this is just true, the clinical value of tests of this kind decreases when the current predictive ability is high. So, if you just make a cartoon of this, the clinical value of a test versus our current ability to predict a drug response-- and this is now in an article that Bruce Meyer and I put together.

Let me just make some examples here partly to get discussion going and partly to be deliberately provocative. If we look at beta-blockade by a beta-blocker aimed, obviously at a beta receptor for the treatment of hypertension, most clinicians, wrongly I might argue, would believe that they can get a first year medical student to measure blood pressure. So, they would put someone on a beta-blocker and in a matter of days later they could tell you if the drug had worked. They could tell you. The blood pressure changed or the blood pressure doesn't change. So, why do you need some genetic test to tell you whether it has gone down or not? I am trying to overstate the point to make the point.

Now, there are important things there. There are really important things there to consider carefully. Currently we just routine use race in that context. I have an African American male in front of me and I don't think about that person in terms of dose, and so forth, in the same way that I might if it were a Caucasian female. I just don't. There are probably really important genetics within that.

So, where I am going here perhaps is a little bit better explained by our collective experience with TPMT. Currently we use azathioprine and 6-mercaptopurine and we have TPMT in the label but, again, if I looked at that

label carefully, and all of you all of you were heavily involved in that discussion and you are all very aware of what the pediatric oncology community thinks. They can follow the white cap. They can follow the white cap. Children with acute lymphocytic leukemia, from their perspective, are followed very, very closely. It is true. They are followed very closely. So, someone who has a very small drop in the white count, frankly, is jumped all over. They get GCSF; they get phone calls from nurses; phone calls to the parents. All kinds of people are concerned about it. I mean, you have something to look at. It is a toxicity certainly, a decrease in the white count, but it is not something like they don't have another tool to follow it.

So, we add a pharmacogenetic test to that and they go, hey, and what the label came out with is recommending we do the testing when we have a serious problem almost as part of the differential diagnostic rather than something that might be valuable prospectively. So, this is something where I think the value as seen by the committee was relatively low. It is part of our collective experience.

I think antidepressants--and this is now thinking towards the future and a situation where we can't predict the response really well right now. You go to a community of psychiatrists and you say, okay, who here can predict

how well X drug is going to work in the first six weeks to three months, which is what it takes to determine efficacy for an antidepressant? We have to remember that only ten percent of the people who start treatment for depression end up on that individual treatment, and most patients treated for depression end up on more than one drug. So, only ten percent end up on the initial treatment. That means that 90 percent end up on a different dose or a different schedule or different drug from what they started on, and we are not good at that.

Obviously, lastly, in the area of cancer chemotherapy--and I might also put warfarin somewhere in here, you know, you find if it has worked when the tumors come back and it is too late. The horse is out of the barn. And, we have all kinds of predictors in some settings for this but basically for chemotherapy we are not good and our current clinical ability to predict response is quite low.

Now, what does this have to do with the label? I think we have to present something like this in the label. Larry referred to it when he was talking about the percent variability explained by a specific genotype and he referred to some studies where there was ten percent of the variability explained by other things, but this is what the clinicians want. They want to know how much extra bang for

the buck they get by using this test on top of--I am sorry, I am a clinician who works in the field of breast cancer so in breast cancer what about if I have nodes already? I know the number of nodes in the breast cancer case already and I also know the tumor stage and grade. That helps me. And, I also know the age of the patient. That helps me. Then, by the time I add in the estrogen receptor I have a bunch of stuff up here.

So, for the breast cancer oncologist is the idea of personalized medicine is actually an old thing and what do you have by doing that? So, I think we need ways of clearly, simply, effectively explaining to them it adds this. You have this much you can explain at the moment and it adds this much.

[Slide]

Now, pharmacogenetics can be used in the same kind of way. Mechanistically when we understand something really, really, really well I think the value of a lot of studies and tests to figure out more mechanistic things are less valuable and there is a series of analogous discussions that can be had about that.

[Slide]

But back to the overall flow here, "clinically relevant" can be over-interpreted; the iterative value of tests is just not presented. The thing these guys want,

the thing they need most, the specific change recommended and then the iterative value of that test. Now, Larry made the point, and it is a valuable point, the constraints in labels. We can't always put everything in we want to because the evidence isn't always there and the evidence specifically for a change from 10 mg to 9 mg might be really hard to support and, therefore, we make a best guess. We put in what we can. We put in the best we can in order to try and help clinical practice but we kind of step back completely and say we are not going to put something in sections that people read carefully.

The last is this--I think I have made the point several times but I made it for a reason--that the simple genetic tests we are considering now actually are the harbinger of tests that are going to be more complicated. We already use in breast cancer a 17-gene panel, expression of 17 different genes to predict who will respond to individual therapy--a 17-gene panel and that is scored. And, I think that is an important precedent. It is a scored thing. You don't list those 17 genes but you provide a series of scores and a therapy is recommended on the basis of that.

[Slide]

This is just a cartoon of the same thing. So, we have multiple genetic changes down here and we have

sufficiently creative statisticians that they can come up with ways of putting patterns together of these. You can imagine these in different colors or different intensities of these lines, such that we get a predictive pattern for a clinical outcome here, and we have to get beyond the idea of just presenting one SNP or one gene as a change, and we have to try and design things here that will be valuable for these efforts in the future as we put more things in.

[Slide]

This is against just a cartoon to point out that the clinical study section and the clinical pharmacology section are, from my perspective, very valuable things but they are in small print and they are small. Indications and usage is read by an awful lot more people. It is where people really go. And, the big thing is the dosage and administration. So, what I am saying between the lines here is anything valuable we can get into the dosage and administration section, when the science supports it, ought to be there.

[Slide]

Currently we have dose changes recommended in text form usually. I could find two labels going through the entire PDR where the actual dose was represented in graphic form. There are multiple labels where renal function is talked about, for example, and we have tables

and graphs describing how to dose in renal situations and they are used. But a graph of dose versus genotype recommendations should be used where possible.

[Slide]

So, this is a fantasy one just to make the point for the administration section of warfarin. This is if we have the data to support this. This is starting dose here--it is not maintenance dose. All the studies we have so far, apart from one, relate to maintenance dose. The idea here would be that you had a series of genotypes down here that helped people decide which dose they start at. The FDA and the academic community collaboratively have thought hard about how to design a prospective trial to make this kind of information available.

I want to point out two things about this. It tells you clearly what dose to take for a given coded thing down here. I put genotype here because I am trying to get at the iterative value of the genetic test. But one might very reasonably put down here genotype plus other predictors. So, you could put down here particularly age and come up with a predictive formula in the same kind of way as we have in other contexts for if you have age of this and you have genotype of this and the genotype, of course, in this context is going to be two genotypes, 2C9 and VKOR genes at least, and that comes up with some

pattern. And I think this is doable. We have multiple situations where we use, if you like, fudge factors, things like the Gale index in breast cancer, to come up with a number and then that is used to determine the risk and why shouldn't we use it for a drug dose.

[Slide]

So, that is my recommendation for the dosage and administration section. Then, I think in the indications and usage section we have to have data on the specific genetic populations that would be targeted in the same way as Shiew-Mei outlined for Herceptin only in people with HER2neu-positive breast cancer. One might argue that similarly strong language should be used in the thioridazine 2D6 kind of section where thioridazine is contraindicated in people who have a 2D6 variant. Then, in the adverse reactions section there has to be a clear genotype or genotypic pattern to be avoided.

[Slide]

So, overall these are my suggestions. I made points about including in the label the specificity and sensitivity of the data specifically when I was talking about irinotecan and people wanting to know the number of patients in whom individual genotypes occur, and they also care a lot about the question of if you have a specific genotype, then what is your risk, and not everyone UGT1A1

and *7/*7 gets neutropenia. Not everybody avoids it. So, what is the sensitivity and specificity of that data, and obviously, the incidence in specific ethnic subpopulations in the same kind of context. Where this should be in the label I think is up for discussion.

A description of the clinical context, and this gets to this point about the iterative value being a listing of currently approved predictive tests in the clinical studies section. In the example I gave you for breast cancer this would be quite a long list that are currently predictors at the moment. If you go to clinicians, very often they feel they have predictors for a lot of things. Anybody that you went to for warfarin would have age certainly, and body weight, size, even gender which isn't supported by the data, but people have a list of things that are used for warfarin at the moment. So, in that context what is the value of the genetic test?

We should be scientists about this. We need to make an attempt to quantitate the iterative value of a pharmacogenetic test in at least the clinical pharmacology section. So, I think the variabilities that Dr. Lesko was talking about and how much they explain should certainly be there, but it is really on top of everything else--what additional value does this test have? That might arguably be not only included in a drug label but, conceivably, it

ought to be looked at seriously by the device label people as well.

Lastly, a clear clinical consequence--change in dose or the possibility--and we are increasingly going to come across this in environments where multiple drugs are available for the same indication--hypertension, breast cancer and so forth--considering an alternative drug. I will stop there and resume my role as chair of the committee and now you get to ask questions. Yes, Shiew-Mei?

DR. HUANG: Just a clarification on the irinotecan labeling. You mentioned that we don't know what a dose level is. Right after that section there are the other sections where--

DR. FLOCKHART: Yes, we don't read them.

DR. HUANG: Because it depends on whether this drug is given as a single agent or in combination--

DR. FLOCKHART: Yes.

DR. HUANG: So, there are differences so it could be 25 mg/m^2, 30 mg/m^2 or 50 mg/m^2, depending on the schedule or the indication. That is why it says dose level.

DR. FLOCKHART: It is a complicated thing just to put simply up front. One could argue for graphics in that context. But I think the other thing related to that label

is the language later on that really says, you know, we don't have evidence to support any change. Dr. Relling?

DR. RELLING: I hate to bring this up but you showed a slide so, back to the irinotecan again, the genotype is related to neutropenia and there is a companion slide showing the genotype having an inverse relationship with response. So, the labeling change for irinotecan is based exclusively on trying to avoid toxicity and how is the issue of the potential decrement in response dealt with?

DR. FLOCKHART: Excellent, excellent point that I omitted to make, and it was the whole reason for bringing that slide up with both things. So, the label spends a lot of time on dealing with neutropenia but, obviously, the clinicians are most concerned, as you made the point in another context, with response. It is not all about toxicity, especially when you are talking about cancer chemotherapy. So, they want the balance. They really want both and they want information from both sides of that thing, and I think a major issue with the irinotecan label is that it doesn't have that addressed sufficiently. It is a really important thing, very important point. Dr. Sadee?

DR. SADEE: I think this was a very clear presentation and I like the concept of uncertainty, and so on. So, in this slide you have made a strong point of

providing all the scientific evidence in the clearest way possible. The question I have is that the biggest jump is from that to the clear clinical consequence. In the label, what does it mean? As you said, a lawyer is going to read this very differently from a physician or a scientist. So, how do we go from the scientific evidence that is available to a recommendation or a clear decision-making?

DR. FLOCKHART: Let me outline my own thinking about this, and this is my own purely personal thinking. Good science needs to be in the clinical studies and the clinical pharmacology section. I think we do need to be careful about what goes in that. We need a bit of a barrier, a filter for things relevant to clinicians, relevant to providers, relevant to prescribers. But the full scientific support for what we plan should be there. I am arguing for something more than that. I think what we tend to do at the moment is put everything there, everything there including the substance of what we are trying to recommend, and I am trying to recommend that important things relative to genetic test effects be included in the dosage and administration and indications sections which lots of people read, higher above that. So, I am not in any way wishing to dilute the scientific support for those things, but to bring them out into the visible area, if you like.

Because right at the moment there is one really important point I think, the chances of something getting reimbursed, if it is not in the dosage and administration or indications section is very low. It is not a universal statement but if it doesn't make it to that point--for example, thioridazine where we have said very clearly it is contraindicated the test isn't reimbursable. It is not in the dosage; it is not in the indications and usage section. And, that is what you get into when you are talking to Medicare.

DR. SINGPURWALLA: I would say this to you off-line but I have been repeatedly told not to talk to any committee members off-line so I am going to say this here. Your graph, clinical value decreases when current predictive ability is high, could use a little change.

DR. FLOCKHART: Okay.

DR. SINGPURWALLA: Because the way you have drawn it, it says that if the ability is very high the clinical value essentially is zero, which obviously is not true.

DR. FLOCKHART: Well, actually I think it is true. I think if the real clinical--

DR. SINGPURWALLA: Can I finish, please?

DR. FLOCKHART: Sorry.

DR. SINGPURWALLA: I would like to suggest that you use an exponential curve rather than a straight line.

DR. FLOCKHART: Right.

DR. SINGPURWALLA: That is a suggestion. The second one is about your fantasy dose. That could be slightly misleading. The reason is the following: With genotype, if there is some other factor which interacts the continuous increase may take a change. So, a graph like this is a very good idea but I don't think it is going to be practical because graphs like this should be multi-dimensional. They should have other factors also going in and it is difficult to show those things. So, unless genotype has nothing else which negatively interacts with it, this graph is fine.

DR. FLOCKHART: I am glad you brought it up. I tried to make the point--and I agree completely. Let me just walk through two possibilities here. The obvious thing is age. Right? So, you could have 2C9 or VKORC1, some code for them down here, and you could make a third dimension which is age. Right? And, that would be okay but it would be hard to look at. You know, 3D is hard to look at--2D. But I think there are ways to make 3D 2D by simply calculating modeling, if you like, a different formula. So, instead of genotype here--and I tried to make this point verbally, you would have genotype plus other predictors down here and you would come up with a formula,

age plus 2C9 given some weight, plus VKOR given some weight, and you would come up with some fudge factor.

What the clinicians are doing, they are really saying, okay, so I do the calculations on my patient and I come up with a fudge factor and it is that. That is the idea. That is the communication.

Now, I also take your point, a very good point, that this might not be linear. But what I was trying to do here was arrange the doses in a linear manner so that someone could use it, rather than having a series of things down here based on the 2C9 or the VKOR or the age. But your point is very well taken. It might well not be linear.

DR. LESKO: I sometimes wonder about how we are going to get there based on some of the things you have recommended. It strikes me a little bit like an exceptionalism with regard to genomics in this sense. I can appreciate the iterative value of adding a test, but the difficulty I see, and I am not sure we can solve the difficulty, is that it has no context. So, if I were to take, for example warfarin, I would say what is the iterative value if I consider age, or if I consider ethnicity, or if I consider something else?

Coming up with a number for a genotype is going to be very specific to the study design, the patients

studied, the drugs they were taking, the other cofactors. So, to put something in the label as a general statement about iterative values is going to have a lot of pitfalls there because it may not be relevant to my patient, depending on my patient. So, I am wondering if we don't have iterative values for other things we use in decision-making in clinical practice, say, for selecting a dose or even selecting a dosing adjustment I am not sure how we can position the perceived value of an iterative value for a genomic test.

DR. FLOCKHART: Well, I think I would say two things about that. I think actually you have provided yourselves the scientific approach to this by thinking about the percent variability. I think warfarin is a particularly difficult drug because it is used in so many clinical contexts. Others would be simpler I think. But what really speaks to me most is this is what the clinicians want most. They want what extra does it make.

I think I can see two situations, one situation where we would be very specific where a drug is used in a specific situation with a specific group of patients and it is FDA legally indicated for that thing--and we are not going to get into the off-label--in that context one could provide those numbers.

Another context where it is much harder because the drug is used in multiple contexts or those numbers are harder to come by, where one could very legitimately walk along the path you yourselves have already laid out, and that is to step back from it a bit, step back from the quantitative and say but you should consider these other things. In the case of breast cancer I would list the estrogen receptor HER2neu and all these other things, and the value of that test in that clinical setting. That is okay. But, you know, I am always going to be trying to make it more quantitative, to segue to this afternoon, but the more specific the recommendation we can make, always the better. I think there will be scenarios, Larry, where we are able to do that. Warfarin is a really hard one.

DR. LESKO: I have one more comment. Do we have time?

DR. FLOCKHART: No, but go ahead.

DR. PHAN: We could save it for discussion later.

DR. FLOCKHART: One last question? Dr. Powell?

DR. POWELL: It seems to me that we are bumping up against the constraint of a paper label in a two-dimensional format. I mean, what you are talking about is trying to force a number of dimensions into this to come up with a dose. I mean, it seems to me that in this situation having a multi-dimensional--I mean, you have what dose do

you start with and there are a number of dimensions that would impact on that, and then you have a feedback piece that has to come back. I mean, it just seems to me that it is almost like a program.

DR. FLOCKHART: Well, I think I would say this about it, it gets to usability again. I mean, one of the pluses or minuses of the paper is that it is two-dimensional. It is communicated simply as a number and a dose. In this context, in the warfarin context it is particularly hard because there is a feedback of the INR going on as well. But generically what I am trying to do is get to a situation where, whenever possible, it is possible to look at a label and say, okay, fudge factor, if you like, to dose--whatever we call fudge factor, whether it be a combination of age and two genotypes or more complicated ones later. Whenever I think we can say very simply it is this genotype, we should but I think increasingly we won't be able to.

DR. POWELL: So, my question to you is based on what you know about warfarin and wanting to diminish adverse events and speed getting to therapeutic effect, do you think that it is actually feasible to put all that information in a label that a physician can actually use to get to those endpoints?

DR. FLOCKHART: Yes. I think what is really hard actually is to ask someone to actually themselves put age in, put 2C9 genotype in and then from a separate thing put the vitamin K genotype in and come up with some gestalt, without guidance, that tells them what to do. I think that is a really hard thing. Mary?

DR. RELLING: Along the same lines, you have done a good job of saying what do clinicians want and let's listen to what clinicians say they want, but maybe what we as clinicians want isn't necessarily the right thing to do to advance drug prescribing in the next century. A perfect example, as you brought up, is the thiopurine methyltransferase labeling proposals and the rejection of a few pediatric oncologists in the committee room, which ignored the evidence and, despite their contention that dosing could be based on neutropenia, they were wrong. There are multiple drugs that cause neutropenia and they would make the wrong decision and all of the evidence points against that.

So, we have to weigh our desire for things to really happen in the clinical community with embracing the reality that, as we learn more about pharmacogenetic determinants of drug dosing, prescribing is going to get so complicated that it doesn't fit on a little piece of paper that is pasted onto the vial of drugs that are stored in

the pharmacies in the future. That is not the way it is going to happen. And, you can easily see with warfarin alone, with all of the clinical and genetic determinants of drug dosing that we know, that you will need a whole lot of those graphs to give some clinician an easy place to look and say, okay, I use this dose. So, the fact that they want it doesn't make it the right thing to do.

DR. FLOCKHART: I take the point absolutely. I think that largely we have swung way the other way though. I think we definitely need not to have information I am talking about not just be communicated and be wrong but serve as a guide also towards better prescribing. I really believe, and I know you do too, that pharmacogenomics can be used as a tool to notably improve prescribing. I think really TPNT is a very unfortunate precedent. I tried to make those points at the beginning. There is a series of unfortunate precedents we have at this point and I hope we can really go forward in a way that is not only scientifically responsible and leads the scientific community in a way that can improve rational prescribing, but also is from a public policy point of view effective.

We have a break.

[Brief recess]

DR. FLOCKHART: Shiew-Mei?

Topic 1B: Current Evidence Related to the

Pharmacogenetics of Warfarin as a Potential

Basis for Label Updates Evidence

Supporting Relabeling of Warfarin

[Slide]

DR. HUANG: We have heard earlier this morning that warfarin is one of the most prescribed drugs and also is one with among the highest reporting of adverse events. Dr. Lesko has given some background on warfarin disposition, the clinical management for safe and effective use of warfarin, and he has also shown data to indicate that there is room for improvement in the management of anticoagulation.

What I would like to do is to briefly introduce some data to support relabeling of warfarin to include pharmacogenomic information in the labeling. Right now other factors that may confound the safe and effective use, such as age, sex, body surface area, concomitant medications causing drug interactions or dietary supplements used, co-morbidities have all been in the labeling. What is missing is the genetic factor.

[Slide]

There have been many studies recently published to correlate the warfarin maintenance dose and the genotype of 2C9, which is the enzyme that is responsible for clearance of S-warfarin which is a more active form of

warfarin. As shown in this particular study, where Higashi has done the analysis of 185 patients from a pharmacy at University of Washington Medical Center where the median time of follow-up is about 500 days, ranging from 14-4000 days. They showed a correlation between the maintenance dose and the genotype. These subjects with wild type *1*1 has the highest dose. As soon as you have at least one variant allele the dose is decreased; with the subjects with at least one of the *3 allele with the lower maintenance dose. This is based on a long-term study.

[Slide]

If we look at the induction phase, here is another study looking at this Italian population where 125 patients were evaluated and a similar analysis was done, you can see that after day 4 they started to see a trend of the doses that were adjusted based on INR and, again, the wild type has the highest dose and patients with either 1*2 or *3 started to have decreased dose, again, subjects with one variant allele *3 with lower doses between days 4-24. So, these are just some examples of data correlating the maintenance dose and genotype of CYP2C9. There are other analyses correlating genotype with some other efficacy endpoints or safety endpoints.

[Slide]

For example, these are the graphs based on some of the data again in the Higashi paper where two other efficacy endpoints were looked at. Here is looking at therapeutic INR or patients reaching therapeutic INR or subjects reaching stable dosing. The way the graph was presented is the proportion of subjects without therapeutic INR. You can see that there are no apparent differences between the wild type, which is the green triangle, and the subjects with at least one variant allele CYP2C9, and that is the yellow square.

So, this may have implications for the other parameters that we look at because, clearly, the subject even initially shows no difference in the time to reach therapeutic INR will have differences in other parameters. The top curve are the subjects with at least one variant allele. If you look at 50 percent of subjects without stable dosing, between subjects with one variant allele and the subjects with wild type there is about 95 days difference. So, it takes a longer time for patients with a variant allele to reach stable dosing. In this study they have calculated the hazard ratio, a standard statistical calculation, and the ratio os 0.65.

[Slide]

When you look at two other parameters which may represent the safety parameters, it will be subjects with

an above-range of INR. So, for subjects in this study, if their target is 2-3, then the above-range would be the time when they have INR above 4. If the target is 2.5-3.5, then the data will represent when the INR is above 4.5.

