DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

ADVISORY COMMITTEE FOR PHARMACEUTICAL SCIENCE

CLINICAL PHARMACOLOGY SUBCOMMITTEE

Thursday, November 4, 2004

8:05 a.m.

Hilton Washington, D.C. North
620 Perry Parkway
Gaithersburg, Maryland

C O N T E N T S

P R O C E E D I N G S

[8:03 a.m.]

DR. VENITZ: Good morning, everyone. For the second day of the Clinical Pharmacology Subcommittee meeting, we have half a day agenda for today. And I would like to point out that we don't have anybody signed up right now for the open hearing. If anyone in the audience wants to do that, please contact Ms. Scharen as soon as possible so we can lock you in.

The first order of business is to review the conflict of interest, and Ms. Scharen is going to do that for us.

MS. SCHAREN: Good morning.

The following announcement addresses the issue of conflict of interest with respect to this meeting and is made part of the record to preclude even the appearance of such. Based on the agenda, it has been determined that the topics of today's meetings are issues of broad applicability, and there are no products being approved.

Unlike issues before a subcommittee in

which a particular product is discussed, issues of broader applicability involve many industrial sponsors and academic institutions. All special government employees have been screened for their financial interests as they may apply to the general topics at hand.

To determine if any conflict of interest existed, the agency has reviewed the agenda and all relevant financial interests reported by the meeting participants. The Food and Drug Administration has granted general matter waivers to the special government employees participating in this meeting who require a waiver under Title 18, United States Code, Section 208. A copy of the waiver statements may be obtained by submitting a written request to the agency's Freedom of Information office, Room 12A30 of the Parklawn Building.

Because general topics impact so many entities, it is not practical to recite all potential conflicts of interest as they apply to each member, consultant and guest speaker. FDA

acknowledges that there may be potential conflicts of interest, but because of the general nature of the discussions before this subcommittee, these potential conflicts are mitigated.

With respect to FDA's invited industry representative, we would like to disclose that Dr. Paul Fachler and Mr. Gerald Migliaccio are participating in this meeting as nonvoting industry representatives acting on behalf of regulated industry. Dr. Fachler's and Migliaccio's role at this meeting is to represent industry interests in general and not any one particular company. Dr. Fachler is employed by Teva Pharmaceuticals, USA, and Mr. Migliaccio is employed by Pfizer.

In the event that the discussions involve any other products or firms not already on the agenda for which FDA participants have a financial interest, the participants' involvement and their exclusion will be noted for the record. With respect to all other participants, we ask in the interests of fairness that they address any current or previous financial involvement with any firm

whose product they may wish to comment upon.

Thank you.

DR. VENITZ: Thank you, Hilda.

Before we proceed with the scientific agenda, we will pay a tribute to one of the seminal members of this Committee, who passed away earlier this year, Dr. Lew Sheiner, and Dr. Lesko and Dr. Blaschke will pay tribute to his contributions in clinical pharmacology.

DR. LESKO: Thank you and good morning, everyone. Welcome back. We had a long day yesterday filled with a lot of heavy duty intellectual discussions, and it's nice to see you all back and I think refreshed.

Anyway, we would like to pause at this moment and remember our colleague, Dr. Lewis Sheiner, who was what I would call a founding member of the Clinical Pharmacology Subcommittee. I remember inviting him to join the Committee a couple of years ago, and he said to me I'll only come if it's going to be intellectually stimulating. And after each meeting, I would ask

him was that intellectually stimulating? And he would say yes, and he came back to every meeting.

Dr. Sheiner, as everyone knows, and Jurgen mentioned, passed away unexpectedly in April of this year, and Lewis, we all know, was many things to many people. He had an important role as a member of the CPSC. He provided us with an extraordinary dimension of opinions on many different subject matters, always challenging us to dig deeper into our intellect.

He was great as a member of this Committee. He focused on solutions, and he didn't dwell on the problems very much. I remember last November, and many of you do, too; we were discussing the end of phase two-A meeting, and I think we spent about three or four hours of discussion, and I still remember his question, which came at the end of that discussion, and I think it exemplified his way of spicing up a Committee meeting. He said Larry, it sounds like a good idea somehow, but I'm not sure exactly why.

I think that was his way of challenging us

to think clearly and fully about what we were proposing at this meeting. And I think the topic that we will discuss later this morning would have been very near and dear to his heart. So I know that I speak for many of you, members and audience alike, all of us at FDA, when I say that it would be an understatement of the highest proportion to state that Lewis is sorely missed today.