Also looking at the major bleeding event, you can see that there are significant differences between patients with wild and subjects with variant with either of the parameters. If you look at the hazard ratio, it is about 1.4 for patients with one variant allele compared to patients with wild type 2C9. When you look at the proportion of subjects without the bleeding effect, if you are comparing subjects with one variant allele, at least one variant allele compared to the subjects with wild type the hazard ratio is about 2.5 within the length of time that was evaluated. The median time is about 500 days, ranging from 14-4000 days. If you look at an earlier time point, the first three months, the hazard ratio is higher. It is 3.9.

The authors indicated that the genotype of 2C9 alone is an independent predictor of a bleeding event. With genotyping we really hope that these subjects, the yellow square, eventually, if the dose is adjusted early, the curves may overlap with individuals with the wild type by taking away that genotype variability.

[Slide]

Again with the Italian population study, with the effect of the genotype on warfarin with patients with the above-range INR, and here the definition of above-range is when the INR was above 3, and this is during the induction period. The data I showed earlier is from 14-400 days and this is within 24 days. Again you see that the patients with variant alleles have more significant percent with above-range INR, which is, again, a safety measure of the warfarin treatment.

[Slide]

So, I have shown you the data of the correlation of genotype of CYP2C9 and either the maintenance dose of the efficacy measures, such as time to therapeutic INR or time to stable dosing and the relationship to safety measures, such as time when you reach the above INR level or time to major bleeding. What about additional factors? Later on Dr. Gage will discuss more about an additional gene and how that genotype will affect some of the parameters that I have just mentioned, the maintenance dose and various safety and efficacy measures.

[Slide]

But I will just cite data from one study--this was published last month--where the effect of 2C9 and VKORC1 were looked at together for the warfarin dose, the maintenance dose. So, they looked at 297 patients and this

is the distribution of CYP2C9 genotypes. You can see that within each CYP2C9 genotype there is again differentiation of the doses when you look at their VKORC1 genotype. If you look at the A/A type, which is a variant allele, which accounts for about 20 percent of the population, you can see that within the wild type the A/A genotype is lower for the maintenance dose and the trend continues for the others. I did not plot areas where we only have one or two subjects so this graph looks slightly different from the original publication. But here it shows that VKORC1 is another factor to consider when we look how to safely and effectively administer warfarin.

[Slide]

This same paper looks at the modeling of warfarin dose to look at the effect of age, and they found that 17 percent of the variability can be explained by the differences in age; 18 percent by CYP2C9 ; 15 percent by VKORC1; and 16 percent by the height. When you look at all four parameters within one equation in this modeling, it shows that these four have 50 percent to explain the 50 percent of the variability. So, genetic polymorphism in these two genes in particular, CYP2C9 and VKORC1, accounted for significant inter-individual variability.

[Slide]

So, the questions for the committee, looking at these data and later on with Dr. Gage's presentation, are that we would like to ask you does the committee agree that sufficient mechanistic and clinical evidence exists to support a recommendation that we use lower doses, lower starting doses of warfarin for patients with genetic variations of CYP2C9 that lead to reduced activity, and the same question for VKORC1.

[Slide]

Our second question will be does the committee believe that genotyping some or all patients prior to beginning of warfarin therapy will reduce adverse events and improve achievement of stable INR--these are some of the efficacy and safety endpoints--in patients with genetic variability in CYP2C9 and also patients with variations in VKORC1.

[Slide]

Finally, we ask does the committee believe that existing evidence of the influence of 2C9 genotypes warrants relabeling right now to include the information either in the clinical pharmacology, clinical studies section or other recommendations.

If yes, what information should be provided in the label? If not, what additional information is needed to provide the necessary evidence for a labeling update?

We ask the same question for VKORC1 genotypes. Again, if yes, what information should be provided in the label? If not, what additional information is needed to provide the necessary evidence for a labeling update?

[Slide]

I would like to acknowledge that for the labeling recommendation the document was prepared by a subgroup from various centers, which is a subgroup of the FDA pharmacogenomics working group, which is chaired by Dr. Lesko, and for the warfarin discussion this is the group that originally worked on the warfarin which led to the current discussion on how and when we should relabel warfarin to include the pharmacogenomic information in the labeling, and also other members of our office pharmacogenomics working group, led by Dr. Atik Rhaman who contributed to our discussion. That is all I have

DR. FLOCKHART: Thank you, Shiew-Mei. I think in the interest of having a substantive discussion later, we will go right on to Brian Gage's presentation. Dr. Gage is from Washington University School of Medicine, where he has some really important ground-breaking work on warfarin, and his presentation is entitled new insights on warfarin: how CYP2C9 and VKORC1 may improve benefit-risk ratio.

New Insights on Warfarin: How CYP2C9 and VKORC1 Information May Improve Benefit-Risk Ratio

DR. GAGE: Thank you.

[Slide]

I am going to begin by briefly discussing cytochrome P450 2C9, but mostly I want to talk about the vitamin K epoxide reductase, complex 1, the gene for that we call VKORC1. Then I will talk about derivation of a pharmacogenetics-based warfarin dosing algorithm and present a validation of those data.

[Slide]

I have shown here schematically warfarin. The S half of commercially available warfarin is taken in and metabolized almost exclusively by cytochrome P450 2C9 and that is why the *2 or *3 polymorphism in C29 results in over-accumulation of S-warfarin. What happens then? Well, then there is greater inhibition of vitamin K reductase and that inhibits the vitamin K cycle, and when this vitamin K cycle doesn't work properly because of inhibition by warfarin or inadequate vitamin K, then what results is that there is inadequate production of the functional forms of vitamin K-dependent clotting factors II, VII, IX and X and also other proteins like C, S and Z.

[Slide]

So, we just looked at these data from Higashi and colleagues. I don't want to go over them again, except to show you that there is a significant difference here

between these two curves in patients that have one of the variants, either *2 or *3. It took those patients longer to achieve stable dosing. On the right side you will see time to first serious or life-threatening bleed and those patients suffered a bleed more quickly.

[Slide]

It turns out that there has been a number of studies of CYP2C9 and we don't have time to go through them individually so I thought I would show you this meta-analysis that was published earlier this year. Sanderson and colleagues took a number of studies, and actually there are a few new ones, but they came up with essentially the same finding, that patients that have one or two alleles of CYP2C9*2 have lower requirements for warfarin. On average, they required a dose of warfarin that was 17 percent less.

[Slide]

They also looked at CYP2C9*3 and they found that the presence of that allele correlated with a 37 percent reduction in the therapeutic dose of warfarin. Again, there are numerous studies but basically they confirm these results, although this was published as well this year.

[Slide]

So, knowing that CYP2C9 is clinically relevant and affects the therapeutic dose of warfarin, let's spend the rest of the time talking about VKORC1. We know that

mutations in this gene cause warfarin resistance and multiple coagulation factor deficiency type 2. It is a very rare syndrome. We know that VKORC1 synthesizes vitamin K epoxide reductase or VKOR, which resides in the endoplasmic reticulum of the hepatocyte and other cells. I just showed you how VKOR is inhibited by warfarin, especially by the S-warfarin which is biologically more potent than the R form. We know that VKOR activity is required for the post-translational modification of Glu residues on several clotting factors. We are now sure if VKOR is part of a complex. We call it VKORC1 because it sort of leaves the opportunity of additional protein.

[Slide]

So, the hypothesis from the study that we did is that informative SNPs in this gene would correlate with the warfarin dose. We collaborated with Mark Rieder and Allan Rettie at University of Washington to sequence this gene in archived DNA from CEPH families and from the Coriell depository, and to correlate informative SNPs, that is inferred haplotypes, in 186 patients, and then to correlate these four tag SNPs and inferred haplotypes in a larger cross-sectional study.

[Slide]

If you look here, there are five SNPs that are in high linkage disequilibrium so we could have picked any one

of these SNPs to come with the inferred haplotype. Then, over here you see a few other SNPs that come up then with inferred haplotypes which we labeled H1 all the way through H9. But in terms of similar dosing, you can think of this as haplotype group A and haplotype group B. This amplification means you can genotype for a single SNP, any one of these SNPs, to figure out if patients fall in the low dose or the high dose group. In fact, that is what a number of authors have done. The nomenclature is a little different between these authors. So, what I have done on this slide is I have shown you, if you take the nomenclature that we use up here, I have given you rs number and I have also given you the name that other groups have used, and I have told you the names of the groups that have used this SNP so you can search it for yourself.

This 1639 is probably in the promoter, and this is probably the functional SNP of VKORC1. This SNP, here, is in essentially 100 percent linkage disequilibrium. Then, these SNPs here are very close to that very high linkage disequilibrium. So, that tells us that we can come up with SNPs that should correlate with the expression. In fact, that is what we have found.

So, if you look at group A, which are the lower does group, these are the frequencies as stratified by race and those patients, on average, have a lower dose

requirement. I will show you how low in a minute. The next line is the group B and those patients, with frequencies fairly uncommon in the Asian populations, tend to have higher warfarin doses.

[Slide]

What does this look like graphically? Let me just show you. So, the frequency of the VKORC1 negative 3853 allele was 37 percent in Caucasian and 24 percent in African American patients.

This was done by pyrosequencing in Dr. McLeod and in Dr. Eby's laboratories.

[Slide]

Here are the results. On the top you can see the results from the Seattle cohort and top left you see all patients, that is, not stratified by 2C9. As they have one or especially two of the higher dose alleles in VKORC1 they have higher dose requirements. So, those patients have the highest expression of this VKOR gene and they require more warfarin to achieve therapeutic dose. On average, they require about a 6 mg dose, whereas patients down here require about 3 mg or less.

In our cohort, in St. Louis, we have almost the identical results, replicating the findings, in our cohort in 306 VA patients. In the middle and right you will see the same results but now stratified by CYP2C9 so CYP2C9

wild type patients are in the middle column and CYP2C9*2 and *3 are in the far right. So, you see, yes, those patients with CYP2C9 have lower doses because the S-warfarin is not as rapidly metabolized but you still see that the VKORC1 effect is still independent of that. So, that tells us these two genes will both work by a different mechanism and both ought to be synergistic in terms of predicting a therapeutic dose of warfarin.

[Slide]

D'Andrea and colleagues had a study that was published recently in Blood. They looked at VKORC1 1173 genotype and they looked for common SNPs in that, and they found that the 1173, although it did not affect mRNA processing, that is, it was not in the promoter the way negative 6239 is, it did correlate very nicely with the warfarin dose. Here I show you the genotypes from that study, CC, CT and TT. I give you the frequency of those genotypes and I show you that the warfarin dose went from 6.2 mg per day to 4.8 mg per day to 3.5 mg per day. So, what you see is that their results are almost identical to the results that we got. Now, we used a different SNP but because those two SNPs are in high linkage disequilibrium it doesn't seem to matter. The point is that the VKORC1 can help us predict the therapeutic dose of warfarin.

[Slide]

Wadelius and colleagues in their paper in Pharmacogenomics genotyped 200 warfarin-treated patients also for common SNPs in the same gene. They found that same SNP, 1173, explained 29 percent of the variability in the warfarin dose. When they combined this SNP with the CYP2C9 *2*3 SNPs and with clinical factors they derived a regression-based dosing equation that accounted for 56 percent of the variability in the warfarin dose.

[Slide]

these are the results graphically. This is the SNP that was most employed in the study, the VKORC1 1173, and what you see here is that these lines are the median dose and they have graphed it in terms of milligrams per week. So, you just divide this by seven. You will notice that these patients are just under 3 mg per day, and over here divided by seven is about 6 mg per day. So, again, very similar results to the studies by D'Andrea and the studies that we published.

[Slide]

This is a study by Bodin and colleagues Now, they looked at a different coumarin called acenocoumarol. The half-life differs and the metabolism is slightly different but close enough that it is very insightful to study. They studied the 1639 SNP which is probably the promoter SNP in the VKORC1, and what they did is they

administered a standard dose of the coumarin to 222 patients and after that standard dose they looked and quantified the INR change as stratified by genotype.

What you will see on the far right is those patients with the AA group had a much greater INR change than those patients with the GA or GG genotype, again showing that this increased sensitivity to warfarin is genetically determined by the same gene VKORC1.

[Slide]

Here is the work of Sconce and colleagues where they derived in 297 patients and validated in an independent cohort of 38 patients a dosing algorithm for warfarin using VKORC1, CYP2C9, age and height, and they ended up with an R squared that was very similar to the 56 percent I just showed you.

On the horizonal axis is the calculated warfarin dose from the equation that they derived in 297 patients, and then the actual dose in those 38 validation patients is shown on the vertical axis and the correlation is pretty good. It is not perfect. Clearly, we will still need to monitor the INR but it is a nice place to start.

[Slide]

We performed a similar but larger study. We derived the dosing equation in 900 participants taking a therapeutic dose of warfarin and validated it in 100

patients selected at random from the same patient populations. The age of the two cohorts was similar, about 65 years. The body surface area on average was similar. Most of the patients in the study were Caucasian. A little more than half of them were male. The average target INR was 2.5. The average therapeutic dose was just under 5 mg per day and about 3-6 percent of the patients were taking amiodarone, a drug known to interact with warfarin therapy.

[Slide]

We did a step-wise dosing equation to figure out which factors were independent predictors of the therapeutic dose of warfarin. The very first factor to enter was a SNP in VKORC1 and 6853 is a step we looked at. We could have used negative 1639 or 1173 since they are in high linkage disequilibrium, and per each of those alleles we found that the warfarin dose decreased by 27 percent. The R squared after that single factor was 22 percent. That is the percent of variability that is explained by this VKOR SNP.

One thing that was interesting is that although that SNP was statistically and clinically significant in the Caucasian patients, and you see entering model number six was a different SNP, the 5808, that was statistically and clinically significant in African American patients.

The next variable to enter the equation was body surface area, then the CYP2C9 alleles, then age, then the target INR, then use of amiodarone, then race, then smoking status and finally simvastatin or fluvastatin.

So, this is the way we came up with the dosing equation. I know this seems a lot more complicated than the slide that David Flockhart showed in terms of estimating the warfarin dose, but really I think this is where we are headed in terms of estimating the warfarin dose. It is not nice and simple. It probably needs to be programmed in my palm pilot even for me to use it. But once that is done, it should be readily accessible.

[Slide]

How accurate would this be? Well, in our validation cohort we got an R squared of 56 percent, lower in African Americans probably because there are other SNPs in that patient population and we hope to study those.

Just a clinical model if you want to talk about the iterative or incremental benefit, just take out the genetics, repeat the exact same process and we get an R squared of 26 percent. So, the difference between those two is 30 percent. That is the extra added predictive accuracy that you have by taking the SNPs and two genes.

In terms of how the error comes out if we gave everybody 5 mg per day, our average mean error--this is

absolute value--is 1.79 mg; a clinically-based dosing regression would be 1.55 mg. The full pharmacogenetics model would be 1.31 mg. So, by decreasing this error from 1.79 to 1.3 mg for an initial estimated dose we ought to be able to increase the time and range. How much? It is hard to estimate. We have some models for that. It is pretty modest. It is about a 2-4 percent increase in the time to range in the first 30-90 days of warfarin therapy. That sounds pretty small, and the reason for that is that patients who have a more common genotype don't get any benefit from pharmacogenetic testing. It is a subset of patients whose dose is quite different than you would have otherwise estimated that really has the most to gain. So, for those patients, of course, the gain in time and range is much larger but on average it is probably only 2-4 percent.

[Slide]

Where are we headed with future studies? SNP discovery in targeted genes; APOE, calumenin, clotting factors II, VII, IX and X, gamma carboxylase and quantifying the relationship between new these new SNPs and warfarin dose. The reason why I mention these is just to let you know this field is a moving target. We will come up with something now but it is clear that in the future there will be other SNPs, probably not as important as VKOR

or CYP2C9 , that will help us adjust or estimate warfarin dose. Prospective validation of a pharmacogenetics dosing model would be helpful. To do that, we would need a platform that could very rapidly and inexpensively genotype patients for the SNPs of interest.

[Slide]

So, in conclusion, the maintenance warfarin dose can be estimated from clinical and pharmacogenetic factors that can be obtained at the time of warfarin initiation. At least half of the variability in the warfarin dose can be predicted from regression modeling using two genes, the VKORC1 and CYP2C9, in conjunction with clinical factors, but not dosing algorithm incorporating these genes has been prospectively validated.

So, the relationship between SNPs and these genes and the therapeutic warfarin dose, although biologically and statistically compelling, is not yet validated in terms of benefit for outcomes. My best guess estimate is about a 2-4 percent increase in time to therapeutic INR range in the first month or couple of months of therapy, with probably less benefit thereafter when the therapeutic dose is figured out. After one month or two the pharmacogenetic knowledge could still be helpful and it would allow for more cautious dose escalation in patients that are poor metabolizers. That is, patients that have a *2 or *3

polymorphism in CYP2C9, because they metabolize S-warfarin more slowly, probably need a more cautious dose escalation. I will stop there. Thank you.

DR. FLOCKHART: Thanks, Brian. Let's move right along. You mentioned, if I could segue to Dr. Caldwell, the importance of thinking about prospective studies and Dr. Caldwell is the medical director at the Marshfield Clinical Research Foundation where they have a huge patient database that they have had for many, many years and, as a result, we have looked to them really for a large study in this respect. Dr. Caldwell?

Commentary on Current Status and Next Steps with Integrating PGx Information into Safe and Effective Prescribing of Warfarin

DR. CALDWELL: Thank you, Dr. Flockhart.

[Slide]

My task was to comment on the current status and next steps with integrating pharmacogenetic information into safe and effective prescribing of warfarin, and the approach that I took was to review the literature based on genotype and warfarin dosing, and to give you basically a timeline which I have passed out for you and that I have here, on this slide, knowing that it wouldn't project as easily as if you had it in your hands.

Much of what is listed on here has already been said so I am not going to try to be redundant of many of the studies. I will point out that probably the first known study linking cytochrome P45 2C9 genotype to warfarin dose was done in 1995. Subsequent to then and up actually really until about 2003 or actually to 2005 with the meta-analysis of Sanderson. At that point, there were about 11 studies that had been done that Sanderson shows for the meta-analysis. Nine of those 11 studies, 2,775 patients--the data from that have already been presented to you twice, showing that patients with *2 and *3 genotypes have lower mean daily warfarin doses and a greater risk of bleeding.

Subsequent to that, and what I think is clear from reviewing all of these studies, is the fact that they have all been retrospective by and large. There are another five or six studies after Sanderson's meta-analysis, again retrospective studies, again corroborating and validating the same data.

There was one prospective study of perioperative pharmacogenetic-based dosing of warfarin that is by Voora, in 2005. These were orthopedic patients who were having hip replacements. There were no controls. All of the patients in the study were genotyped. As I recall, it was about 38 patients. There was no standard treatment group

to compare to. What they demonstrated was that the patients achieved a stable therapeutic dose quickly. They still, however, had a risk. Those that had the mutant alleles still had an increased risk for adverse outcome, which they defined primarily as an INR of greater than 4.

Subsequent to that, and with the description of VKORC1 and demonstration that its mutations cause warfarin resistance, in 2004, we have had studies beginning to include VKORC1 and 2C9 as a part of predicting what would have been the patient's stable dose, initially Wadelieus showed that you could explain 29-30 percent of the variance by adding the VKORC1 versus 12 percent only with CYP2C9. We have, as I will show you in a moment, some data that are very similar to exactly what Dr. Gage has just presented to you as far as a comparison with VKORC1 and CYP2C9 and warfarin dosing.

One of the issues that came out of the only prospective study that I am aware of was the question of whether or not genotyping could be done in a clinically relevant manner. In other words since, indeed, all of these studies have been retrospective genotyping was relatively easy to do after the fact and it was a post hoc analysis, but the question was can you actually genotype prospectively quickly enough to get the information to physicians so they can make the appropriate decision in a

relevant time span. We published a study this year showing that you, indeed, can genotype within about four hours for cytochrome P450 2C9 and get that from the time the blood is drawn to the time the information is back to the physician in about a four-hour turnaround. So, we feel that you can, indeed, do that.

The data that Dr. Gage just presented is now beginning to predict that the combination of age, cytochrome P450 2C9 and VKORC1 genotypes can explain about 55 percent of the variance in warfarin dose.

[Slide]

I will show you a little bit of our data. We really are primarily interested in initiation of warfarin therapy and being able to predict the final stable dose for a patient. Instead of doing what most physicians do at this point, which is pulling a dose out of the air or taking a standard 5/5/5 mg approach for 3 days to use something a bit more helpful.

[Slide]

So, this shows you the approach if you initially just using stand clinical data as far as being able to predict the dose. Age, body surface area, whether or not the patient has a cardiac valve replaced, and male gender even though I think that is variable in our studies and it has certainly not held up in others as well, predicts a

small portion of the stable dose of warfarin but, as you can see, there is a huge amount that is unknown.

Speaking of graphics, these are the kind of graphics that initially translate I think to practicing physicians as a way of being able to understand this process. So, clinical data--and this is what I think Dr. Flockhart was alluding to--what does genotype add to standard clinical data itself?

[Slide]

This is stable dose of warfarin as predicted by genotype from our data from 438 patients. As you can see here, and as has been seen in other data throughout the day, there is an association of genotype with stable dose of warfarin.

I want to stop for a moment and just point out what I think is obvious to you when you look at this slide, and that is that the variance around the mean for the wild type is considerably greater than it is for the mutants. I think that if that holds up in other studies and as others look at their data, I think that holds some merit for the way that we think about monitoring these patients, and the conclusion that I draw from that, whether it is valid or not, is one that comes from some inductive reasoning and that is that if you reduce the enzyme activity of 2C9 down to 10 percent and if you are measuring the effect of the

enzyme activity by something that is a bit distant, such as INR, you can pick up a change in INR much more easily when you are working at 100 percent of enzyme activity than when you are working at 10 percent of enzyme activity. IN other words, if you cut your enzyme activity in half, when are at 100 percent activity you are now working at 50 percent. If you cut your enzyme activity in half and you are only working at 10 percent, you are now working at 5 percent and INR is much more likely to pick up the former than the latter.

If that bears out, and I think it is worth considering, it may actually suggest that wild types may need to be monitored in the long term once they are on stable doses of warfarin more closely than with alleles. Alleles may be more important for initial dose but ongoing monitoring may be more important for wild type. Regardless, genotyping would be important in the process.

[Slide]

In your booklet that was handed out for this meeting you have the results of our prospective study, a very small study looking at prospectively prescribing warfarin using genotype as part of that process. This was the study that we basically did to see if, indeed, you can genotype in a clinically relevant manner.

The size of the study was not enough to really come to conclusions regarding complications of warfarin therapy, but these data, which are shown in your pamphlet--I gave you this slide because it is important. The bottom of the slide talks about blue lines and red lines and yours is xeroxed in black and white so I want to give you the real slide that shows you what is going on.

In this case, these are all patients with at least one variant CYP2C9 allele, and looking at a difference in final dose between the model dose, which is the open circles, and the dose of 35 mg per week, which would be 5 mg per day as a standard dose. You can see that the basic take-home message is that for most of these patients with variant alleles the model dose is much better at predicting their final stable dose than giving them 5 mg a day as the standard dose.

[Slide]

So, CYP2C9 adds to the process considerably and reduces the amount of unknown stable dose of warfarin. VKORC1--we have also done similar genotyping and overlaid the VKORC1 changes in genotype here with the genotypes for cytochrome P45 2C9 and, as you can see here and as you have seen in other slides today, particularly at those genotypes that are either homozygous or heterozygous for wild type,

VKORC1 seems to have a significant effect on the stable coumadin dose by genotype.

[Slide]

If you look at our model which includes things such as age plus whether or not you have had a cardiac valve and cytochrome P450 2C9 genotype, and then if you look at it with or without VKORC1 to see what the R squared is as far as your ability to predict doses it is, just like what everybody else is seeing, about 56 percent that can be predicted by adding VKORC1.

[Slide]

So, in a surgeon's simple way of looking at it, we are now explaining over half of the stable dose of warfarin from that approach.

[Slide]

My conclusion from looking at the literature is that at this point, using dosing algorithms for warfarin that include at least age, cytochrome P450 2C9 and VKORC1 genotypes should more accurately predict stable therapeutic dose and should reduce complications, and the genotyping ca be done with clinically relevant turnaround times.

[Slide]

The issues that I think are still left--some of the issues that are still left are that we need further definition of the genotypes non-Caucasian ethnic groups so

that we can ultimately develop a standardized panel which could be used across the general population. Then, we would need to make certain that the timeliness of that assay is still clinically relevant.

A cost benefit analysis is needed I think as well, to get back to Dr. Flockhart's nasty relevance to reimbursement but that is one of the key realities that we live with, and that is that unless it gets reimbursed it doesn't get incorporated into the clinical practice.

[Slide]

I suggest that the next step should include a prospective trial which predicts stable dose; examines the complications which would have to use surrogate markers to get the number of individuals small enough to be able to feasibly do this in an affordable way; and examines the cost-benefit. I think this prospective trial needs to be done as soon as possible, the reason being, and perhaps this is controversial and, if so, fine, let's talk about it, but I think soon, if not already, it may not be considered ethical by institutional review boards to do a study that has a cohort in which genotyping prior to dosing is omitted. Clearly, that is the standard of care but, clearly, if you read the literature the effect of particularly cytochrome P450 2C9 genotyping, its relevance to the ultimate stable dose and its relevance to

complications is such that IRBs I believe are going to start calling that into question. Those are the major comments that I have.

DR. FLOCKHART: Thank you, Dr. Caldwell. Again in the interest of having a substantive discussion, we are going to postpone most of the discussion until after the open public hearing so that we have all the information on which to have a productive discussion. To begin that, we need first to have the guidelines for the open part of the meeting read:

Both the FDA and the public believe in a transparent process for information gathering and decision-making. To ensure such transparency at the open public hearing session of this advisory committee meeting, the FDA believes that it is important to understand the context of an individual's presentation. For this reason, the FDA encourages you, the open public hearing speaker, at the beginning of your written or oral statement to advise the committee of any financial relationships that you may have with any company or any group that is likely to be impacted by the topic of this meeting. For example, the financial information may include the a company's or a group's payment of your travel, lodging or other expenses in connection with your attendance at this meeting.

Likewise, the FDA encourages presenters at the beginning of their statement to advise the committee if they have any such a financial relationship. If you choose not to address this issue of financial relationships at the beginning of your statement, it will not preclude you from speaking.

Now, in this context, I would like to invite Dr. Hu, from IMS Management Consulting, to come up and make a statement about the industry context in which discussion is being held. Again, this is the public part of the meeting. Dr. Hu, could you clearly state your name for the record?

Open Public Hearing

DR. HU: Shiang Hu is my name. First, I would like to thank the committee for--

DR. FLOCKHART: I am sorry, we will need your affiliation also.

DR. HU: Yes, this is Shiang Hu. I work for IMS Health, Management Consulting Division. I would first like to thank the committee for allowing me this opportunity to voice some of the comments, brief comments I have.

In terms of conflict of interest, the division that we work in, which is management consulting, otherwise in general we advise pharmaceutical companies related to strategic decisions for their product commercialization.

So, we are constantly involved with or engaged with various pharmaceutical companies all the time.

We are talking about pharmacogenomics in clinical studies, and I realize that the results, although different from pharmacogenomic studies, can vary a lot, for example, in terms of number of trials; the size of the trial; the power of the identified correlation or association; and also maybe the percentage of the patient response that could be explained. So, what is not totally clear to me is what are the criteria that should be used in reviewing all these different trial results before the committee or FDA could decide whether or not the trial information should be or could be included in the label. So, that is the first area of comment.

The other one is to realize the benefits of pharmacogenomic study results, certainly it is necessary to have corresponding tests, clinical tests and to make them available to clinicians. What is also not entirely clear to me is for all the tests different specificity and sensitivity--how good those measures have to be before those tests could be considered to be acceptable? So, those my comments. Thank you.

DR. FLOCKHART: Thanks, Dr. Hu. I think next we have Dr. McLeod--I don't think, I know that next we have Dr. McLeod for ten minutes, just from his perspective again

on warfarin. Then we will move to the committee discussion of questions. We are about five minutes behind. The discussion is so important that I will actually delay lunch for ten minutes.

Commentary on Current Status and Next Steps with Integrating PGx Information into Safe and Effective Prescribing for Warfarin

DR. MCLEOD: Well, since I am wasting away I wouldn't want to delay lunch too much! One of my chins may disappear! I want to make a couple of quick comments based on what I have heard this so far. One is that the data that has been shown has been very consistent, and it has been consistently seen across ethnic groups. Some data that wasn't presented, most recently from Hong King and Taiwan, show that the VKORC1 CYP2C9 genotype relationship with warfarin is something that is not just unique to the Caucasian population or the African American population, but also is being seen in the Asian populations. That is a unique finding in that some of the previous pharmacogenetic variables that have been included in the package insert have been looked at in exclusive populations and not had this wide finding.

Secondly, I think that the data really reflects the state of the art of anticoagulation in that the data that has been presented was mainly retrospective data,

small amount of prospective data, but all from anticoagulation services. So, if you look at where most warfarin is dosed, unfortunately, most warfarin is dosed in a less than official way. In other words, expert services are not being used to dose warfarin in many instances.

So, when we look at the data that has been presented, remember that this is value for genetics in the context of expert anticoagulation. Think about how valuable this might be in the context of the regular prescriber that is trying to manage this in the context of a lot of other aspects of medical care.

I think that clinical need is really highlighted in those findings, and that everyone is great at dosing warfarin when you talk to them about genetics. In talking to a lot of folks that dose warfarin, the first thing they say is, well, we don't really need genetics; we are really good at dosing warfarin. Really there is not a problem. But I have been having issues with this patient and that patient and pretty soon virtually every patient needs something better than they currently offer. I think the reality is that patients are currently not receiving adequate therapy. Genetics has some value. There are still a lot of unknowns, as was highlighted by Dr. Caldwell, but still some value with the genetic information.

The next point I want to make is that INR optimization is sufficient. I think that we all would like to have prospective data demonstrating that fewer people bleed in their head or their gut by genetic-guided therapy. That would be a great thing to have and, hopefully, in our lifetime we will have that information. But the reality is that the development of all of the guidelines for warfarin have been done using INR as an endpoint. INR is sufficient for all these large clinical trials that are changing the face of the management of patients in diverse medical areas. Why is it not sufficient as a biomarker in the context of the points that we are making with pharmacogenetic markers?

Another point I want to make is that Larry highlighted that the need for action is sufficient. Even if we don't know the exact dosing guidelines, the exact milligram that an individual patient must need, the fact that there is an issue is sufficient. We need to at least point out that there is a problem as we work towards a solution. Both Brian and Mike highlighted, and David, that this is a very dynamic field. The algorithm that we come up with today will be changing tomorrow. So, knowing that there is a problem is a step forward. Knowing what to do about the problem is even better. But knowing that there

is a problem is adequate in terms of defining safe and effective medicines.

There are a number of unanswered questions and there always will be a number of unanswered questions with every example that this committee faces and any other FDA committee faces in terms of this context. I would just hope that we don't let the unanswered questions keep us from progressing forward. Unanswered questions include specific dosing schedule, not just what dose but at what frequency patients should be receiving warfarin for this specific example. But don't let progress be held up because we don't have a 100 percent prediction.

The last point I want to make is that the label does not equal a dosing algorithm. Having a label that highlights issues, putting the information in the proper places--David highlighted that--is a step forward. Ultimately, we want good algorithms for guiding care but I am not sure that is the remit of the FDA, nor is that going to be the remit of a static document such as the package insert. So, I think we need to do the best we can with the current information and nothing less. Thank you.

Committee Disease of Questions

DR. FLOCKHART: Thank you, Dr. McLeod. Shiew-Mei, could I ask you to come up and walk us through the specific questions for the committee because I think that

is better done with the slides of the questions, and we may need some back and forth here between slides?

DR. HUANG: The first three questions are related to our background information about our recommendations on how, where and what information about pharmacogenomics should be put into drug or device labeling.

So, the first question is does the committee agree with our labeling recommendation, as delineated in the document of background information for topic 1A, in particular, those related to metabolizing enzymes? If you need a reminder?

DR. FLOCKHART: Yes, I think so. So, the specific part of the text that we are talking about, why don't we jump to that from Shiew-Mei's presentations so people know what we are talking about?

DR. HUANG: Our recommendation is clinical relevant information in the clinical studies section or clinical pharmacology section. Then, depending on the implications of this information of safety and effectiveness, they may be in different sections. When they result in the first five sections, as listed here, we would propose that we also put it into the highlighted section. That is where we have our final rule and associated guidance published about physician labeling. Do you want to see more?

DR. FLOCKHART: No, that is good. Let me ask a couple of questions. What we are getting to in this discussion is where stuff goes and I think this is important. So, in this context, Shiew-Mei, could you or Larry talk a little bit about how the FDA normally interprets the term important implications for safe and effective use? You know, it is kind of like talking about clinically relevant. You need some kind of guidance as to how you determine that. What determines whether something has an important implication for safe and effective use?

DR. HUANG: In general? The important implications? That is how we try to use examples in our document. For example, if it is important that you can only use it in subjects with a certain genotype you must have a test before you do it. Unless you are asking about the data behind making this conclusion?

DR. FLOCKHART: Well, I am just talking about in general. If we are going to use this in general for lots of situations--right?

DR. HUANG: Right. But then under each section we have described specific situations when you would put it in this section specifically about genotypes. We would consider genotype to be maybe very similar to other specific populations like hepatic impaired, renal impaired, a subject using certain concomitant medications, pediatric.

But depending on the recommendation, the actions that we would make based on those data, then we would put it into different sections.

DR. FLOCKHART: Yes. I guess everybody has the general gist here that I am trying basically to improve prescribing by getting this information on genetics in particular. I could make the same argument for renal higher up, but I would not argue against having this information in detail in the clinical pharmacology and clinical studies sections at all. But I will get off my platform and entertain points from the other members of the committee. Yes, Dr. Lesko?

DR. LESKO: Dave, to get to your question, I think there are sort of two dimensions to the question. One is, is the information that we are talking about putting in the label descriptive or, secondly, is the information that is going to the label actually--

DR. FLOCKHART: Usable, yes.

DR. LESKO: It is all usable in my opinion, but it depends--

DR. FLOCKHART: Actionable. Actionable is a better word.

DR. LESKO: It wouldn't be in the label, we don't think, if it wasn't usable. But I think it is a matter of evidence that substantiates what goes into the label. On

dosing specifically, I think you would imagine that a recommended dose would be derived from, let's say, Phase III randomized, controlled trials where frequently one or two doses may be studied. You then begin to ask the question how were doses in that trial adjusted for different cofactors that may have some uncertainty in the required dose.

So, I was thinking of your fantasy dose graph where you have on the Y axis the dose and on the X axis covariates--let's call them--A through J. Typically what we would do is to take various covariates, whether they be age, renal function, drug interactions and I might even add genotype, and look on the Y axis the area under the curve increase or decrease in that subpopulation of special population. That might come from a specific pharmacokinetic study that was conducted. Then, if we coupled the area under the curve increase with what we know to be the dose response for efficacy and safety, or maybe a PK/PD relationship for some adverse event, we can draw a horizontal line across your fantasy graph and say a threshold exists above which you need to increase the dose or decrease the dose. That dose, in and of itself, is not tested prospectively in any subset of the population. Rather, it is an assumption that exposure leads to dynamics and dynamics leads to clinical outcome. That is an

inherent assumption. So, the dosing recommendations that go into dosing and administration for pediatric or for elderly or for renally impaired patients come from that type of analysis of the data.

DR. FLOCKHART: That is very good. That was very clear. So, that is how you would approach putting things in that section, yes. We have to welcome Dr. Kearns to the committee. Dr. Kearns, could you introduce yourself quickly?

DR. KEARNS: Good morning. I am Greg Kearns, chief of clinical pharmacology at Children's Mercy Hospital in Kansas City, and also professor in pharmacology and pediatrics of the UMKC School of Medicine.

DR. FLOCKHART: Thank you, Dr. Kearns. Dr. Singpurwalla?

DR. SINGPURWALLA: Thank you. I just want to repeat a question of clarification and then give you my reactions. The labeling is for the benefit of both the clinician and the patient. Is that correct?

DR. FLOCKHART: Yes.

DR. SINGPURWALLA: Well, here is my reaction to the question and it has nothing to do with statistics. I do have questions on the statistics that are being presented but these are reactions as a patient occasionally. I am against the idea of excessive labeling.

It creates information overload. The consequence is that critical information will be ignored. It is like getting a set of instructions to replace a plumbing shaft in your toilette tank. There are 50,000 instructions so you ignore it and you start doing it yourself. Excessive labeling will also increase the risk of law suits towards the physicians. There are too many labels. Lawyers will read them--and I wrote this before our chairman's comment. Excessive labeling will also create advantages to the educated elite versus those who do not have the ability to read these things.

So, my suggestion is that only labeling which pertains to adverse reactions should be put on the medicine, and the other information should be made available to clinicians by means of bar codes where the information on the medicine is coded in a computer and the physician has access to it either through hand-held devices or through larger devices. One should take better advantage of computer technology in this context than what you have been proposing.

DR. FLOCKHART: Thank you. Other points? Dr. Relling, you look like you were going to have a point?

DR. RELLING: I guess I vehemently disagree with the concept that information should be withheld from the label because it is complicated or because it gets an

advantage to people who can read. I mean, I think the purpose of the label has to include summarizing what the scientific community thinks, for which the public is subsidizing those results, and has to be made available. You are right, I mean the use of technology should be enhanced but I believe that it is not necessary to create any new technology or new bar codes. People just have to get used to reading package insert material on-line and, by searching for key words, they can improve their prescribing without having to read reams and reams of information. That is one of the advantages that the computer provides for us in prescribing.

I guess I had a question for Dr. Caldwell, and maybe this was sort of answered by Dr. McLeod. What do you think the purpose of conducting a prospective trial is at this point? I wasn't sure what the purpose of that is because I think Dr. McLeod eloquently illustrated that we probably already have sufficient information using INR as the endpoint to indicate that this genotyping information can be a useful adjunct to dosage individualization.

DR. CALDWELL: The reason I made that comment is because all of the data, by and large, that have been collected to date are retrospective. My contention is, particularly from the standpoint of being able to understand the cost benefit and the efficacy of

pharmacogenetic prescribing, that we need a prospective trial that allows us to do this in the real setting, and to really be able to understand the practicality of this process; its cost benefit analysis. And, I am very concerned that unless we have those types of data--and I acknowledge and think, quite rightly, that surrogate markers such as INR greater than 4 is an appropriate marker. We don't have to wait until people bleed in their head. But a prospective trial that looks at that is needed. We really have none. We have one that I know of that was the orthopedic trial I described, and we have one that we did which was 28 patients and was under-powered to really make the case for whether or not we reduce complications and how predictive we are of the ultimate stable dose.

DR. FLOCKHART: I am nervous that we are getting a little bit off the subject here. This is a very important discussion about whether or not we have a prospective trial. I think it is in the context of the kinds of data we would need to have for this and I would just like to emphasize that point. Dr. Barrett first and then Dr. Relling.

DR. BARRETT: Regarding the labeling, the labeling has a necessary redundancy and connectivity that we have come to know, but with this topic in particular I

think, Dave, you brought up the idea of the iterative value of incorporating the genomic data as part of the label, which I also support. But one of the things that I think needs to be clear there is that there does need to be a section--you called it highlighting, or whatever, in which very clear guidance is synthesized so that you have information that involves the most current information and recommendations are provided in a clear context, and in a section that people become familiar with as far as looking for the synthesis of that information.

I mean, one of the things that strikes me, and I guess it will come up this afternoon is, you know, if you had to develop warfarin in 2005 what would that label look like even with the backdrop of some of the historical information but in terms of what you would do prospectively.

DR. FLOCKHART: Good point. Dr. Relling?

DR. RELLING: Yes, I guess another thing I was wondering is exactly what is included in the labeling for warfarin right now. But I believe that for warfarin and for many other drugs relevant to this discussion there are suggestions about dosage individualization based on things like renal dysfunction or liver dysfunction, for which there are no prospective trials to validate that doses with and without renal dysfunction makes a difference. I

believe if that is the standard to which we are going to hold every pharmacogenetically-based decision, we will all be dust in another universe by the time we have any labels changed. So, I think in cases like this where there are clear-cut retrospective data, and this will be the case for many, many covariates that are included into labeling information, for dosage in particular it is not going to be feasible to conduct a prospective trial nor, as was pointed out, ethical to conduct a prospective trial for every one of these adjustments to labeling as we get better and better at determining a million genetic covariates.

DR. FLOCKHART: I couldn't agree more. I think that increasingly we will have information based on retrospective data, and I think it is clear we are holding genomic testing to a much, much higher standard. We are asking for a much higher bar. But, having said that, if renal testing were put up for the first time now it might be seriously examined, and we are looking at doing this the first few times. First Dr. Caldwell and then Dr. Kearns.

DR. CALDWELL: Thank you. I just wanted to be really clear. I see a difference between the necessity to change the labeling currently and the need to ultimately have a prospective trial. I think they are asking different questions. I believe if we wait much longer before we do a prospective trial and about the ethics of

being able to do one, that speaks to what I feel strongly is a need to change the labeling currently. I think that there are data that will be extremely helpful to us in the ultimate delineation of the proper doses of warfarin for our patients that a prospective trial will help us do, and I think it can be done in a small enough number of patients to make it reasonable.

DR. FLOCKHART: Just to clarify your point a little, in the long run it would be more germane to reimbursement, which we are not here to discuss. I mean, normally were a prospective trial to be done that convincingly demonstrated the value of this, that would influence clinical practice and it would influence reimbursement unquestionably. But right now we are discussing what evidence would be required to get something on the label and I think, as Dr. Relling has pointed out, the history on that is that there are many, many things that have valuably been put in the label without organized, randomized prospective trials.

DR. CALDWELL: And I don't dispute that at all.

DR. FLOCKHART: Dr. Kearns?

DR. KEARNS: Just a comment and a question. I think certainly Dr. Relling is right. Information is good. Knowledge is even better and understanding how to use it is

the best option. And, all the information we can put in the label helps us at some level do that.

But I have a question for the agency. In pediatrics over the last five or six years we have done a lot of studies. We generate a lot of information, a small amount of which is put into labeling. Much of the information generated from those studies was not necessarily indication driven but, yet, it doesn't appear in labeling. If the agency were to take the position of putting pharmacogenetic or genomic information in the indications section of a label how might this impact studies in special populations like pediatrics who, by law, are required to study the drug within the context of the adult indication? And, if some of the stuff got in there kind of willy-nilly might make life very difficult, if not impossible, for kids.

DR. HUANG: I think Dr. Shirley Murphy would like to address this question.

DR. MURPHY: Thanks. Hi. I am Dr. Shirley Murphy. I am the deputy director of the Office of Counter-Terrorism and Pediatric Drug Development. I think your comment is a very important one. Let me just tell you a few things that we are doing to get negative data into the public domain.

First of all, through BPCA there is a disclosure requirement that any drug that gets exclusivity has to have the clinical summaries posted on the web. We now have 56 summaries of all the clinical trials posted on the web.

Second of all, our own advisory committee, of which I don't think any of you are members, recommended to us a year ago that we start putting this negative data that you so wisely suggested into labels. So, we have been doing that. If you look at the more recent labels, you will see in the clinical trials information that even though the studies are negative, you will see the results of the trials in the label.

So, I think your comment is a very important one and we have heard it at the agency. As for new genomic data, I think it would have the same impact and we would continue along this path of when children are used in trials and study that the data does get into the labels.

DR. KEARNS: If I could just respond, I think when children participate in studies, as opposed to being used--

DR. MURPHY: I know that was the wrong word. Thank you for correcting me.