I have invited Dr. Terry Blaschke, who was a close friend and colleague of Dr. Sheiner to pay him a tribute on all of our behalf.

DR. BLASCHKE: Well, thanks, Larry. This actually is a harder talk to give than the one I'm going to give later this morning.

Larry did ask me to pay a tribute to Lewis, and I think we really did lose a visionary leader in drug development in April. Lewis died shortly after receiving the Oscar B. Hunter Award of the American Society for Clinical Pharmacology and Therapeutics, which is really one of the premier awards in clinical pharmacology, and I think Lewis was very pleased to get that award. I

had the pleasure of introducing him for that award.

Many of the people, of course, in this room, not just on the Committee but in the audience, knew Lewis and had an opportunity to interact with him, and I think if you had that, you really knew what a wonderful person, enthusiastic and exciting as Larry has just expressed.

But one of the things that he really did want to do and did do, I think, not only in this Committee but elsewhere was really get involved in improving the process of drug development. And one of the things I'd like to do during the next few minutes is really talk about some of those concepts that he championed and I think have become very important in the whole field of clinical pharmacology and drug development.

But I'll start out with a little bit of a background about Lewis, for those of you who don't perhaps know some of his background. He was born in New York City, and in fact, it took many years for him to evolve his California-like approach to discussions like this. Those of you who knew him

early in his career probably remember that he could be pretty acerbic as a critic of presentations and so forth, and certainly, as he grew older, he became much more of a mellow individual when it came to his discussions.

Lewis received his bachelor's degree from Cornell University, his medical degree from Albert Einstein. He was then an intern and a first-year resident at Columbia Presbyterian Hospital in New York City. He then, as many of us did in that era, go to the NIH, where he was a research associate at the National Institute of Mental Health.

There, Lewis actually published two papers in the Journal of Biological Chemistry, and I think but for a change that I'll tell you about in a moment, he might have been a molecular biologist or a molecular pharmacologist. He had planned to return to Columbia University Medical Center to finish his residency training and called down to the chair of medicine when he was about to complete his tour of duty down at the NIH and was told that he should have called earlier; that basically, they

weren't ready to take him back.

So instead of returning to Columbia, he joined the NIH Division of Computer Research and Technology, where he, I think, had his first exposure to computers in medicine and to modeling and possibly a simulation at that time, but the SAM program. This actually led to his first publication, which had to do with the computer-aided long-term anticoagulant therapy, which was published in 1969 in Computers and Biomedical Research.

After completing that additional two years at the NIH, Lewis came to Stanford, where he completed his medical residency and then went to UCSF as a clinical pharmacology fellow, joining the faculty there in 1972, and spending the rest of his career there, where he was professor of laboratory medicine and biopharmaceutical sciences.

Of course, Lewis is widely recognized as a pioneer in the field of pharmacometrics, and his career at UCSF really focused on the mathematical and statistical methods applied to the problems of

clinical pharmacology. During the early part of his career, Lewis was involved in the whole area of therapeutic drug monitoring, which was then becoming established at many hospitals through the country.

Through Ken Melman, Lewis was introduced to Bar Rosenberg, a brilliant statistician at Berkeley, and this really represented another pivotal point in Lewis' career and really marking his entrance into the field of the world of statistics. And this particular paper, again, published in 1972 in Computers in Biomedical Research, represented this first paper, actually, I think it was the second paper along with Bar Rosenberg in which the focus on individual pharmacokinetics and computer-aided drug dosing was first published.

Now, this introduction to Bar and interest in computer-aided modeling of drug therapy led to this paper, actually, two papers: a paper published in 1973 in the New England Journal of Medicine on computer-assisted digoxin therapy and

then this paper with our colleague, Carl Peck, Lewis Sheiner, Bar Rosenberg and Ken Melman again that appeared in the Annals of Internal Medicine.

This work really, I think, led, as it inevitably would, to Lewis' interest in developing methods for predicting pharmacokinetics of drugs in individuals using sparse data sets; in other words, using just a few drug concentrations obtained during the patient's hospital stay, and I think as a result of that, together with his colleague Stewart Beal, Lewis developed and applied the NONMEM program, which I think is probably most associated with Lewis' work, and I think most of you are familiar with NONMEM as a Bayesian forecasting tool incorporating population pharmacokinetic information to predict pharmacokinetics.