DR. KEARNS: --as well, I would posit, Dr. Murphy, that not all the data are negative. I mean, there are many positive--

DR. MURPHY: Oh, correct--

DR. KEARNS: --pharmacokinetic information, pharmacodynamic information, some pharmacogenetic information that is out there but in instances has not made it into the label because it was done in studies that may not have necessarily been driven by an expanded indication for children. That is why I get back to the word "indication" as it relates to what we are discussing here today. What does that mean in the context of drug development and regulation?

DR. MURPHY: Well, your point again is well taken and, as you well know, the agency has not wanted to give medications an implied indication by putting pharmacokinetic data in the label when the actual clinical trials are negative. But, again, that is being re-looked at. If you look at Arava, a drug for rheumatoid arthritis where the clinical trial was negative the pharmacokinetic dosing information is in the label, as well as the clinical trial results because it was felt that that is a drug used by specialists and the data needed to be in the label. So, it is really still a case by case basis but I think we have heard you, Dr. Kearns, and others and we are moving forward slowly.

DR. FLOCKHART: I think, Dr. Kearns, if I could summarize a larger point that is very, very relevant to

pharmacogenetics. We have a specific subpopulation of pediatrics where there is a large amount of information generated relatively recently. It gets to this point of what does go in the label, and I think, Greg, you have appropriately focused it on the indications section. I understand the FDA's sensitivity to the indications section. Expanding an indication is a big deal.

On the other hand, I think in many, many contexts here we would be doing something a little bit new, and that is more narrowly defining an affected and a potentially hurt population and, thereby, improving therapy by doing so.

But I think we have had a good discussion of this and if I could summarize, useful information, period, would be included in the section and I think Dr. Lesko gave a quite elegant description of why that would be done and how we would define it. So, if we could move on to the next part of the question?

DR. HUANG: Yes, the second question is what is the best way to present genetic information in the labeling for use? We have shown you examples where there is not a consistent way of presenting data. We have labels where phenotypic information is given, where specific alleles are given but we don't have anything for CYP2C9 yet but we do have specific alleles for UGT1A1. Shall we consider giving

nucleotide changes? Along the line of if we give phenotypic information, specific alleles, nucleotide changes, do we also provide ethnic racial prevalence information in the labeling?

DR. FLOCKHART: Thanks, Shiew-Mei. It is open for discussion. Dr. McLeod?

DR. MCLEOD: Since I neither raised my hand nor indicated I wanted to comment, thank you for calling my name!

[Laughter]

I think that the most thorough information that we have is the most appropriate one in these instances. So, for example, with CYP2C9 if you are going to do allele specific information you might as well add a few more pages to the package insert. Whereas, the poor metabolizer phenotype, even if derived through genetic means, would be more explanatory. In other cases where there is very clear information, such as CYP2C9 or where there are the *2 alleles that have been evaluated at the exclusion of most of the others, there is very specific information. The more detail you have for small cases, the better. So, for CYP2C9 C is going to be a great example because you can put very specific information in. In cases like 2D6 A is going to be the best example. I don't think B is ever going to

be the right example because really B and C are equal. I think basically B and C are the same thing in my mind.

In terms of section D, I am of the opinion that race and ethnicity is a surrogate marker of something objective and that it only has value when there is no objective information available. There have been cases recently where race was used as a biomarker and that was all that was available and, therefore, it was included in the package insert but certainly in the cases we are talking about I think race and ethnicity is too hard to define and, therefore, should not be used.

DR. FLOCKHART: Thank you. I think my point about this would be not to shy away from a fifth category, and that is a category which includes combinations of these things and other factors. That is partly where I am going with this with the idea of the iterative value of the test. But there are many, many situations like warfarin where clouding physicians with lots and lots of genetic plus other data in different categories stretches their ability to understand it. I don't think that withholding it from them but I think it also needs to be packageable in some kind of fudge factor or predictive thing so it makes it clinically usable. We will have Dr. Singpurwalla, Dr. Barrett and Dr. Sadee in that order.

DR. SINGPURWALLA: I think this question is also connected in some sense with the first question. But I have two points to raise. Point number one, the statistical evidence that has been presented in some of the talks shows an R squared of about 50 percent. If the basis of moving in this direction is that particular evidence, I consider 50 percent to be a very small explanation of the variability so I would be very hesitant to move along in this direction based on that evidence.

The second thing, and I think our chairman has said it in kind of a different way, there is a very important result in information theory which simply says that the utility of information is concave. That means the more information you have the less valuable it is. Therefore, I must disagree with some of my colleagues who are advocating excessive amount of information that it goes against the theory.

DR. FLOCKHART: Dr. Barrett?

DR. BARRETT: I wanted to put in a plug for race and ethnicity being part of this too. I would hate to see that left behind, even Dr. McLeod's statements notwithstanding. I agree that it is probably less informative when you have good genomic data, however, people don't necessarily look at themselves as being a certain genotype, and where you don't have that information

I think it is informative to the level of the patient to have some idea of the likelihood of them falling into one of these categories based on their racial or ethnic background. Likewise, for the physician seeking to make those kinds of associations that information is actually very reasonable. But it has to be put in the right format. I would definitely agree with that in terms of the ranking of that information. My only point to all of this would be that there is some standardization that occurs and perhaps ranking of the information as it becomes more available.

DR. FLOCKHART: Dr. Sadee and then Dr. Lesko.

DR. SADEE: I would like to come back to the point also that to me the most important piece of information is that we do know that, for example, there are poor metabolizers or there are high responders or low responders, and that is in part genetically determined. So, that needs to be stated because it is the most easily understood and it allows us to judge as to how careful one has to be, whether a standard dose may do it or whether there is a need for consideration of actually dividing the patients into subpopulations.

I just heard the statement that for 2C9 we have very precise information because we have two alleles that appear to be predicting the metabolic activity. Well, I disagree and I think I would like to turn the question from

what can we explain to what have we not explained. If you look at the dosages for warfarin that are in the so-called wild type population, which is the majority of people, the spread is tremendous. There is a huge amount of variation and I think the most important question here is, that we are not addressing, is that spread caused by genetic factors or environmental factors. If genetic, then we should be addressing this. We should find out what it is. In fact, for 2C9 the two polymorphisms that we do know account for we don't know what fraction of the genetic variability in that gene that those two alleles account for. We don't know that. I think we should find that out.

DR. FLOCKHART: Point taken. Dr. Lesko?

DR. LESKO: So, for the question of information in the label, there are two points I thought maybe I would ask for comments on. The first point is, again recognizing that many of you on this committee use these tests in a clinical practice setting, so the question is how closely would it benefit to have the drug label and a test label aligned with each other in terms of the way information is being reported out or, to add a third dimension to it, how laboratories report out the information, say, in your clinical setting? So, if laboratories are reporting out information in a certain way, does that map to the label in

the same sort of language? That would be sort of one area I would question.

The other is when we know negative information about the effects of an allele do you consider this worth putting in the label so that if I know that, for example, 2C9*7 has no implication with regard to reduced activity, albeit perhaps on limited data, is that worth putting in there in the event that technology sort of catches up and can report on many different alleles, not just for example 8@ and *3?

DR. FLOCKHART: If I could respond first to a number of points, I am really trying to focus still on how to do this but you have expanded the question a little beyond, Dr. Lesko, to including devices within this. I think there are two subjects here. I think the first is this percent variability point, and I would make a statistical point to Dr. Singpurwalla to expand on your point about 50 percent being small. Fifty percent of one is small; 50 percent of 2000 is not small. So, if you have a huge amount of variability, 50 percent of it can contribute very significantly in the minds of prescribers. On the other hand, if you have something that is very tight, where dosing doesn't make a huge amount of difference, then 50 percent of that--you are right, it isn't a big deal.

In fact, if I could carry on, germane to that I think is this little dialogue that we are having here about race between Dr. McLeod and Dr. Barrett. So, I am going to ask Dr. McLeod a question. It is possible that there are settings where a genotype is very common in one population and hardly present in another where one might recommend genotyping more intensely in the very common population than in the other. In that sense race is helping as a practical economic predictor of who it is worth genotyping. I take the point that it is a surrogate and it is self-reported, and all of that. But would you take the point that it also might be clinically useful in that kind of setting?

DR. MCLEOD: I think if there is clear evidence where you can enrich for utility and it has been reproduced, then I am okay with that. I think the problem is if you say we only genotype one particular group, that means that you have looked at all the other groups and we don't need to genotype them. The second most common group in the U.S. right now are Hispanics; the third most common are African Americans. Yet, there is hardly any information on genotype, as well as most other variables, in the Hispanic population and yet they are a very commonly occurring group. And, there is not enough data on African Americans as well, but at least there is some.

The problem I have with race is more that there is no such thing as race. I mean, if you take the golfer Tiger Woods, you think, oh, he is African American. Well, he is 50 percent Thai, he is 50 percent African American and 50 percent native American. So, depending on which way you want to look at him--

[Laughter]

DR. FLOCKHART: He is more than 50 percent!

DR. MCLEOD: He is an amazing guy. He is 50 percent Thai, 25 percent African American and 25 percent native American, and each one of those is exceptional. So, if you choose to genotype him, which do you pick? Do you pick the Thai part?

DR. FLOCKHART: We have the point that it is a self-described thing anyway so people may even be wrong about that. Dr. Singpurwalla, you might want to respond to my comments.

DR. SINGPURWALLA: Just very briefly, I am not questioning the scientific truth about the analysis that is done. The scientific conclusion may still be valid. Using 50 percent as an argument to make a decision is what concerns me. The chances are that a more sophisticated analysis of the kind my colleague, Dr. Davidian, does will be able to boost that 50 percent to some higher number. I am not sure about it because I have not done that, but I

think just making a big decision based on 50 percent evidence bothers me. That is all. It may still be true.

DR. FLOCKHART: I think we share the same thing, 50 percent should not be in stone. Dr. Relling and then Dr. Gage.

DR. RELLING: To address the question that is up right now, and Dr. Lesko mentioned this briefly, I think that where possible we should try to provide the label with synonyms because it is possible that there will be some laboratories that will provide results back that say *28 and some laboratories will provide results back that say 7/7. We have already developed guidelines for pharmacogenetic consults at our own institution and that is a big part of what we have done, is make those tables that create the synonyms that will map the laboratory results back to information that can be provided in the form of a consult. So, I would say don't try to decide on any one method but be open to the idea of helping clinicians interpret what *28 means, what poor metabolizer means, etc. through use of tables.

DR. FLOCKHART: No graphs?

DR. RELLING: Or graphs. Graphs would be good. In terms of whether 50 percent is enough of a threshold to estimate variability in response, I guess my question would be what is being proposed to try to establish some kind of

hard number of what variability is before the FDA is allowed to ask people to indicate what to include. In my mind, 50 percent seems high enough so I guess we must be getting close to that threshold.

The other thing is that I think I heard somebody mention that they would be loathe to include genetic information in the label while there could be so many other as yet undiscovered genetic predictors of variability, and that our goal should be to be able to account for the majority of genomic variability. My response would be that that is not our goal. Our goal is to determine what accounts for the greatest percentage of variability in a clinical parameter such as dose requirements, which is present in the paper in Blood that you reviewed or INR. So, in fact, when dose requirement was the dependent variable by using age, 2C9 genotype, VKORC1 genotype and height, they predicted over 50 percent of dosing variability, and I think that should be our goal, not accounting for overall genomic variability.

DR. FLOCKHART: Yes, very good. I am going to try and move this part of the discussion to wrapping up. Dr. Gage and then Dr. McLeod and then I am going to ask FDA whether they have sufficient input on this point or whether they want more. Brian?

DR. GAGE: So, my advice in terms of labeling is to remember that labeling is used primarily by clinicians and that ought to be our primary motivation, and if the information is likely to be clinically relevant we ought to include it. So, the CYP2C9*2 and *3 would meet that definition but the *7 would not because it is not likely to be clinically relevant.

My second point is that we need to be careful. Although this information is clinically relevant when we know it, in the absence of prospective studies or more thorough analysis we don't want to be necessarily advocating genotyping. In particular, we want to be careful that our clinical colleagues don't delay starting an important drug while they wait for genotype information which is maybe not going to contribute largely to their dose estimation.

DR. FLOCKHART: That is an important point with TPNT in mind because that really contaminated the discussion about TPNT in a negative way. So, I think we might need some more thinking about how to have those discussions, have fully informed discussions and communicate that point. Howard, last point?

DR. MCLEOD: It really comes down to some of the points that Dr. Singpurwalla was making. You know, 50 percent certainly is not adequate but it is a whole lot

better than 23 percent and 23 percent is where we were and 50 percent is better and let's keep going. So, I think as long as we are going in the right direction, that is a step forward.

In terms of the specific information, restrict it to the evidence. That is where you should stop, and that comes back to Wolfgang's point that there is evidence for CYP2C9*2 and *3. If there is evidence for something else in the future, then an amendment is ready.

DR. FLOCKHART: So, Dr. Huang and Dr. Lesko, has this provided you with enough information or are there specific parts of this discussion that you would like more guidance on from the committee?

DR. LESKO: No, this has been great. I don't have any other questions. I think we got a good sense of where people are on the two questions.

DR. FLOCKHART: We can directly address your question about the label, should the labeling language be the same? If I could just comment on that myself, I think where possible, obviously, it should be the same. But I can imagine situations, and UGT1A1 is one, where it is going to be relevant to multiple drugs. At the moment we have information, a lot of information about irinotecan. So, I can imagine a humongous device label eventually does include it. I think maybe we need to think about ways in

which that can be done in such a way that people can access the useful information about a specific drug in the label, some way to refer effectively to the drug labels. But I am not sure that all the information that is in every drug label needs to be in a device label. That, to me, risks a significant amount of information overload. Next question?

DR. HUANG: This actually has been briefly mentioned by Dr. Lesko, how should the results of a genotype test be reported when technology allows measurement of genotypes where clinical significance is uncertain or incomplete? Do we rely solely on the evidence? I guess the evidence would be clinical genotype response association data to report uncertain genotypes, or would in vitro data be sufficient in certain cases where alleles are rare and clinical data are difficult to obtain? Dr. Lesko has used *7 or we could say *6 of CYP2C9 where it is null alleles; *2, *3 is not a null allele so what if we actually measure *6? Would we include that information? We do not have clinical data--

DR. FLOCKHART: Well, not to preclude but rather to accelerate the discussion, I would note Dr. McLeod's last point, where it is meaningful. Where there is really a change, where the evidence supports that we should include it. But I would invite particularly Dr. Sadee and

Dr. McLeod to think out loud about that. Would you agree with me?

DR. HUANG: Just to clarify, we are discussing device reporting.

DR. SADEE: Well, certainly in vitro data we have enough evidence that there is a null mutation that is rare. Then the information helps. I mean, that practically everybody is at best heterozygous but with another allele being there, such as *2 and *3, your heterozygosity becomes important because then you talk about compound heterozygosity. So, if you have a very rare mutation and you combine this with *2 or *3 that actually be relatively frequent and you might want to include that because it can be easily projected as to what the in vivo impact would be.

DR. FLOCKHART: I think we get into a mess here though. It is vulnerable because we have a particular, let's say, phenotype and I like Mary's use of the word "synonym." I can imagine multiple possible synonyms to describe a phenotypic outcome and I am not sure that we need to iterate the process we every rare allele. If you think about that for 2D6, every time we add a new variant allele we would have to change the label. That doesn't make any sense to me.

DR. MCLEOD: But the way the question is worded--one way of interpreting the question the way it is worded

is that the assay for the common variant happens to give you information for the less common variant. For example, in the UGT1A1 situation where if you try to measure the 6 repeat and the 7 repeat you get information on the 5 repeat and the 8 repeat. You may not have meant to but that is what the patient. The current situation on the device label is that it ignores that fact or it certainly doesn't emphasize that fact and, therefore, the data is derived by the assay. I don't know what is currently done with the data, whether it is reported or whatever, but the data is available yet the package insert is not allowed to comment because these individuals make up such a small proportion of the clinical studies that there is not enough evidence to really address it. In that situation, which has occurred in real life but may be rare--I don't know, we will see in the future I guess, it is really difficult. I mean, this really lays out that issue because you have data suggesting that the 8 repeats will have a high risk of neutropenia but they are so rare that there is no clinical evidence to say whether these people should be at risk or not. If an oncologist orders the assay they could get that information and they would have to figure out what to do with it. It is really not helping them make a decision.

Now, you could argue that it is rare but, I mean, objectively the medical evidence does not allow you to make

a clear decision, but the clinical evidence suggests that these people should be monitored more carefully. So, do we put that in? I mean, it is a really hard question.

DR. FLOCKHART: Well, I think actually the FDA is not asking for stuff in stone here. They are asking for a context like that I think where some judgment would have to be applied. I mean, again, my overall response to this would be to try to communicate the results of a genotype in a communicable way. So, I think for 2D6 that is a really hard thing. You have to deal with simple phenotypes, simple synonyms, if you like. At our center too we are in the business of communicating this stuff clinically to patients and it has to be done in such a way that it comes in three or four communicable categories. I expect for warfarin we are going to have to do the same thing in communicable categories.

So, when you say how should the results of a genotype test be reported, I think we should not have a list of star alleles and put that in stone. I take Howard's point earlier on, there are a few genes where that is legitimate where there are relatively simple genetic or hepatypic structures and it is possible to do that. But in many, many situations increasingly, as we move towards multiple SNPs and multiple genes, we will have to use synonyms.

DR. HUANG: You have discussed several times using a synonym and maybe putting in a table of a drug label, I believe--

DR. FLOCKHART: Yes.

DR. HUANG: --for example, for 2D6 you could say poor metabolizer--

DR. FLOCKHART: In the dosage section.

DR. HUANG: --but for CYP2C9 this may be difficult because we only have data on *2 and *3 and those are not null alleles so theoretically we wouldn't be able to call that poor metabolizer. The poor metabolizers are the African Americans with *6.

DR. FLOCKHART: Okay, we have to come up with a different synonym. That is okay. You know, decreased metabolizer or something like that. Dr. Caldwell?

DR. CALDWELL: Not being used to thinking about these questions, I might be in the realm of fools rushing in where angels fear to tread, but is this a place for your graph? That is, could not the laboratory report out an A, a B, a C, a D or an E which has incorporated the genotype, has incorporated the genotypic panel, if you will, has incorporated age and whatever else becomes relevant, and reports out something an A, a B, a C or a D which then translates to the physician as a dose range for warfarin? So, if you get an A back from the laboratory you know that

this is a range that you should start your patient off with a dose? Does that open difficult problems?

DR. FLOCKHART: Some.

DR. CALDWELL: Are they insurmountable problems?

DR. FLOCKHART: Wolfgang?

DR. SADEE: I think that would be obscuring information. You either bring the real information out and maybe attached with some summation that is practical. But I think we do face the reality that as genotypic assays become less expensive, whether you measure one SNP or 100 SNPs or 1000 SNPs, it is going to be the same price because it is the quality control behind it and there is not going to be much difference. Therefore, we are systematically going to face the need to include less frequent polymorphisms. Right now we are just talking about very frequent polymorphisms that are functional, which by itself actually is quite unusual.

In any case, we need to look forward and have a policy where we can deal with the less frequent polymorphisms and how they interact with the more frequent ones, and how much of the information would be in there, and I don't see any problem with just mentioning in the label only the frequent functional polymorphisms, what is known about them, and then have an attachment saying that there are also less frequent polymorphisms if people want

to look into this, and the measurements are available because it says there when technology allows measurement of this. Well, technology does allow the measurement right now at no cost. So, too much information? I don't know. Leaving it out? Well, I would like to have it. So, as we get more sophisticated it will be very easy to add it.

DR. FLOCKHART: That gets to where it is--

DR. SADEE: Yes, in an attachment because now you are talking about more sophisticated analysis that the clinicians are currently not necessarily prepared--

DR. FLOCKHART: Yes. Dr. Kearns?

DR. KEARNS: I would posit, with respect to the first paragraph, that in the clinical arena more often than not the information is incomplete--well, no, it is relatively certain but incomplete in terms of interpreting it in the context of the patient decision. I mean, we get a point of light with this information that when used by people who have the ability to interpret it can be very, very useful. When put out in the context of people who don't have sufficient knowledge to interpret it as a multifactorial, it could be useless and almost dangerous.

I think at best what we get with this information clinically is some sort of dimensional direction where things are going to go. We know is it likely to be high; is it likely to be low; what are the implications with

respect to the drug. So, some direction, again, is going to be good with respect to prescribing but it "ain't" going to answer the question about clinical utility of the information.

DR. FLOCKHART: Dr. McLeod?

DR. MCLEOD: I need to come back to the last paragraph there talking about the in vitro evidence. I forgot to mention that point when I was talking. I think that in cases where there is clear in vitro evidence in the absence of clinical information because of the rarity of the alleles in this situation, that information should be included somewhere in the package insert, not in the indications section obviously but somewhere because at least it gives some guidance by which someone could at least look further and try to get the relevant information. Again, in the UGT1A1 situation that information would be quite useful because it at least gives someone a flag that there is a particular problem that they need to chase up, and there will be other examples that will be coming forward. For the null allele, as you mentioned, people should at least have the ability to know that information.

DR. HUANG: Can I clarify? Are you talking about discussing this data, incomplete data, in the drug label?

DR. MCLEOD: Yes.

DR. HUANG: And also report out in the device test?

DR. MCLEOD: I am talking about the device test. This question is only for the device label.

DR. FLOCKHART: Well, it is not really incomplete data. If you have clear in vitro data that something is a null, it is not incomplete.

DR. MCLEOD: It is not as good as in vivo data.

DR. FLOCKHART: Right, I see what you mean. Last point, Wolfgang. Then we have 20 minutes for the large, difficult questions.

DR. SADEE: Including this information also makes the point that actually there is more genetic variability or that the two alleles you might want to assay are not accounting for all. One of the major problems in medical genetics, if you look at the literature, is that once there is polymorphism there are 500 clinical assays that eventually make you believe that there is something to it and often it is wrong.

DR. FLOCKHART: Thanks to everyone. Next question I think.

DR. HUANG: We have three questions related to warfarin. I would like to ask the committee does the committee agree that sufficient mechanistic and clinical evidence exists to support a recommendation to use lower

starting doses of warfarin for patients with genetic variations in CYP2C9 that lead to reduced activities?

DR. FLOCKHART: I am going to start the discussion by being provocative. Actually, I think that the evidence to support actually using lower starting doses doesn't exist We don't have enough. We have very, very little information about starting doses. We have huge amounts of information about the maintenance dose. What I would like to see in the long run is a setting where I think it is practical to start with a 5 mg dose in the vast, vast, vast majority of people and, at the time that dose is done, to have a blood draw and then, within a matter of 24-48 hours, because not all settings are the same as Dr. Caldwell's where we can get a 2C9 genotype back in four or five hours--we don't know how pharmacogenetic testing is going to be done from a business point of view nationwide or, for that matter, internationally so my thought would be to try to move to a setting where--and this is much, much more practical than expecting places to set up a very fast turnaround for 2C9 assays--so, is there enough mechanistic and clinical evidence to support such a recommendation? I would argue that we have very, very few trials that address the question for the first dose, the induction phase.

DR. HUANG: Dr. Lesko, in his presentation, modified some of our questions. Would you consider to start with the dosing but then genotyping--

DR. FLOCKHART: Absolutely.

DR. HUANG: --and change the escalating dose--

DR. FLOCKHART: Right. I think that ought to be what we do.

DR. HUANG: --genotype dependent?

DR. FLOCKHART: Right.

DR. HUANG: With that you would agree?

DR. FLOCKHART: Yes, I would.

DR. SINGPURWALLA: Point of information, what do you mean by mechanistic evidence? I understand what you mean by clinical evidence. By mechanistic, do you mean running data--

DR. HUANG: Warfarin is metabolized by 2C9--

DR. FLOCKHART: She means in vitro laboratory data.

DR. SINGPURWALLA: I didn't get you.

DR. FLOCKHART: I didn't mean to speak over Shiew-Mei or confuse the transcription process but I think she means laboratory in vitro mechanistic data explaining why 2C9 genotype influences warfarin kinetics. Correct?

Now, this is like we have a motion on the table because we are going to need to vote on this. Do you want

to vote for them all collectively? Which would be the most helpful to you? Probably individually?

DR. LESKO: I think it would be individually on each question. But I am not clear on the answer that has been provided to this question.

DR. FLOCKHART: Well, I think because the question has been changed, amended, if you like. Without using specific wording, if I could propose that what we are really voting on here is does the committee agree that sufficient mechanistic and clinical evidence exists to support the recommendation to use initial genotyping, followed by a maintenance dose determined by genotype. Is that where you are going, Larry?

DR. LESKO: Yes, I think the evidence that I presented and one of the ways of possibly integrating genotyping into the clinical situation, particularly when there was concern about delay in starting therapy in a patient that needs it, was to begin with a fixed dose but very quickly obtain the genotype information--

DR. FLOCKHART: Right.

DR. LESKO: --and as Dr. Caldwell pointed out, depending on the setting this could occur in four hours or it could occur in five hours, but the important point is to incorporate that information into what is done next in conjunction with the INR. I think to ignore that kind of

information and simply start out with 5 mg--I think the evidence that we have seen to date indicates that the maintenance doses are not predictive in the general population from that starting dose.

So, the question I think is what are we saying by starting dose. I think what we are saying is the dose that we initiate in the induction phase therapy. It doesn't mean the exact first dose in the warfarin regimen but it would be the first doses in the early part of therapy. I just want to make that distinction.

DR. FLOCKHART: So, that has clarified what you mean by starting dose here. You don't mean the very first dose. With that clarification, are there any other, if you like, points of information from the committee in order for people to understand the motion on the floor?

[No response]

In that case, let's go ahead and vote. I am sorry, there is a question.

DR. RELLING: So, you have suggested a wording change there because I think it is important to get this wording right.

DR. LESKO: So, I think the warfarin question is related to I guess part A of that question.

DR. FLOCKHART: Yes. I think we can focus very simply, Larry, on the term "starting dose."

DR. LESKO: Right. If we take out "starting," to use lower doses of warfarin for patients with 2C9 and leave it at that. I think the confusion is around the "starting" part where people are thinking of it as first dose as opposed to the initial induction phase of therapy or dosing in the initial phase of warfarin dosing.

DR. FLOCKHART: Do we say something in there about using 5 mg first, and so forth, or not and just avoid that subject?

DR. LESKO: I think we can avoid that.

DR. FLOCKHART: This is like having an amendment on the floor. So, are there any questions about just using the term "lower doses?" Any questions about what people mean by that? Everybody understands the amendment? Any arguments for or against the amendment?

Having heard no arguments, then we would include the amendment and delete the term "starting dose" and the original motion, if you like, is back on the floor. Any speakers against the whole motion?

[No response]

Any objection to adopting that language by consensus?

[No response]

We got through without a vote. Good. Next question?

DR. HUANG: So this will also be amended, does the committee agree that sufficient mechanistic and clinical evidence exists to support a recommendation to use lower doses of warfarin for patients with genetic variations in VKORC1 that lead to reduced activities?

DR. FLOCKHART: Questions about that language?

DR. SADEE: It is a bit ambiguous. What you would like to say is that lead to reduced enzyme activities.

DR. FLOCKHART: So the term "activity" is ambiguous. VKOR activity is what you mean. If we amended it that way it would be clearer. Any objections to that amendment, relatively minor?

[No response]

Any other questions so everybody fully understands the language here? Would anybody like to speak against including this?

[No response]

Objections to consensus to include it? No objections to consensus. So, hearing none, we will move on to the next one.

DR. HUANG: The second question is does the committee believe that genotyping some or all patients prior to beginning warfarin therapy would reduce adverse events and improve achievement of stable INR?

DR. FLOCKHART: So that people understand the question before us, any question about that?

DR. SADEE: We just talked about giving the first doses and then obtaining a genotype to maybe guide further dosage development. So, the prior is the first question.

DR. FLOCKHART: So, how about establishing stable warfarin therapy? In other words, fundamentally you are saying arriving at a maintenance dose. So, if we modify that as genotyping some or all patients prior to establishing stable warfarin therapy would reduce adverse events, with that deal with that?

DR. LESKO: David, I was just thinking of the conversation we have had and the terms here are important but I think the balance here is between "prior to beginning" and in the early part of therapy as we just discussed with the prior question. The intent here was that genotyping information would be available as early as possible in the sequence that would initiate warfarin therapy.

DR. FLOCKHART: So, what language would you recommendation?

DR. LESKO: I would take out "prior." I would amend this question and take out "prior to beginning warfarin therapy" and include does the committee believe that genotyping some or all patients at the initiation of

warfarin therapy, which may be before any dose is given or shortly thereafter--we might borrow language from that first question that the committee voted to amend and incorporate the same sort of language.

DR. FLOCKHART: I am just worried about actually preventing people doing this by making it sound like they have to get a genotype, have to get a fast turnaround.

DR. LESKO: Yes, that is not the intent of the question, that it is a mandatory test or that it is required before the drug is begun. But I think, on the other hand, getting the test early is essential to benefiting the most from that inference.

DR. FLOCKHART: So, how about some or all patients in the early--no--

DR. LESKO: I am going back to the evidence we presented today. I think we are referring to the induction phase, having a genotype in the induction phase of warfarin therapy would be the appropriate language perhaps.

DR. FLOCKHART: Induction phase is good. Dr. Kearns?

DR. KEARNS: I am a bit confused about what is meant by "some or all." I mean, if you are going to do some how do you pick them out? Hair color? Eye color? You know, I get "all" but I think it has to be one or the other.

DR. FLOCKHART: I think we are dealing with one amendment at a time and I think we have dealt with the "prior to" so we are moving on to another amendment here. You are essentially asking that if we say "some" we have to define "some." We could just take it out--genotyping patients prior to warfarin therapy and delete "some or all." That would be the amendment that you proposed?

DR. KEARNS: Yes.

DR. FLOCKHART: Any disagreement with doing that, including the FDA? Dr. Gloff?

DR. GLOFF: My comment was just on the wording of "prior to." I was going to make an alternative suggestion but it sounds like you think that is fine.

DR. FLOCKHART: An alternative suggestion is another amendment and you can do that.

DR. GLOFF: It is not important it is just that there was so much discussion going on and I had what I thought was some simple wording that would resolve it, but I think it has been resolved.

DR. FLOCKHART: It has been resolved.

DR. GLOFF: Okay.

DR. FLOCKHART: So, the wording that we would finish with I think, if you could help me, Larry, is does the committee believe that genotyping some or all--sorry, genotyping patients in the induction phase of warfarin

therapy would reduce adverse events and improve achievement of stable INR? That is the question.

DR. LESKO: I think that is correct.

DR. FLOCKHART: Would anyone like to speak against that? Emphasize points for it? Objection to consensus? Any objection to consensus of adopting that language?

[No response]

Good. Seeing no objection, we will move to the next one.

DR. HUANG: This will be about VKORC1.

DR. FLOCKHART: I think we have had substantive discussion, unless anyone would like to make separate points about VKOR? In particular, since Dr. Gage is here, any extra points about that?

DR. GAGE: I want to make sure the wording is clear. This would be rephrased as does the committee believe that genotyping patients in the induction phase would reduce adverse events and improve achievement of stable INR? I think that is a clear question.

DR. FLOCKHART: Everybody understands? Any objection to consensus for adopting that language?

[No response]

We will move right along. We have five minutes. The third question?

DR. HUANG: The third question, 3a is CYP2C9 and 3b is VKORC1. Does the committee believe that existing evidence of the influence of CYP2C9 genotypes warrants relabeling of warfarin to include genomic and test information?

DR. FLOCKHART: First of all, questions about what is being proposed? Does everybody understand that? Does anyone want to speak against it or modify it? Emphasize points for it? Adopted by consensus, any objections to doing that?

DR. SINGPURWALLA: Yes.

DR. FLOCKHART: Would you like to speak to that point?

DR. SINGPURWALLA: No, I have spoken already.

DR. FLOCKHART: Well, given that there is not going to be complete consensus, we need to take a vote. So, if I could ask first just to see the hands voting for adopting this into the labeling language? Visually, a majority of the committee would be for that. You can take your hands down one second. So, we need to individually go around and vote yea or nay. Dr. D'Argenio?

DR. D'ARGENIO: Yes.

DR. FLOCKHART: Dr. Capparelli?

DR. CAPPARELLI: Yes.

DR. FLOCKHART: Dr. Sadee?

DR. SADEE: Yes.

DR. FLOCKHART: Dr. Singpurwalla?

DR. SINGPURWALLA: No.

DR. FLOCKHART: Dr. Kearns?

DR. KEARNS: No.

DR. FLOCKHART: For myself, yes. Dr. Barrett?

DR. BARRETT: Yes.

DR. FLOCKHART: Dr. Relling?

DR. RELLING: Yes.

DR. FLOCKHART: I am sorry, I can't see around there.

DR. GLOFF: Carol Gloff, yes.

DR. FLOCKHART: Dr. Davidian?

DR. DAVIDIAN: Yes.

DR. FLOCKHART: Thank you, folks. Let's go to the next part.

DR. HUANG: If yes, what information should be provided in the label?

DR. FLOCKHART: Well, I think we have had a pretty extensive discussion about what information should be provided in the label and where, including discussions of graphs, predictive things, iterative stuff and so forth. So, you have all kind of feedback on that. "If no" is no longer relevant because we voted yes.

DR. HUANG: So, we will go to 3b, the same question, does the committee believe that existing evidence of the influence of VKORC1 genotypes warrants relabeling of warfarin to include VKORC1 genomic information and testing information?

DR. FLOCKHART: Let me try and accelerate the process by asking Dr. Singpurwalla whether he is going to disagree with this too.

DR. SINGPURWALLA: As a corollary to the first one, yes.

DR. FLOCKHART: Of course, you are consistent. Would anyone like to speak to it one way or the other, for or against?

[No response]

We can't have a consensus so we are going to go around again. Dr. D'Argenio?

DR. D'ARGENIO: Yes.

DR. FLOCKHART: Dr. Capparelli?

DR. CAPPARELLI: Yes.

DR. FLOCKHART: Dr. Sadee?

DR. SADEE: Yes.

DR. FLOCKHART: Dr. Singpurwalla?

DR. SINGPURWALLA: No.

DR. FLOCKHART: Dr. Kearns?

DR. KEARNS: No.

DR. FLOCKHART: Myself, yes.

DR. BARRETT: Yes.

DR. RELLING: Yes.

DR. FLOCKHART: Can you state your names because the record needs it?

RELLING: Yes.

DR. GLOFF: Carol Gloff, yes.

DR. DAVIDIAN: Davidian, yes.

DR. FLOCKHART: Thank you. That concludes our proceedings I believe for the morning. We can move to lunch and return in the afternoon, when Dr. Barrett will be the chair. Thanks to everyone.

[Luncheon recess.]

- - -

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

Topic 2: A Critical Path Pilot Project in Pharmacometrics (Qualitative Methods)

Call to Order

DR. BARRETT: My name is Jeff Barrett. I am going to call to order the afternoon session. We are going to again start with a roll call, so if we could go around the room and just state your name and affiliation for the record.

DR. D'ARGENIO: David D'Argenio, University of Southern California.

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

DR. SADEE: Ohio State, Wolfgang Sadee.

DR. SINGPURWALLA: Nozer Singpurwalla, George Washington University.

DR. KEARNS: Greg Kearns, Children's Mercy Hospital and Clinic.

DR. JUSKO: William Jusko, New York University at Buffalo.

DR. GAGE: Brian Gage, Washington University in St. Louis.

DR. MCLEOD: Howard McLeod, Washington University, St. Louis.

DR. BARRETT: Jeff Barrett, Children's Hospital, Philadelphia and University of Pennsylvania.

DR. PHAN: Mimi Phan, executive secretary.

DR. FLOCKHART: Dave Flockhart, Indiana University.

DR. RELLING: May Relling, St. Jude Children's Research Hospital.

DR. GLOFF: Carol Gloff, Washington University and independent consultant.

DR. DAVIDIAN: Marie Davidian, North Carolina State University.

DR. POWELL: Bob Powell, FDA.

DR. LESKO: Larry Lesko, FDA Clinical Pharmacology.

DR. WANG: Yaning Wang, Clinical Pharmacology, FDA.

DR. BARRETT: This afternoon's session is going to move a little bit further down the path of the critical path project, specifically on the topic of pharmacometrics. We are going to have an update on the end of Phase II progress by Dr. Robert Powell, and then we are going to be treated to a case study by Dr. Yaning Wang, focusing on quantitative approaches to assess a genomic design and a biomarker titration design for a Phase III clinical trial.

So, there is some continuity with the morning session but this afternoon we will be focusing a lot on the quantitative aspects and using, I think, the approach of model based drug development and quantitative tools in general as a vehicle to communicate the progress of drug development as these agents move down the pipeline. I thought the morning's discussion was very good in terms of framing some of these discussions, particularly as it pertains to pharmacogenomic testing so I expect there to be a lot of overlap. Without further ado, Dr. Powell?

FDA Experience with End of Phase IIa Meetings:

An Attempt to Improve Drug Development Decisions

DR. POWELL: Thanks. I know this was presented at the advisory committee, end of Phase IIa meetings, before as a concept. Larry's recollection was that it was in 2003 when this was presented as an idea.

What I am going to do is to go through our experience to date. End of Phase IIa meetings is a pilot project and we will be summarizing where we are after the current series of meetings that we are going through now, and then making recommendations within the FDA in terms of what to do next. So, getting your recommendations and observations on what Yaning and I will present will be useful to us. And, I would like to thank you, Jeff, for granting me a Ph.D.

DR. BARRETT: It is an honorary degree, Bob!

[Laughter]

DR. POWELL: So, the whole idea I think is--I tried to capture what is the need for this type of meeting. At the bottom are phases of development with flexibility in learning on my left and research and development expense and revenue on the right. My sense, at least from my years in R&D, is that flexibility and learning tends to decrease as you move out closer to market, and the expense of R&D tends to increase over time. You could draw this slightly differently I suppose, but once a drug goes to market it certainly tails off. Revenue, obviously, takes off once a drug goes to market--no-brainer.

But the piece of what we are trying to do is near the end of Phase I through about Phase IIa to address a sweet spot when a company's development program might still be flexible enough that they could take the information that they have from preclinical and Phase I through Phase IIa and when they have some information on proof of concept and understand some of the pharmacokinetic and dynamic characteristics, to then sit down with the FDA and say, okay, based on your experience what shall we do next. Largely, what should be done next is usually in the form of trial design. I mean, that is the explicit translation of what is next.

[Slide]

So, 50 percent of people have indicated, primarily out of Boston, that the data would say that there is about 50 percent clinical trial failure rate. Is it true? Well, I think most of us with experience would say certainly it can be fairly high. It is going to be different from one therapeutic area to another. It is probably lower with antibiotics, higher with CNS drugs. What people would like to be able to do is to increase, if you can, finding true positives. If you can predict true negatives, then you might not want to engage in those trials unnecessarily if you could avoid it. Of course, you would want to decrease your chance of running into false positives and false negatives.

[Slide]

So, the root causes for finding these effects, if you can plan for it, that you may not wish to encounter would be a lack of efficacy and higher toxicity than might be appreciated; a placebo effect; baseline effects; dropout rates and patient selection. So, if you can account for these factors approach priori then you might be able to decrease late-stage trial failure rates which are fairly expensive.

So, the idea is simply that if you can account for known failure sources from prior information, including

the information in the development plan, then you might be able to include that in clinical trial design.

[Slide]

That is science. Another reason that probably bad decisions are made has to do with these factors that come from what are called smart choices: Conspiracy of optimism is where you have a champion, project champion. In drug companies a lot of times these are Phase II/III physicians and they are going to drive that thing to an NDA. Anyway, there is a bias in that direction.

Framing the problem too narrowly; not involving the right people; avoiding uncertainty--we have all seen these whether you are in academics or the FDA or drug companies. So, it is just another form of bias.

[Slide]

So, then that leads into what is model-based drug development. This thinking is seminal to what an end of Phase IIa meeting is where the objective is to improve decision quality by employing drug disease models in clinical trial simulation. Everyone knows kind of what a model is. Drug disease models can be empiric or mechanistic. I will show you an example in HIV of a mechanistic model in a second.

The disease portion of the model would account for what you know about the relationship between patient--a

lot of things we talked about this morning--biomarker and the relationship to morbidity and mortality. The drug portion simply adds on to that, what you know about the drug in terms of the dose, the combination, placebo effect and patient characteristics, size, age, adherence, and adverse events to the disease model.

Well, what do you do with it? Once you have this sort of model you have to do something with it. Inside a company you can use this for go/no go decisions and there are companies that are doing that today, and you can use it for dose and regimen selection. It is used to some extent inside the FDA for approval decisions, another form of go/no go decision, and also for labeling.

[Slide]

The end of Phase IIa meetings--Larry and Peter originally conceived it, the purpose is to decrease unnecessary late-phase trial failures. Probably the most important adjective is that it is non-binding. A company can come in and basically discuss their development program and we will discuss and we will, in effect, do a quantitative modeling or simulation with them and we give them recommendations but they are not down to whatever comes out of that. It is really a learning type of meeting.

What we have done--and I will show you this and Yaning will show you some of this in a second--is we take the relevant company's prior data from Phase I through IIa and we also take data from the literature on that disease or drugs in that class, and we have also extracted data from inside the secret FDA database, which doesn't exist, but it is data that we can extract from prior submissions, not to give us information on another company's drug but to tell us about placebo effect and dropout rates and that sort of information that may be critical for this specific application. So, it is a way we are working to share that information which companies speak about that they would like to be able to do.

So, we make recommendations on the sponsor's trial design. We will also look at alternative designs, patient selection and dosing regimens. Once we are finished we share all of our code with the sponsor and we encourage them to extend the work, the models, when they come back for end of Phase II meetings or for the NDA. We explicitly answer their questions, along with our clinical colleagues and biostatistics colleagues. The time course is about six weeks from the time we receive the data from the sponsor, and the key participants I have listed here but I will expand more in a second.

[Slide]

So, the process is basically that there is a meeting request that comes into the therapeutic division. There is a letter with some questions. The FDA evaluates that and will grant or not grant--we have turned down some applications, more so when we felt that mainly of a marketing orientation. We have a phone meeting with the sponsor to explain the purpose of the meeting and the data that we will need. Basically, we need their data up to that point. We need their next trial design, what they are planning to do, be it IIb or III, and we need a summary on the drug to that point. That is summarized here.

Once we get the data, as I mentioned, it is about six weeks from that point until we have the face-to-face meeting. We begin analysis. We invariably run into questions from the analysis. We ask questions sometimes weekly of the sponsor. Sometimes the sponsor will ask other questions, other than what the trial design would be, so we try to answer those before the meeting in writing. If we think the sponsor has the resources, we tell them the form of the simulation that we are interested in with trial design alternatives, dosing regimens, sample size, and so on, and then we have the meeting. It is about an hour and a half meeting where we focus on drug disease modeling and trial simulation for the trial. It is about a 30-40 minute presentation between the FDA and the sponsor. Then we have

about an hour of dialogue based on the development plan and what the science says.

[Slide]

Roles and responsibilities--the project manager holds everything together, coordinates sponsor communication, FDA meetings and documentation. The physician is primarily responsible for primary endpoints, disease information, pointing us to where to go in the FDA to extract information, trial design and draft guidance. Statisticians are key for the trial design, and also knowing--because they are usually oriented to the therapeutic area--where to go for the best dropout rate and placebo information, and they participate in the simulations. The clinical pharmacometrics people tend to focus on drug disease modeling, dosing regimen, drug interactions and simulations.

[Slide]

Now, this is the case study that I will speak to. It is in HIV. These are the questions that the sponsor was asking. It was a new drug with a new mechanism of action. Nothing is on the market with this mechanism of action. They were asking for a given area under the curve based on the relationship between viral load suppression from a ten-day treatment Phase IIa trial, is it reasonable to select the best dose for Phase III? They asked is testing BID

giving the drug once a day appropriate? And, could they make a decision from their IIb trial at four weeks to select the dosing regimen for their Phase III trial? So, those were the questions.

[Slide]

This is not their data but it sort of looked like this, where looking at area under the curve with a number of different doses that they had looked at, first is suppression of viral load and they wanted to achieve about a 1.5 log drop in viral load. So, they selected this area under the curve and figured that if they figured out how to dose the drug to that endpoint that that would be okay.

What we did was--actually Jenni Zheng, Yaning and Joga Gobburu worked on this--that strategy would not allow you to understand anything about the time course and the schedule. So, they set up a mechanistic model that I will describe in a second to look at time course of drug effect, the schedule, drug interactions--obviously, these drugs are always given with two other drugs at this phase of development--and to incorporate adherence, drug interactions, dropout rate and resistance into the model to help with the choice of trial design.

[Slide]

So, this is what they were proposing. They were going to give their drug, drug X in a parallel study, all

arms with lopinavir and ritanovir. At 1 mg BID, 2 mg BID, 4 mg once a day and then the standard of care was Combivir with lopinavir ritanovir.

[Slide]

I should have also mentioned that this was done in collaboration inside clinical pharmacology in addition to physicians, with Kelly Reynolds and people from the anti-infective team.

So, this is the model that was put together. Their mechanism of action is not listed here, but this is CD4 counts being affected by the virus, going into a latent phase or going into an actively infected phase where more virus is created and the infection goes on with (N)NRTIs and protease inhibitors acting at these points. Abbott and I think Pfizer use this model in their development programs. Probably some other companies do as well.

[Slide]

These are from that ten-day trial that I talked about. These are the mean data and I will show you individual fits in a second. The fit curves are for 2 mg once a day, 4 mg once a day, 2 mg BID and 6 mg BID. You see from this clearly dose response in suppressing viral load going out for the ten days of treatment and then stopping drug. But the other thing that you see is a schedule dependency. So, you see giving exactly the same

daily dose, 4 mg once a day versus 2 mg BID, you get more of an effect with giving the drug twice a day. As I mentioned, this was applying the previous model to that.

[Slide]

These are individual patients. This is single drug therapy. This is just giving the drug X alone. So, this is one patient with their viral load versus time, and this is the computer fit that Yaning and Jennie did. So, they fit pretty well for individuals.

[Slide]

For a lot of protease inhibitors, for example, drugs in this class, there is tradeoff. So, as you increase the dose you can drop viral load, but the other problem is that as you increase the dose you increase gastrointestinal adverse effects. I mean, if someone is nauseated and throwing up they can stop taking the drug and decrease adherence. So, understanding these therapeutic tradeoffs was important. I will show you a bit on the dropout rate and some of these other effects in a second.

[Slide]

Yaning simulated the effect of lopinavir ritonavir, trade name Kaletra, BID which was given to everyone, looking at viral load over time as they were going to be using it in this trial, looking at the drop in viral road in combination with 1 mg and 2 mg BID and 4 mg

once a day. You can see here the drug effect using this model, but you don't see that much dose dependency.

Knowing that this drug did produce significant gastrointestinal side effects, then what Yaning went on to simulate is if they used a lower dose than they were planning to, the model indicated that it would probably be about as efficacious and, from their clinical data, would have less GI adverse events. So, we ended up recommending that.

[Slide]

This is data that Frazier Smith, the statistician that we were working with in antivirals, extracted from another NDA. This is looking at the dropout rate over a one-year study--so, don't pay any attention to this out here, You see this biphasic dropout rate occurring on a drug that had some characteristics that were similar. So, this is something that we then applied to our clinical trial simulations.

[Slide]

There were a number of scenarios that were put together. Here Yaning looked at 20 simulated trials. The conditions were a 2-log drop, 90 percent adherence with no dropouts. There were other ones that looked at different dropout series. This suggested that 2 mg BID was the most likely winner. This is using trial simulation software.

This is the chance of being a winner at different weeks going out for the one-year study. As you can see, as you moved out it began to appear that the 2 mg BID was winning over the other treatments. But early on it was hard to make that discrimination.

[Slide]

So, our recommendations to the sponsor in this case were, is the target area under the curve of 950 reasonable? Well, okay, but it is probably more informative to use the model and individual concentrations and then from that you could extract BID versus QD.

Secondly, if you want to discriminate between BID and QD, what we saw was that the BID regimens were preferable. In this case, like in many indications, there is market pressure to give the drug once a day but, clearly, you can see that you get better performance by giving it twice a day.

In addition, we recommended lowering the dose for at least one of the treatment arms. Can you make a decision at four weeks to select the dose for the Phase III trial? Both based on what the clinical teams--the way they operate, as well as from this model, there was convergence that you can't make that decision at four weeks reliably.

In addition, we indicated that the Kaletra effect was so strong that it may be difficult to demonstrate dose

response, and we thought that the Phase III trial was adequately designed to determine dose response.

[Slide]

Over the past year we have conducted five of these end of Phase IIa meetings and three are currently in progress. They have been in these therapeutic areas, HIV that I just went through with a new mechanism and there was a question of dosing; prostate cancer where there was a formulation question and dosing; type 2 diabetes where there was a genotype question with dosing; anticonvulsant, new mechanism of action; hot flashes; pain; weight loss.

The workload is about 5-7 person-months per project. Then we do a post-meeting evaluation which is kind of odd, I suppose. We actually do a post-meeting evaluation with the FDA team as well as the sponsor, and it is a number of questions, one of which is just a score where 5 would be pivotal and 1 would be worthless--what is the value of this meeting to you in your development plan? They have ranged between 4.1 to 4.3 from the sponsor. Inside the FDA a number of people are saying, well, why are we doing the company's business to help them out here?

[Slide]

The sponsor's comments--now, the case that I presented, this is exactly what they presented at their senior research and development meeting where they

recommended inside that they wanted to take the higher priority projects and begin to get them set for this sort of meeting. So, we said, you know, people at the FDA seem to be serious about this. I am not paraphrasing this, this is what they were saying, that it aids in selection of doses; the FDA is inviting sponsors to participate for certain drugs. We have not been discriminative today but if the demand outstrips our supply, then we would probably start wanting to go for high impact sorts of diseases.

FDA's preparation was extensive, and our preparation must also be extensive in the company. We need to get the data. We don't need mean data; we need the data. So, sometimes that creates a bit of a delay.

[Slide]

Which has gotten to another issue. Inside the FDA we tend to get data in a fairly heterogeneous way. So, what we have begun to do--Peter Lee is working on this system for clinical pharmacology, to begin getting our data sorted out fairly quickly and then channeling and knowing what to do with it. For data input currently there is an electronic document room. But data comes in in a heterogeneous way and I will describe a bit about CDISC in a minute, but that is what has to happen in the near future so, instead of Pfizer and Glaxo and whatever having different nomenclature for the same term, that has to

become the same nomenclature so that we can then set up a database in a more meaningful way.

Peter has begun to work with data visualization once the data comes in, then lodging the data that we need to analyze in aa data warehouse, then setting it up for data analysis. If we are reviewing an NDA, then we are beginning to create disease modules, screens to help reviewers review the data to ask, traditional in Larry's and clin. pharm's. vernacular question-based reviews, does the drug work; what is the risk-benefit; dose response and so on. And, then to lodge that back into the data warehouse.

A second path is for end of Phase IIa meetings to use disease models to assist in clinical trial simulation, and sending that back into the warehouse and then assisting in end of Phase IIa meetings.

[Slide]

The CDISC, it seems to me, is sort of like you have all these languages and you are at the World Bank or United Nations or French business school, INSEA or Novartis--they can't speak all these languages so it is going to be one language that they speak so that they can communicate inside their organization. Well, you know, whether it is safety or efficacy this, to me, is a tremendous problem that really needs to be addressed fairly

quickly because it really limits our ability to put data together in a timely fashion with it is for safety issues or issues like I am describing here. That is what CDISC gives. This is something that I was actually not aware of prior to coming to the FDA but I can certainly see the importance now.

As I mentioned, we are developing drug disease models, in part from the NDAs that we review but also from these end of Phase IIa meetings. We extract information from the literature, prior NDAs. There are some other diseases I haven't mentioned, like kidney transplant rejection, systemic lupus erythematosus, and in development we are working with osteoporosis and non-small cell lung cancer. We are thinking about how to share these models in a public sort of way so that people can then have dialogue around these disease models.

[Slide]

This is one that Hae-Young Aen and some other people have put together in diabetes. Simply, it explains the relationship between fasting plasma glucose and hemoglobin A1c time course and magnitude of change at this point from three different drug mechanisms of action. Through this, we are talking about adding on some adverse drug reaction non-mechanistic models with this.

[Slide]

This is some of the information. This model was applied to a prior NDA that is sitting in the bowels of the FDA. This is looking at the fit between fasting plasma glucose and hemoglobin A1c for 1000 patients.

[Slide]

So, we started talking about, well, how are we going to save these diseases and communicate. Well, partly this is creating a quantitative disease library with the objective of improving trial outcome. There are levels of information, the first of which is clinical trial data, and the FDA can't share a sponsor's data. But what we think we can do--the sponsor could do that or the NIH can share their actual data. What we think we can do that would have some value is to take derived quantitative information from trials inside the FDA and share that data without revealing anything about the sponsors, as well as drug disease models.

The derived quantitative information would be more than this, but basically what is the central tendency for variability and the time course. But to do that we have to create some standards. Clearly, some of this would be modeled information around the disease endpoints, the biomarkers, who are these patients, what was the trial design, and what was the disease or disease phase, placebo, dropout rate, adverse drug effects and covariates.

We have chosen two diseases to create as a prototype, and what we are planning to do is to expand this dialogue inside the FDA and, if people are comfortable with it, then begin to have conversations outside the FDA.

[Slide]

In summary, end of Phase IIa meetings--we have completed a number and we are coming to the end of the pilot program. We will be summarizing that inside the agency with recommendations, some of which we hope to get from you. A guidance has to get out on this. With model based drug development, which is kind of part of the end of Phase IIa meeting, we are looking at ways of exploring disease models and sharing that more openly, and creating a software system to do the sorts of things that I talked about in this talk. And, we welcome your comments and recommendations.

Committee Questions to Speaker

DR. BARRETT: We have a slot here to ask some questions so why don't we do that now and then we can always come back and ask more. Dave?

DR. D'ARGENIO: Bob, in your experience so far, what is the biggest time sync for doing this? Building the PK/PD models or does that come to you from the company? Is there any one part of this process that is the most time consuming?

DR. POWELL: Well, let me go to one of the guys that has done this. What would you say in doing the quantitative work is the biggest time sync for you? This is Yaning Wang.

DR. WANG: When I did that HIV, my part was mainly the trial simulation part. Of course, others also work on the dropout and adherence model. For my part, the most time consuming part is--well, of course, this is my first case in HIV therapeutic area. I also had to spend quite some time in getting to know the basics of those viral dynamics. That took me probably one week. Then, the next most time consuming part is to set up the model in the simulation by including all the information, including dropout, adherence and all the variability because, as you can see, we used TS2 from Pharsight to simulate the clinical trials. Therefore, we have to translate our model results from Nonmem, which is different model software, to the TS2, at the same time including adherence, dropout and everything. But I would say it is modeling, construction, Nonmem and trial simulation set up. Also, when we do run the simulation we only run, as you saw, 20 replicates. That also would take a lot of time. It requires computation power if we want to see 100 or 200 replicates. That is my opinion about where the time is mainly located.

DR. POWELL: One of the things, as Yaning said, this was actually the first one that we both worked on and it is the first time that we have done HIV. If we do more there should be some efficiency.

I forgot to mention this bullet, send an academic friend to the FDA to create a disease model.

[Laughter]

Larry has some money and Pascal, from France, is working in cancer. He sent one of his graduate students in to work with some people in our group for the non-small cell lung cancer model. We linked up with a physician in oncology and we now have something that we can work with. It still needs to be extended but what we have to move towards is having these models in place so we can use them over and over again.

DR. JUSKO: I have two comments. When you ask for feedback from the companies, I wonder if this was done in a blinded manner--

[Laughter]

--When we ask students for evaluations, when we know who they are, they always give us high grades so we do it anonymously.

DR. SINGPURWALLA: How do you manage that?

DR. JUSKO: Computerized evaluations where students don't identify themselves. But my more serious

question is this quantitative information database sounds like a superb idea and I wonder if there are possibilities of your merging information from various sources and sharing it with the academic communities as well. But as part of my question there also, have you had any objections from companies in doing this?

DR. POWELL: Well, we haven't done it yet to have objections. Actually, I introduced that slide, the idea, a couple of weeks ago at a PhRMA meeting. You know, my sense is that R&D people will generally say, yes, if we can get better information to plan trials we are up for it. But it hasn't been tested in a broader way yet. I think, for example, when clinical pharmacology does the reviews of a given NDA, then that information becomes publicly available. So, if we extract this information and we put it in the clinical pharmacology review we ought then to be able to take it into a more structured public database. That is why I put the slide in of what they said inside their meeting inside the company. You know, to me, sure, you run that risk. So, that is what they said inside the company and someone told me about it later and gave me the slide. So, that is why I showed it.

But the other thing is you can see from the comments that there is heterogeneity from the company statisticians, project managers, clinical pharmacology

people, and some of them don't think it is such a hot idea and some of them think that it is just really cool. So, there is that bias of course.

DR. SINGPURWALLA: Two questions. I take it this is voluntary on the part--

DR. POWELL: Absolutely.

DR. SINGPURWALLA: Because it is voluntary and because you are using a model to simulate, suppose that your simulation reveals results that are not attractive to the manufacturer, what choice does the manufacturer have?

DR. POWELL: They have the same choice that they had before the meeting. I mean, we have seen some where, if I was on the inside, I would probably favor more of a decision not to proceed but that is not the FDA's place to say, well, you should stop development. So, we don't get into that. We just give them quantitative information and talk about how to proceed.

DR. SINGPURWALLA: But giving them information that is questionable to their interests--

DR. POWELL: Right?

DR. SINGPURWALLA: --if I were the manufacturer, I would come and start questioning the models that you have used because giving them questionable information is also a signal that when they come back one or two years later on you are liable not to give approval.

DR. POWELL: Yes, you raise a couple of things. I think you have raised a couple. Let me start at the end. Our intention would be that the people that do the NDA review would not have worked on the end of Phase IIa meeting so that, you know, we don't want to get into an internal conflict or conspiracy of optimism by us continuing to work on the project as it goes to the NDA. You know, that is one concern.

I think it is a model so that we are not only giving them the model but we are giving them other information, like these dropout rates or in some cases looking at baseline from prior submissions. So, they are getting a lot of information that they might not be able to get otherwise. You know, we don't own stock in the company so it is a learning experience.

DR. BARRETT: Dr. Davidian?

DR. DAVIDIAN: Not really following up on that but sort of following up on that, I just have a question. To what extent do you carry out, say, a sensitivity analysis, maybe changing the model a little bit, tweaking some of your assumptions and seeing what effects? Another question, and maybe some of this will come up later too, as far as dropout rates go for example. How do you even simulate dropouts? Do you have different mechanisms by which individuals drop out? Is it purely this percent

dropout and you just randomly drop them out, those kinds of things? So, I am just kind of wondering.

DR. POWELL: Why do people drop out? Is that what you are asking? In the example that I showed you, I mean, we took a drug that had a similar adverse event profile, similar reasons, and we just took that and used that dropout rate in the simulation. It would be much more informative if you had a dropout model in a disease. You could do that but we haven't had the time to do that. Yaning, do you want to talk to sensitivity?

DR. WANG: Actually, Bob didn't show the full results. Actually, one of the sensitivities we did is on the dropout. We actually had two scenarios for the dropouts and Bob only showed one. Actually, that is the ultimate determining power for the efficacy and one of our conclusions is based on that sensitivity analysis because we showed that the ultimate efficacy result was heavily dependent on the dropout. The major reason for the dropouts is adverse events. Therefore, one of the conclusions is saying that the ultimate dose selection may be dominated by toxicity which would drive the dropout instead of the efficacy. We did run the sensitivity on the dropouts for the HIV case.

DR. POWELL: Also, we are limited to six weeks so the clock is running. At that point in time Yaning had a couch in his office--

[Laughter]

--so it does limit some of the things that you might do. One of the things--I don't know why, but it was amazing to me that at the FDA the cycle times are much shorter than the industry. So, the longest cycle time on a project --I mean, it is six weeks on these but for an NDA people have to do their work really in seven months or less. In the industry you have years to sharpen these sorts of tools.

DR. BARRETT: Let's do one last question. Edmund?

DR. CAPPARELLI: Actually, it is two parts. One of them, hopefully, is quick. If you could expand a little bit on the level of interactions, both during this process iteratively with the company, the sponsor and review groups, is this something that, you know, has a lot of interaction or is it really done at the level of clinical pharmacology?

Then the second, which is just a specific clarification of an example that I am very interested in; it was very nice in the presentation, the rebound on the

model for the viral load, is that because therapy stopped or was there some sensitivity in terms--

DR. POWELL: It was a Phase IIa trial, monotherapy with the drug. So, it was really a proof of principle trial. At the end of ten-day dosing they stopped the drug and you could see the rebound, which is informative information. The other question you asked?

DR. CAPPARELLI: The interaction. This was done in a short time frame and, obviously, the more interaction one has--

DR. POWELL: That is right.

DR. CAPPARELLI: --the more global the model but, in a sense, you are limited as well in terms of getting results.

DR. POWELL: Yes, when there is a lot of analytical work to do we tend to be talking with the sponsor on the order of weekly, not more than every two weeks over that period of time. The sponsors have been surprised by that. There has been some resistance inside the FDA to that additional workload.

DR. CAPPARELLI: How about also within the disease groups?

DR. POWELL: That is what I mean.

DR. CAPPARELLI: So, the interactions then are at the level of the pharmacology group talking to the sponsor?

I am also interested in the FDA talking within the different groups.

DR. POWELL: Yes, usually when we are talking to them--it depends on the question at hand but we need to be engaging usually the physician that is responsible for that next trial and what they are thinking about. Another thing that we ask for is what is the product profile that you are looking for so we have a pretty clear view of where they are going with the next trial, and we want to be talking to some decision-makers in this thing. If we are not talking to decision-makers, then it doesn't work very well.

Likewise, inside they always want to make sure--the sponsors want to make sure that the division director or someone that is in the decision-making authority inside the FDA is engaged as well.

DR. BARRETT: The committee will have additional time for more specific questions, and we have been charged to answer a couple of questions from FDA. At this time, let's move to Dr. Yaning Wang to present a case study: a quantitative approach to assess a genomic design and a biomarker titration design for a Phase III clinical study.

Case Study: A Quantitative Approach to Assess a Genomic Design and a Biomarker Titration Design for a Phase III Clinical Study

DR. WANG: Good afternoon, everyone.

[Slide]

I am sure all of you must know that the FDA critical path initiative identified many opportunities to improve the overall drug development and our review process. Some of the opportunities involve the application of pharmacogenomic information and biomarker data in trial designs and model based drug development. Today I will present a case that will include all these components, specifically a quantitative approach to assess a genomic design and a biomarker titration design for a Phase III clinical study. I believe that by the end of this presentation you will be convinced that FDA is trying to apply all these approaches to improve the overall drug development and review process, and we are here today to ask for your advice to help us do a better job.

[Slide]

In the next 30 minutes or so I will start with a brief introduction about the background of this new drug product; followed by some clinical pharmacology features that are relevant for this case. Then I will present how we developed two potential trial designs, specifically stratification by genotype design and titration by a biomarker design based on pharmacokinetics and pharmacodynamics modeling and clinical trial simulation.

Finally, I will summarize our findings from the simulation and list the questions for the committee.

[Slide]

Drug X was developed to treat a chronic disease and early Phase I PK studies have shown polymorphism in a metabolic enzyme for this drug. Seventy percent of the patient population is classified as extensive metabolizers, or EMs, and 30 percent is poor metabolizers, or PMs. The mechanism of action for this drug is to reduce the level of a biomarker which will eventually reduce the level of a surrogate endpoint. So, the question was how to manage a genotypic influence on drug clearance in dose selection in a Phase III clinical trial design.

[Slide]

The early Phase I PK studies have shown that the genotype difference caused a significant difference in the drug exposure among the three genotype groups, as indicated by this plot. As we can see, the PM group achieved a significantly higher drug AUC compared to the EM groups, and this pharmacogenetic difference is significant enough to cause a difference in the surrogate endpoint.

[Slide]

As we can see, due to the high exposure in the PM groups more response or more surrogate change from baseline was observed for the PM groups compared to the EM groups,

and the response observed for the two genotype groups within the EM category is comparable.

[Slide]

In order to quantitatively incorporate the influence of genotype on drug exposure and subsequently on the response into the Phase III trial design, we modeled both the pharmacokinetics and pharmacodynamics for this drug. Specifically, we used the Phase I data to establish a population pharmacokinetic model and then used Phase II data to update this PK model.

Due to the lack of long-term efficacy data for drug X, we used a dataset from another drug of the same class to describe the relationship between the biomarker and the surrogate by simultaneously modeling these two endpoints but the final model was updated with the limited drug X data.

[Slide]

We used the model first with our assumption to describe the pharmacogenetics for this drug and based on the mechanism of action the concentration of the drug is reducing the elimination rate of the biomarker, and the level of the biomarker is determining the production of the surrogate. As indicated by this simplified plot, there will be a delay between the drug concentration and the biomarker response. The drug's action on the ultimate

surrogate level is further delayed. We will see later that these delays are important factors to explain the difference between the genotype and biomarker designs. These are the equations that can be used to describe how the biomarker level and surrogate level change over time.

[Slide]

These are the modeling results for drug Y. Based on a dataset of 900 patients for drug Y, both the biomarker and surrogate were simultaneously modeled. The top two plots are for the biomarker and the bottom two plots are for the surrogate. As indicated by the diagnostic plots of individual predicted versus observed levels, the model can describe the data reasonably well on an individual level.

[Slide]

Then we generated a hybrid dataset that included 400 patients from drug X and 100 patients from drug Y. As you can see, we only have 12-week efficacy data for drug X. We are trying to borrow some long-term information from drug Y. The final parameters were updated with this hybrid dataset.

[Slide]

These are some individual fits for drug X. The black lines and circles are the predicted and observed results for the biomarker level. The red lines and the triangles are the predicted and observed levels for the

surrogate levels. Overall, the model can predict the individual data fairly well for both biomarker and surrogate.

[Slide]

Then all this model was used in a clinical trial simulation. Specifically, we used the two-compartment model for the population PK to generate the drug concentration with the drug clearance dependent on the genotype. Then an exposure-response model was used to link the drug concentration to the biomarker reduction which will eventually reduce the level of the surrogate. Now we had these two trial designs in stratification by genotype and titration by biomarker. I will explain the details in the next two slides.

The inclusion criterion for this clinical trial is that the patients should have a baseline surrogate level between 70-100 units. The final analysis plan is a simple response rate summary at week 26, and a surrogate responder is defined as a patient with surrogate reduction more than 10 units at the end of week 26 therapy. We simulated 100 clinical trial for each scenario.

[Slide]

This is the scheme for the stratification by genotype design. All the patients will be randomized to four groups, one placebo and three different dose levels.

At each dose level the patients will be first genotyped to determine what dose to take. For example, at the 40 mg dose level the PM patients will take 40 mg but the EM patients will take a higher dose which is 120 mg and the dose will be fixed for the whole time period of the trial.

[Slide]

In the titration by biomarker design all the patients will also be randomized to four groups. At each dose group, however, all the patients will start with the same dose. For example, at the 40 mg dose level all the patients will start with 40 mg. But at week 12 the patient's biomarker level will be evaluated to determine who are the biomarker responders and who are the biomarker non-responders. A biomarker non-responder is defined as a patient with biomarker reduction less than 13 units at week 12. Then, those biomarker non-responders will triple their dose and keep the higher dose until the end of the trial. The biomarker responders will keep taking the original dose for the whole time period of the trial. For both biomarker and genotype designs different dose regiments, like BID versus QD, will also be evaluated during the trial simulation.

[Slide]

This is the response rate results at week 26 for various dose regimens. The blue bar is for the genotype

design and the green bar is for the biomarker design. I want to make three points from this plot.

First, we can see a clear dose response for the BID regimen within the genotype design. This is also the case for the QD regimen and also for the biomarker design.

[Slide]

Second, let's look at the response rate at one dose regimen, BID 10 mg. As we can see, the response rate is higher for the genotype design than the biomarker design, and we will see a similar patter for all the other dose regimens. The reason for this difference is that most of the patients started with the right dose at the very beginning of the trial for the genotype design, but in the biomarker design most of the patients are under-dosed at the very beginning of the trial and did not get the right dose until 12 weeks. Therefore, if we can have a longer trial we would expect that the difference in response rate between these two designs would get smaller.

[Slide]

For example, if we can have a trial for 38 weeks, then the response rate difference between these two designs will be much smaller at week 38 compared to week 26. Basically, it takes time for the surrogate level to reach the new steady state under a given dose due to the delay in response we explained earlier in that simplified plot.

[Slide]

The third point is that the BID regimen is always achieving a better response rate than the QD regimen, as indicated by BID 10 mg versus QD 20 mg. This is true for both biomarker and genotype designs.

[Slide]

But the advantage of BID versus QD is also dependent on whether a patient is a PM patient or an EM patient. From this plot we can see that the advantage of BID versus QD is much more obvious in EM patients compared to the PM patients, and this is consistent with the faster clearance of this drug within the EM population.

[Slide]

Even though no major safety endpoint has been identified for this drug at this stage, based on these designs some PM patients could have exposures higher than any observed levels in the earlier studies. Here is the proportion of PM patients that could receive high doses that have never been studied. It is about 3 percent for the genotype design and about 30 percent for the biomarker design.

In the genotype design this is mainly due to the misclassification of PM patients as EM patients. But in the biomarker design this is mainly due to the between subject variability in both the pharmacogenetics and the

pharmacodynamics for this drug. As a result, these PM patients may experience unexpected safety issues under this very high PK exposure.

[Slide]

So, based on all these analyses we concluded that at week 26 the stratification by genotype design will have higher response rates than the titration by biomarker design but this difference will get smaller at later weeks during the trial.

Also, we believe that the BID regimens perform better than QD regimens, especially in the EM population. This conclusion helped the sponsor to confirm the plan for a sustained release formulation for this drug.

Also, we believe that before the Phase III trial is started the high dose safety data in the PM population should be collected in order to avoid some unexpected safety issues in this population under high PK exposure.

Finally, we believe that the biomarker-surrogate relationship established in this case can be applied to other drugs that have similar mechanisms of action.

[Slide]

Before I present all the questions to the committee I would like to show you this summarized post-meeting evaluation from the sponsor. It is similar to what

Bob showed you earlier so you can evaluate it in a different way if you think it is biased.

Basically, there were nine sponsor attendees from the sponsor side, three from the clinical side, one regulatory, two project management, two biostatisticians and one clinical pharmacologist.

For the question of how valuable did you find this meeting, we got a mean response of 4.3. The range was 1, meaning worthless, to 5, meaning pivotal. The clinical folks believed that the meeting was pivotal because they were at a junction in development that required detailed discussion and feedback from FDA. The biostatistician responded also positively and he said the technical discussion about the modeling simulation approach is very helpful and high quality, and they are glad to know that FDA is supporting the use of new metrics in drug development.

In terms of whether the meeting changed the development plan, seven answered yes and one said no. Both the laboratory and clinical team believed that the modeling offered alternative development scenarios that could ultimately change the development plan.

Eight out of nine believed that this meeting will have value in decreasing Phase III attrition and designing better Phase IV trials with a better success chance. Also,

the majority of them believed that the time and effort required to prepare for this meeting was adequate and worth the results.

[Slide]

Finally here are the questions to the committee: We would like you to comment on the quantitative approach used in this case study, including both the pros and cons. We also want you to comment on how we could incorporate and evaluate genotype clinical trial design recommendations in the following scenarios, metabolism genotype; pharmacodynamic genotype; disease genotype; and narrow versus wide therapeutic index. Thank you.

Committee Questions to Speaker

DR. BARRETT: We have time allotted later in the program to address these questions, so why don't we open this up now for questions on the specific case study? We will go right to Dr. Kearns. You must have some questions.

DR. KEARNS: No.

[Laughter]

DR. BARRETT: Can we note this in the minutes? Bill, go ahead.

DR. JUSKO: I have some comments that relate to what you have on the screen but these are my original questions in the first place. This is a very nice modeling approach. It is a superb modeling approach. I have seen

the same model applied previously by several companies to modeling glucose and glycosylated hemoglobin, and we once used rosiglitazone and have applied it for a couple of other companies' drugs as well. So, it is a beautiful model that looks very nice.

As you were showing the data, it became clear that the changes in glucose and glycosylated hemoglobin that you were seeing were not any better than those from rosiglitazone. So, my question--maybe you don't want to hear this--is why would you bother suggesting that a company continue further if there is a genotype phenotype question for this drug, creating a complication that doesn't exist for the drugs that already work very well in this class?

DR. WANG: I masked the magnitude of the response for the drug. I am not sure you can read those numbers and compare them with the approved drug. But, so far, the sponsor believed the drug is going to give a better--well, at least will beat the placebo with the defined margin that FDA has. More importantly, we identified a major source for the pharmacokinetic variability. So far, we haven't found a major safety issue for this drug at the given doses. So, I think they believe there is some value in developing this drug but I won't comment on whether this drug is almost as good or not as good as approved drugs.

DR. POWELL: Could I address that? While we were completely cognizant of what the effects of the other drugs were, we don't really get into a discussion at this meeting in terms of what your drug is going to do versus other drugs that are on the market unless their development plan was to beat other drugs, let's say to be as efficacious and safer, or something like that. That was in their development plan. Anyway, we probably shouldn't talk about what the other characteristics of this drug would be.

DR. BARRETT: David?

DR. D'ARGENIO: A couple of points. One of the other reasons you saw a difference between the biomarker feedback strategy and the a priori strategy was where you chose also to make that biomarker measurement at 12 weeks. Because of the inter-subject variability and the biomarker dynamics and the PK, undoubtedly, you would see them or not see them. So, that critically entered into why you are seeing the difference here between the feedback and the a priori, as well as the horizon of the whole study.

But my main point is that in presenting these results you showed that in the end you looked at the average or the mean response value. And, one of the real benefits of these kinds of simulations where you incorporate everything you know about the inter-subject variability, adherence, dropout is that you can also look

at the distribution in this case of the surrogate marker in the whole and that is much more telling than looking at the median.

As Stephen Jay Gould reminded us, the median is not the message in these kinds of studies. We want to ask what is the probability that the surrogate marker will be above or below some critical value. That is a very big utility of doing these kinds of studies. It begs the message of how reliable is the certainty that you put in the model, but that is okay, I mean, we have to answer those questions anyway. But by looking at those kinds of probabilities you can look a lot more than I think looking at just the median response of the competing strategies.

DR. WANG: Well, I think the response rate itself is already a probability above a certain cutoff value.

DR. D'ARGENIO: Maybe I didn't understand it, but it is not showing you the response rate in the distribution in the population. So, if you took your one strategy, your a priori adjustment, you are going to get the distribution of the response in the whole population.

DR. WANG: Well, I guess you are trying to see what are the ranges of response rates because right now I only showed one response rate. You are asking--let's see, I repeated the clinical trial 100 times. For each clinical trial I get a response rate. You would like to see the

distribution of those response rates across the clinical trials instead of the mean of all the response rates.

DR. D'ARGENIO: Yes, in part but what we ultimately want, of course, when we do this in the population at large is what is the distribution of this response in the population. There are going to be some subjects that don't respond or have a low response, some that have a high surrogate--I won't even use response rate but surrogate marker. So, what is the probability?

DR. WANG: We could have included a plot to show what is the distribution of surrogate levels at the end of the trial.

DR. D'ARGENIO: That is right. That can be very useful.

DR. BARRETT: Greg, did you have a question?

DR. KEARNS: Yes, thanks. As part of the case you mentioned that one of the decision points after the simulation was to recommend that the sponsor go ahead and collect side effect information data on the PMs who had received the highest dose of the drug.

Now, as I recall, the incidence of PMs was fairly high in this example, 20 percent or so, and we spent the morning talking about warfarin. If this example were warfarin where the bad side effect potential was real, could you see this process maybe making a different

recommendation to a sponsor, which is don't give the higher dose? What I am saying is that the old approach would be to make a plan. We are going to give doses. We are going to bash through it and see what we find. But if it was the sense of FDA clinical pharmacology in doing this exercise that there was potential harm, would you go so far as to go to the sponsor and say don't give that higher dose; we really think that the potential for badness is there?

DR. WANG: See, so far we don't know. They haven't seen any major safety issue under the current dose. But they are planning to give this kind of dose to, let's say, the PM patients and we were asking them to at least use a smaller population to test. If there is a serious safety issue, then we probably would ask them not to do this but so far we don't have the kind of data to support that decision.

DR. POWELL: If I could address that as well, so, it would be usual at this phase when the full safety profile is not known--in fact, that might take years. Right? What we were concerned about was that if the sponsor chose the development path to genotype first and then put people on doses, that there is always the chance of mis-specification either from the assay, or whatever. So, we thought that it is worthwhile even in that scenario to know what would happen to a PM that is given an EM dose

sometime in the development plan. I think that is partly what we were thinking in that recommendation. What do you think of that?

DR. KEARNS: I think the answer is it depends on what your adverse event is.

DR. POWELL: Well, it was clean. In this scenario, the drug is pretty clean at this point. Obviously, if people had been dropping like flies you wouldn't want to do that.

DR. KEARNS: Right. I am trying to think about it in the context of our IRB as they would make this decision given that information. I could see some people being very pensive about the chance of badness happening, especially if a priori you have a signal that says it may happen. You know, it gets back to a lot of the morning discussion of the value of this. It depends I think.

DR. POWELL: Right.

DR. BARRETT: Edmund?

DR. CAPPARELLI: Just one other point, it really comes up in this example as well as the other, but one of the things that was implied but not explicitly--well, maybe explicitly stated but maybe not emphasized--is that these processes really show where the holes are in terms of knowledge that are critical for the pathway development. A few of the things that you mentioned in terms of one of the

conclusions going on to a longer study, well, that is really going to link to some of the assumptions of dropout, adherence, and while you didn't go into those models, clearly those are going to be components that come into play. Toxicity obviously is one. But, you know, as you go further on, waiting that long for when patients have feedback in the case of glucose monitoring for dose changes to occur, is going to also complicate these simulations and really, the real-world outcomes of these trials.

DR. POWELL: It has been a while since I have thought about this, but for this indication sponsors generally have to do 26-week studies. So, the reason the 26 weeks was there was partly due to convention. That is the conventional standard to get a drug approved for this particular indication. So, that is why we were playing that game. You know, if you could say to the sponsor, well, you need to put these on for, like, 42 weeks, if they don't need to do that they probably wouldn't do it.

DR. CAPPARELLI: I guess my point was that the comment that if we went out more than 26 weeks, you know, you are maybe falsely leading them to think that those extra 16 weeks or if you took it out longer you would have activity when, in fact, some of the other models that are the least well understood about the dropouts, the feedback and how the patients respond become bigger and bigger in

terms of your outcomes. So, you have to be careful in those realms. Again, the idea that it points out that you need more information about the toxicity, adherence and dropout stuff I think is key.

DR. WANG: I think that is a very good point. Actually, in this simulation we did not include the dropouts. So, therefore, the conclusion is based on perfect adherence but in reality that probably never happens. In this case we did not have, like in the HIV case, a dropout or adherence model that we could borrow from other drugs. Therefore, we just assumed this and gave sort of a clean picture but, absolutely, your comments about the reality definitely should be incorporated in the modeling. Also, the conclusion is dependent on, well, this is perfect adherence.

DR. CAPPARELLI: Well, it is biased and a certain group is going to dropout more.

DR. WANG: Yes.

DR. FLOCKHART: And that might have affected the fundamental results. It is possible that a genotype drops out more than another.

DR. BARRETT: If there are no questions--

DR. SINGPURWALLA: I do have--

DR. BARRETT: Sorry.

DR. SINGPURWALLA: We were asked to comment on the quantitative approach used, but before one can comment, certainly myself, we need a better understanding of the approach. Now, I get the general impression that you have a differential equation on slide number seven which seems to drive the quantitative approach. I also hear from my colleague on my left that this kind of differential equation has proven to be valuable under many, many other scenarios. It seems that you are hanging your hat on one hook, namely, this differential equation.

Second point, you have a slide there which compares the predictive versus the actual. Now, the question that goes through my mind is are you using this differential equation on a set of data, training the equation with respect to the set of data and re-predicting with respect to that set of data, or are you using this differential equation to predict something in the future that you have not seen? That is the second point.

The third point is that I completely agree with my colleague at the end that you should really give a predictive distribution rather than just give a mean or the median. You should give the entire distribution if you are doing any kind of simulation.

But the most serious question is are there other models, other than this differential equation, that could have been used and could have perhaps been better?

DR. WANG: I can't remember all the questions. For the last question--that is the one I clearly remember, we reviewed several similar models in the literature and this is the one we selected. Yes, there are many models that can be used to approximate--because we all know all models are wrong but as far as the model is useful and can predict what we need, that should be enough. So, we compared different models and this is what we finally chose.

In terms of how we predict the data, we are not predicting outside of the data observed because we have a large database from drug Y from which we developed this original structural model and also the parameter estimates. What we did next is only use drug X limited, short-term 12-week data to update those parameters. Even the structural model is fixed. The only thing we are changing is the parameter estimate and also the variability, the between subject variability, because we are probably dealing with a different population in this new trial. So, those are the parameters we updated based on the new dataset. That is where our prediction came from.

DR. SINGPURWALLA: So, if you were confronted with a new dataset there is no guarantee that this particular model, without being updated with respect to the new dataset, would give you a good prediction. Right?

DR. WANG: Could you repeat that again?

DR. SINGPURWALLA: Let me just summarize. I think what you have done is fitted this differential equation to the existing data, and the fit is very good. It doesn't mean necessarily that the predictability of this model is really good. The fitting is good.

DR. WANG: Well, see, the fitting on the existing data is mainly between the biomarker and the surrogate relationship. We believe that based on the mechanism of action the drug is affecting the biomarker. So, the relationship between the biomarker and the surrogate should be the same. Therefore, we have confidence in the model that fitted the current, like, large database in terms of fitting the future dataset for drug X.

DR. JUSKO: Can I comment and ask a further question? I am not sure you answered correctly when you said that there are other models that you tried that could have been used because this would be the simplest one that has a mechanistic basis, because if the biomarker was glucose, then there is production and there is utilization of glucose. So, a turnover type of concept is the starting

point for any model and any other ones would get more complicated. If the surrogate is glycosylated hemoglobin, which also is the simplest possible mechanistic model for glycosylated hemoglobin, more complex models would exist and you could factor in insulin and the feedback that happens between glucose and insulin, but it probably is not necessary for this drug's mechanism of action. So, I am not sure you are being exactly accurate.

DR. WANG: Well, see, that is why we did not select those models. First, we did not have enough information to fit those models. Second, we think this should be enough or sufficient for this application.

DR. JUSKO: Okay. My further question is did you also utilize information in the FDA data banks for other drugs that behave similarly where you could have provided many of the parameters based on previous quantitative information about glucose and glycosylated hemoglobin? If you did that, then you would be less dependent on just fitting the current dataset because the current dataset was incomplete because you were not showing a rebound like your simulations here, and you may not have had enough power to really resolve all the parameters.

But this is a very nice model to build upon the concepts that Dr. Powell described of merging quantitative information from the disease drug effect databases of the

FDA, and expanding upon that for application to additional drugs.

DR. WANG: I think we did use a much larger database from drug Y to develop this model. Right? Instead of only fitting drug X, a lot of weight is actually on the drug Y database. It is like 900 patients from drug Y.

DR. POWELL: Drug Y was another NDA. It was not used to construct the model. So, we had that data. We applied the model to it and that is what that one figure spoke to in terms of showing the degree of prediction that existed in that new dataset, which I thought was really cool because probably the only place you could do that is at the FDA.

DR. BARRETT: Did you have a question?

DR. DAVIDIAN: Part of my question just got answered, actually. I guess following up on something, in regard to the parameter distributions, and so on--given you fixed your model, which I think is perfectly reasonable, you have information on all these parameters and their distribution in the population. I guess it is a follow-up to the question I asked earlier, you know, that information, of course, in itself is somewhat imperfect because it is coming from fitting these models to some sort of data. So, my question was to what extent did you try

varying some of those assumptions? Maybe increase the inter-individual variation by 10 percent in all your parameters or those kinds of things just to see what the effects might be on the ultimate outcomes and the inferences you drew about the trials?

DR. WANG: That is a very good point. Actually, we did try the sensitivity. You are talking about sensitivity. The between subject variability on this EC-50, that is basically how sensitive a patient is to a given concentration of drug, and we tested if we increased or decreased the between subject variability for this parameter and also for the E-max, the maximum effect the patient can have, we did test the sensitivity of those two on outcome. Overall, the pattern would shift but overall the result will be the same. Actually, if you think about genomic testing, it only takes care of the between subject variability in the pharmacokinetics. When you have the exact same concentration a group of people can still have a very different response. That is the pharmacodynamics between subject variability. That cannot be taken care of by genotyping. That is where we put the between subject variability on these two parameters but we did not look for covariates, like another genotype to explain why this patient is a responder and the other is not. Yes, we didn't do that.

Commentary on the Case Study

DR. BARRETT: If there are no further questions, we are going to come back to the specific questions but prior to the presentation, actually at the time that the background was prepared, I was asked to provide a commentary on the case study.

Most of the points I am going to raise you have already addressed so I am going to try and go through this relatively quickly and just highlight a few points. Again, these are based on the actual background material and not the slides that you presented.

First of all, in the material you provided you described three issues described in this case study, one being the difference in kinetics between populations with different genotypes will lead to clinically significant differences in drug response. Two, the response of a biomarker to the treatment is slow and traditional titration based on a biomarker level requires long time of treatment. Three, there is a need to get the dose right as quickly as possible in the early treatment to ensure the maximum effect on the surrogate endpoint can be observed at the end of the trial. That is 26 weeks.

All of these are extremely well suited for this paradigm in the clinical trial simulation context. I really thought that the case study was very much in line

with the ability to provide the sponsor with good guidance as far as, you know, the next series of experiments.

I did have on issue with that third point about getting the dose right piece of this because it seemed very much attached to this 26-week trial which, as Bob pointed out, is pretty much the next design construct that the sponsor was being asked to provide as the basis for an approval.

That being said, the only thing I thought guidance to the sponsor would have been considerations of other dosing scenarios that may allow you to see this difference between PMs and EMs in order to look at the true response rate.

I thought what you had proposed in terms of this enrichment trial was right on, but somewhat predictable based on the kinetic differences. The other thing I think I wanted to mention was that the other big advantage of a trial simulation approach is the ability to explore population disease and design dependencies all at the same time, which you did a nice job of. So, I think, again, this is very well suited to the approach.

One of the other things I thought, in the context of this morning, was that it did give you an opportunity to explore this genomic kinetic piece very early on. In the morning we had discussions about adding sections to

labeling and specifically on the administration section, so one can see that you could use this as an ability to do labeling exercises right at the time you were designing that pivotal trial. So, I thought, again, there would be tremendous value in that.

As far as points for consideration, again a lot of these were brought up, I thought the criteria for declaring a superior trial really wasn't well defined, at least in the example. You have summarized the data as far as response rates very nicely and it is very clear. But one of the things I though might be helpful more as an exercise would be can you a priori define what is going to be your criteria for a successful trial design, or a scenario relative to others.

You know, the data is summarized very nicely and I think it is very clear so I would agree with all of your results and characterizations on the conclusions. However, on the topic of a priori criteria and deciding which of these trials had superior characteristics, maybe there could have been some intermediate metrics. We talked earlier about this iterative value of pharmacogenomic tests identified as metrics possibly being incorporated into this. Again, it is really outside maybe of the guidance that you wanted to provide to the sponsor at this given time but it is something you could have explored with

the trial simulation piece. You get a lot of nice additional benefits from going through this exercise.

One other thing I thought is that the genotype incorporation is really very compelling if the outcome-based hypotheses are defined as part of the trials. Now, I know at this stage of the game you are really trying to convince the sponsor to do this in the first place which, again, I think is very clear in this case. Could you, in fact, have made this part of the hypotheses in the design to actually do separate randomized by genotype? I don't know if that was part of the original construct. You have done parallel group analysis in the genotype-based trial, but was there any assurance in terms of the distribution with the genotype? That was just one question I had.

Then the other point I wanted to make is as far as the assignment of pharmacokinetic characteristics with genotype. You know, I think you had used clearance as far as the kinetic parameter to differentiate the genotypes. I know, having gone through this in my own exercises, that it may be more helpful to, in fact, focus on the metabolic pathway and use something like fraction metabolized as the kinetic parameter of interest, as opposed to clearance because it tends to lump other elimination processes. So, just a comment as far as that.

But I really thought this was an excellent example of providing guidance at a very early stage in the context of modifying a development program, and I was encouraged to see a glowing response from the sponsor at that stage.

So, at this stage we do have a slot available for open public hearing. No one at this stage has signed up to do this but if there is anyone who would like to come and present at this stage, the floor is open to you.

[No response]

That being the case, we will take a break.

[Brief recess]

Committee Discussion of Questions

DR. BARRETT: Yaning, can you put the questions back up? The committee has been asked to address the following questions from the second topic: The first question is what are the committee's comments on the quantitative approach used in this case study? I will open it up to the floor.

DR. JUSKO: I think your approach was outstanding. This is an excellent demonstration of use of PK/PD modeling in advising on a study design. In fact, it probably can be carried a little bit further than you have carried it. This kind of model has also been used to identify probably most efficiently a proof of concept

study. The modeling can be used to determine just how soon, if you have a new drug, could you see a change as an initial signal that the drug has potential for further pursuit. So, basically I like it a lot.

DR. BARRETT: Dr. Singpurwalla?

DR. SINGPURWALLA: Since you ask me, I will tell you. I like the idea of doing what has been done, namely, taking a model, simulating it, comparing the model with either actual da or observed data. I like the general concept and general principle.

I cannot comment on the quality of the model, other than the fact that what I hear is that the model has been used under many scenarios. I don't quite understand what is the basis of the model; what equations--I mean, it is a differential equation, what are the basic conditions that drive it; what are its initial conditions; what are its input parameters, and all. So, I don't understand the workings of the model.

So, I will answer the question in the following way, that I like the concept but I cannot comment on the model itself and exactly what it is that you have done with it.

DR. BARRETT: Let me just address your comment. As far as the approach goes, I mean, I thought that the approach was exactly married up with the questions that the

sponsor had. As far as the model goes, one of the things I thought that the critical path document did an excellent job with was framing this approach as a useful technique to communicate decision-making. So, from that standpoint, I thought that this was aligned very nicely with what is identified in the critical path.

In fact, the comments back from the sponsor really highlighted the fact that they saw the merit to this approach, and also the fact that they felt that they needed to be more prepared the next time. So, I would be very curious, as you get more of the repeat customers to the end of Phase II meeting, what their perception is the second time around, after they have had one of these interactions with you. One of the interesting metrics I think would be for you, at the agency, would be to what extent, when you deliver these models, do you feel that they are actually used in the hands of the sponsors once you, in fact, provide them.

DR. LESKO: I am going to add that the end of Phase IIa meeting and the concept of model-based drug development is really work in progress. Some of you may remember that we presented this in November of 2003 to the committee as a concept. At the time, I remember some of the comments from committee members were that this sounds like a good idea but I am not sure why. So, we embarked on

the actual end of Phase IIa meeting. As Bob mentioned, we have had I think six meetings, 12 requests and a couple on the boards, three coming up.

They differ quite substantially from any other meetings that FDA has with sponsors in that they are relatively flexible in the conversations that go on. And, it is going to take a while to see what the benefit of these meetings is. So, when Bob presented that HIV model, for example, we are working with very early data so that the predictability in terms of probability of clinical outcomes that would occur 42 weeks later--technically, the way this would work would be if the company came back with an end of Phase II meeting whereby these models could be updated with additional information that came out of those trials, then presumably you could iterate these models with new data all the time to improve their predictability of what you ultimately want to know as, in that case, the 48-week data.

So, by nature, these meetings are going to take time to evaluate but we wanted to bring this to the committee basically two years later to share with you what has happened since we introduced it as a concept. While the feedback from the companies, as you saw, was very positive, the ultimate feed back that we would like to see is that they have actually reduced attrition in the

clinical trial area in the disease states that we focused on.

Getting back to that question about the model that, Nozer, you asked, I think that models have differed in their complexity in different therapeutic areas. So, what we saw in the last case was relatively, I would say, more empirical and maybe semi-mechanistic. What we saw in the HIV case was more mechanistic. This has been the trend across the therapeutic areas that was on one of Dr. Powell's slides where he listed all the therapeutic areas that have been the subject of IIa meetings.

But I think the idea in the future is to develop specific disease state models that are mechanistic, that have as a basis disease progression parameters that can be utilized with a dose effect model to come up with predictions in a better way.

DR. SINGPURWALLA: Excuse me, can I react?

DR. BARRETT: Sure.

DR. SINGPURWALLA: I am making a general comment, not particular to this one. I have been to several meetings of this kind, not particularly this committee but other committees. Sometimes very complex, technical, hard work is presented in about half an hour and the committee is asked to make a comment, this particular case being a

case in point. I am sure you have thought a lot about it. I am sure you have done a lot.

I have no chance to sit down and evaluate it. I am asked to make a judgment quickly, and I don't think that is something I would like to do. I would like to ponder it, think about it, probably make comments and make a general assessment as to whether this is on the right track or not, or whether I like it or not.

So, it is a general philosophical question about how committees like this operate and something very technical comes up. A few weeks ago we had another meeting on sample size inhaler amount of dose. A lot of statistical issues were there and we needed more time to really think and evaluate these things. So, I don't know if there is going to be a policy change about how these things should be done.

DR. BARRETT: Just as a response to that, I think your point is well taken as far as those instances where the format just doesn't allow a critical evaluation. But one of the things I think, hopefully, was in the spirit of this was to comment on the approach and not so much the actual nitty-gritty details of the analyses because some of the questions that David actually raised about simulating the extremes and the distribution criteria--you know, I had similar comments going through this but I didn't go into

that kind of detail because, again, we were trying to focus on the approach and the dynamic between the agency in terms of using the approach in order to facilitate informed decision-making. So, that is where I got some comfort as far as not needing to know all of the details or the drug. Yes, Bob?

DR. POWELL: If I could, I would like to make a distinction that might help with what you are talking about. In Parkinson's disease, some of the people in clinical pharmacology and Bob O'Neil's biostat group and in CNS have been working on how to improve a drug that would change disease progression in Parkinson's disease. Okay? As far as I know, no drug has ever been given that sort of approval. So, that has resulted in going back to a lot of the literature and what a drug does that produces a symptomatic change versus a drug that actually produces a change in disease progression, using the UPDRS score. So, in a sense, disease progression sort of gets down to a slope question and simulations have been done there so the statisticians inside the FDA and the clinicians and Dr. Temple in this case have all noodled around this over a period of about three or four months. This gets at coming up with criteria that you would approve a drug for that indication. That gets to a much higher threshold of verification from statisticians and clinicians.

In this case there are two other situations--let me get back to the IIa situation. We are looking now at one in weight loss and another one in analgesia. Neither of these would be mechanistic type models. In the weight loss we have gone through and we have extracted from prior submissions what the change in weight is over time for different mechanisms. They tend to regress, as we all know, back to where you started. So, it is more of a description of what has gone before and looking at different mechanisms and, in the time we have available, being able to describe it as best we can. So, it is not as well qualified as what I think you are talking about. That is what we were really trying to describe here, if that distinction helps.

DR. BARRETT: David?

DR. D'ARGENIO: Yes, just a couple of comments. I certainly think the approach that is being taken here is exactly right on. It is the logical culmination of what a lot of people have been working on in the use of modeling and simulation and drug development for a number of years. So, the approach is exactly the right approach. The examples that you presented seem perfectly good examples to illustrate this.

The fact that you can do this in six weeks just astounds me, actually, and I worry a little bit, and I am

sure you do more than we do, about the potential workload for this, and also how much the companies could bring to the table. Certainly, they have done some of the early development of the PK/PD models. I hope they share that with you. That would save you a lot of time.

But the bigger point is Larry's point about the need for having disease progression models that are specifically geared to answering drug development questions, as opposed to the very detailed, extremely mechanistic disease progression models that we might pose just to try to understand the development of disease. I make a distinction between those two, and that goes to the whole modeling philosophy but we really need these models that can help answer drug development questions, which means we have these timelines that we have to deal with.

DR. POWELL: I think part of Larry's nefarious strategy is that if we can do these sorts of things and stimulate people, you know, that this behavior might be good behavior for planning trials and getting companies to do it and bring back that sort of information, whether for special protocol assessments or NDAs, that would be a good thing. In fact, we have pushed the work back to a couple of companies to do the work.

DR. BARRETT: Larry?

DR. LESKO: The other issue that comes up in the context of model-based drug development for us, and maybe people have some thoughts on this, is related to the focus of efforts to do this within FDA. We have obviously limited resources. You heard what these meetings take so people, time, effort, etc. are at a premium. So, given the realities of that, which disease areas would be most suitable for dedicating efforts towards this model-based drug development the way that you have seen it? We have been thinking about criteria to make that decision, and the criteria might range from therapeutic areas where drug development has been less than optimally successful, say, where failure rates in Phase III trials are high--that might be a focal point. Another criteria might be the extent to which we have biomarkers available that point towards disease pathophysiology. That could be another area. Another might be the criticalness of the disease state in terms of finding appropriate therapies as opposed to less critical disease states.

So, we haven't quite focused on exactly where we want to focus our efforts, but anticipating that we are not going to spread ourselves across every therapeutic area, 15 or 16 within the Center, but focus perhaps on six therapeutic areas that would have the highest return.

The other thought we have had is how do we leverage what we are doing at FDA in this model-based drug development disease state models, and the extent to which one could imagine collaborations that could develop between FDA and industry but also between FDA and academia to try to focus on, what I think Dave said, models that are useful for the drug development questions as opposed to perhaps models that might of interest to, say, academic medicine, or something like that.

So, any thoughts along these lines--focal areas for emphasis in terms of disease of therapeutic areas, or any thoughts on ways that collaborations could be built to further leverage this particular initiative.

DR. BARRETT: That sounds like a great segue to the second question: Comment on how we would incorporate and evaluate genotype clinical trial design recommendations in different scenarios. There are four examples provided: metabolism genotype; pharmacodynamic genotype; disease genotype; and narrow versus wide therapeutic index. Comments from the committee?

DR. MCLEOD: In this context you may want to start--just following your comments--with very targeted collaborations in an academic environment because there may not be the right dataset in industry coming forward to you because of some of the sensitivities that you have

identified there. But even what we talked about this morning, you know, some prospective analysis with warfarin may allow you to design a study that would be more of an indication or label changing study that might serve some of these issues as well. That is a very different model. I mean, there is no end of Phase IIa in academia--unless I guess if your grant doesn't get refunded--but there may be some real wins there that could not only feed your system and give it some priors for some of these other studies that are happening, but also help train another segment of the community, academics, in how to do some of these studies in a more meaningful way because often we do academically derived studies that seem great and get us some grants but don't necessarily make the world a better place.

DR. BARRETT: Yes?

DR. GAGE: One organization is the Society of Medical Decision-Making and they are very interested in doing mathematical modeling, including pharmacokinetic and pharmacodynamic modeling. A second thought is that in terms of leverage--I heard you use that word, you may consider particularly choosing areas where you have existing data or datasets, particularly those that are not publicly available. Therefore, in a very limited time the

incremental benefit would be substantially more than what industry could do without those data.

DR. BARRETT: Any other comments?

DR. JUSKO: There has been a lot of attention paid to metabolism genotype, as we talked about this morning. The areas of pharmacodynamic genotype and disease genotype are probably wide open frontier areas, and these areas, along with what Larry just described, probably would be a good subject area for a future meeting like this.

DR. LESKO: I guess what comes to my mind is that much of the genomics coming out of oncology is oriented towards disease pathway genomics as opposed to some of the things we talked about this morning. But even there, there are some very interesting questions about let's do tradeoffs between trial design that would be based upon enrichment biomarkers, for example, and how that might play into trial design, trial size or even dose selection to demonstrate efficacy in a disease where we typically have a low response rate in all patients.

So, that is a good point because as you go through the different therapeutic areas you may be more apt to focus actually on disease markers as opposed to the markers that sort of regulate dose exposure relationships.

DR. BARRETT: Larry, one comment I have on this topic is with the number of institutions that are doing DNA

banking, is there an opportunity there to leverage some of that information either to characterize certain populations of interest or, in fact, do prototypical studies that would serve as the priors for more informed designs? Just one suggestion.

DR. LESKO: Yes, that is a good point and, you are right, there is a lot of DNA being collected and stored, and it is actually hard for us to access when we want to access it for a variety of questions, usually revolving around drug safety questions. But these so-called bio-banks, patient registries that are being set up--we have had conversations with people that have developed and sort of controlled the flow of information through those and in the long term I see some possibilities there, but it is going to take some relationship building I think to access, and figure out what the questions are that we want to ask that would serve, say, a regulatory public health need versus, let's say, a drug development need in identifying new targets.

We have begun to explore the possibility of acquiring DNA samples in the post-marketing period by looking at patients that have experienced rare adverse events, and working with the physician that is responsible for that patient trying to get DNA samples that would allow us to begin to develop some associations. This is

relatively new as an area but it is certainly something that has a lot of merit.

In the same way, we have talked to people, like Kaiser and other people that have large patient databases, to look at specific questions. One can imagine from this morning that one can ask several questions about warfarin and the incidence of adverse events sort of looking at DNA to see what kind of genotype we have in those patients that were hurt by the drug, or something like that. So, we are in the process of developing these kinds of relationships and having conversations about as well as another way of gathering, let's say, association data between not only genomic biomarkers but non-genomic biomarkers and outcomes as well.

Part of the critical path, as people know who have read it, is a heavy emphasis on biomarkers for predicting things, things like what makes patients different from each other when it comes to a response to a given dose, both beneficial and adverse, and these are the kind of things that we are trying to address in the critical path, at least one of the things we are trying to address in the critical path area.

DR. FLOCKHART: Just trying to think from the perspective of the progress of the committee over the last several years, which has been very considerable, I think it

is remarkable that we have considered a series of drugs from a genetic point of view and made remarkable progress, I think, in the biometric marker modeling area, and a lot of this has been on drugs that are already approved. So, I think that is an important thing to think about in the context of what areas to focus on. If one steps back and thinks about warfarin, obviously a drug available for a long period of time but something that really presents a clinical dilemma, that is very difficult to use, I don't want to ignore the fact that the vast majority of drugs that I prescribe and that most doctors prescribe are generic and have been around for a long period of time. But in terms of thinking about ways in which this committee and the FDA clinical pharmacology division might interact with the large approval divisions of the FDA, I think that is a very important way to think about it, not just thinking about it from the point of view of a new drug that comes along that is interesting, that has interesting characteristics, but really a point where one can make an impact--pick the low-hanging fruit first. I think with warfarin you have done an amazing job.

DR. LESKO: As another observation, in contrast to approved drugs that are already on the marketplace--in fact, this committee has been somewhat of a home to discussing those types of drugs because it is very easy, in

the context of genomics, to get wrapped up in the new drug issue and how genomics plays into that. For example, in the area of oncology, by and large, with the new drugs many of those issues are discussed in the therapeutic area committee. So, if we were talking about EGFR inhibition, that goes to the oncology drug committee. But, on the other hand, drugs that are in the market and off-patent, not that they get lost in the genomic revolution, or whatever biomarker we are talking about, it is important that we have an opportunity to discuss them and this committee has been very useful in that respect to review the evidence on some of these older drugs.

DR. BARRETT: Dr. Sadee?

DR. SADEE: I want to comment on the metabolism genotype versus pharmacodynamic and disease genotype. I think, actually, the pharmacogenetics people have done a pretty good job in getting at the functionality of the genotypes of, say, the metabolizing enzymes, and so on. In the pharmacodynamic genotype area and disease genotype area there is a lack of good markers. If we think about what are some of the major problems in anti-schizophrenic drugs the efficacy may be 40 percent, or the patient doesn't get the right drug at the beginning, and yet there should be reasons for that and we don't have a clue what those are.

So, I think from a regulatory point of view, one would have to address the question what do we expect from a marker for these major questions that I think we have to confront next, namely, efficacy of CNS agents, or such. What are the quality criteria? Because what I do see more and more in the literature is that there is some experiment done in vitro where polymorphism of a promoter region is shown to be valid in the kidney cells and heterologous system and it may not apply at all to the CNS. So, we can only accept those biomarkers that really have some solid scientific foundation and they can be defined further. So, maybe that will be a topic for the committee, to say how do we define criteria in the future that will bring us forward in assuring that the efficacy is increased, which in many cases is only 50 percent for the very best drugs that we have.

DR. BARRETT: Any other comments?

DR. POWELL: Well, there is a piece that sort of we haven't touched on. To me, there are sort of three phases of a biomarker in a sense. I mean, either in an academic research place where they are looking for targets or a drug company where they are developing their initial concepts--they might have a number of different biomarkers that they are looking at and trying to sort out what does the model look like. They are running early clinical

trials with the disease and seeing what associates with outcomes. That is one phase.

What we are talking about here is a fairly narrow phase of the FDA looking at trying, on the one hand, to help people out use the information that they already have or that is available to make better decisions, and also to use it for regulatory decisions.

There is a whole other area that is being blocked currently, that I don't think we are talking about. So, when you come up with these biomarkers there is very little of this information that gets into labeling to use as a handle to help people tell who to give the drug to or how to dose it. I mean, I would say that that is systematically blocked by marketing from getting into the label.

We have run into a couple of examples recently. For example, in HIV we looked at the inhibitory quotient for protease inhibitor for resistant virus. You can explain pretty well the change in viral load relative to the inhibitory quotient, which is basically the trough drug concentration divided by the IC-50, like the MIC of the organism. So, we recommended, and it seemed to make sense, that this was the last drug that people could use at this point and to try and allow people to increase the dose if their IQ was low or if they were having adverse events

possibly decrease the dose. But, you know, you would have thought that that was just like the worst idea in the world, even though the sponsor's data spoke to being able to do things like that.

You can make the same case for therapeutic drug monitoring type of situations, which is just another form of biomarkers. There are all sorts of things that are useful in drug development that don't quite make it--and, in fact, they may jump over and academics start using it but it is not embraced in the label.

DR. BARRETT: Well, if there are no further comments at this stage, I would like to summarize what the committee's responses were. For question one, the topic of committee's comments on the quantitative approach in the case study, generally very favorable. The committee endorses the approach, the model specifics notwithstanding. It is a good concept and the logical culmination of work that has been ongoing for some time. I think we were all impressed with the time window over which this activity occurs. There is definitely a need for disease progression modeling which is useful for drug development, as opposed to perhaps more esoteric or academic endeavors.

As far as the second question goes, the committee's recommendations on how we would incorporate and evaluate genotype clinical trial design recommendations in

different scenarios, there were discussions about pursuing targeted collaborations with academic environments. Another idea is that we should areas where the FDA has existing data and expertise. The PD and disease genotype area is wide open and should be one that is perhaps pursued above the others. There may be some specific examples in the oncology disease pathway that may bear fruit. DNA banking may afford the ability to pool population characterization that might be useful priors in this endeavor. Finally, we should choose drugs for which the clinical dilemma is evident, as well as new drugs.

Larry, was there anything else that the committee can focus on at this stage, or would you like to wrap it up?

Wrap-Up of Day 1

DR. LESKO: No, I can't think of anything else but I thought I might maybe wrap-up with some ideas on next steps as to where we might go with some of the deliberations and recommendations that we had from the committee today.

I notice I have 25 minutes for wrap-up. I won't dare take 25 minutes to wrap this meeting up, but my wrap-up will really focus on I would say the three areas we touched on today. The first of those was the way that genomic information is reported in the label. Right now we

have a heterogeneous situation with regard to different approaches. I think that simply reflects the evolution of the science and the evolution of our experience in putting that information in labeling. Nevertheless, we had many good ideas on how to begin to create a framework for consistent introduction of genomic information in the label.

So, one of the things we will be thinking about is the possibility of developing a general framework and a set of recommendations for label language related to genomic information. We already have such a framework, I would say, in a paper that we drafted several months ago, called "A Co-Development Concept" paper which talked about specifically the co-development of a drug and a test. But it had, I think, a semblance of a framework that could be used for stand-alone tests as well as tests that would be developed coincidental with the drug product. So, I think we will be heading in that direction.

The second topic we talked about was more specific with the warfarin and the 2C9 VKOR evidence. The consensus that this evidence is compelling and sufficient to include in the label leads us to the next step of that, which will be conversations with our OND medical division counterparts, conversations with the sponsors of these products, and the whole purpose of that is to begin to

focus on the very specific language that would be entailed in an update, and also where that information actually might go into the label.

Finally, we talked about the model-based drug development pilot project under our critical path, and I think we had many good comments here. We have sort of taken an opportunistic view of this model-based drug development, basically taking on many different disease states and many different requests for the purposes of getting experience. As I mentioned, I would like to begin the focus in this area on the targeted disease models where the payoff would be the best, and we have had some very good suggestions on how to think about that. We will probably come back and talk about that again at a subsequent meeting, I would anticipate, to sort of frame those priorities as well.

So, in wrapping up, I would like to extend my thanks to the committee for very productive discussions and confirm that we have gotten a lot of valuable information from the discussion of the topics today. With that, I look forward to discussing our final topic tomorrow. So, I will turn it back to you, Mr. Chairman--Dr. Chairman or Jeff, as the case may be!

DR. BARRETT: I guess the only thing left for me to do is to formally adjourn. Thank you for your participation. See you tomorrow.

[Whereupon, at 4:30 p.m., the proceedings were recessed, to reconvene on Tuesday, November 15, 2005 at 8:30 a.m.]

- - -