This novel program and novel approach has really led to greatly-enhanced predictions for dosing regimens, for patients in clinical settings allowing for individualization of drug therapy and, of course, I think NONMEM really became the

standard in the industry and at the FDA for characterizing population pharmacokinetic data acquired during clinical drug studies, and, in fact, I think really greatly expanded the entire field of population PK over the last decade or two.

Lewis then moved from forecasting of pharmacokinetics to, I think, another very important area, again, with our colleague, Don Stanski, in thinking about pharmacokinetic and pharmacodynamic modeling. Lewis had a very keen sense of clinical pharmacology, and he really pioneered these new methods to simultaneously analyze pharmacokinetic and pharmacodynamic data, leading to the concept of the effect compartment. I'm showing that basically with this slide.

This, I think, is the typical slide that one would see in many different presentations, both of Lewis and others. This has really become, I think, the way in which PK/PD data is handled by many individuals. As with NONMEM, this worked with his pharmacodynamic PK/PD modeling that has really become a standard both for industry and for the FDA

in analyzing drug response data.

Lewis' overall goal all along was to improve patient care by individualization of dosing regimens. And the work that he did really enabled this to be done in a number of different therapeutic areas. Lewis worked, as many of you know, with anesthetic and analgesic drugs, much of which was done in collaboration with Don and Don's colleagues; worked with me and many others in antiretroviral therapy and antiretroviral drugs and in many other therapeutic areas with many collaborators.

As I mentioned at the beginning, much of Lewis' work was really focused on improving the science of drug development by optimizing clinical trial designs, and his vision was to develop methods that allowed more efficient and more informative clinical trials, optimizing dosage recommendations and optimizing therapy. And one of the things which he did, again, with his colleague Nick Holford was, again, really to focus on understanding the dose-effect relationship and

along, again, with Stewart Beal and Nancy Samble of UCSF, I think this was one of the classic papers of study designs that could be used for dose ranging, particularly in phase two studies, and I've seen this particular study quoted many times at meetings and in the literature.

And Lewis would always say that this was one of his signature slides. If you didn't see this slide, you knew it wasn't Lewis talking. This was his whole concept of a response surface, with benefit-risk response surface, and he had many variants of this slide, but this, I think was one of his, as he said, signature slides and favorite slides.

Now, Lewis really, as I mentioned at the beginning, developed an intense interest in statistics. And this led him, really, to question the traditional approaches to data analysis in clinical trials and this whole concept of--did I pass one slide here?--well, I'll come to that slide. This is a little bit out of order. But in any event, he really got very interested in looking

at the whole issue of statistical approaches to analysis of clinical trials, and this review that was written just a couple of years ago in the British Journal of Clinical Pharmacology was one example; another example was this paper written by Nicholas Johnson and Lewis just a couple of years ago in Clinical Pharmacology and Therapeutics, and he had begun to work very closely with a number of statisticians, including Marie in the audience here and other statisticians at Harvard really asking questions about the analysis of clinical trials.

Now, I think perhaps his most important contribution overall was his paper published in 1997 on the concept of the learn-confirm paradigm of drug development. And I've heard this particular paper and this particular concept quoted again and again as I've talked with people in the pharmaceutical industry and so forth, and I think this really does represent a major contribution that Lewis made to the whole thinking of how one develops drugs, and I'm going to come back to that; I won't talk much about that right now, but I'm

actually going to come back to that later on this morning in my own presentation.

Lewis was obviously very interested in the whole area of drug development and in the role of pharmacokinetic and pharmacodynamic modeling in drug development and published this review in 2000 in the Annual Review of Pharmacology and Toxicology, which I think was--again, it's a highly-cited paper, one that really gives an excellent overview along with Jean-Louis Stymer, of the role of modeling in the whole drug development process.

Now, I'll mention to go on a little bit about Lewis' specific service on FDA advisory committees and committees such as this one. Since 1987, Lewis had been an expert consultant to the FDA Center for Drug Evaluation and Research and had participated in many meetings. He was, and this will become important later on again, a member of the Anti-Viral Drugs Advisory Committee from 1991 to 1994, and as you'll see in my presentation later, this was a very critical time in that field

of antiretroviral drugs.

He was very involved in the whole area of bioequivalents and was a member of an expert panel on the guidance in population PK/PD as well as this expert panel on individual and population bioequivalents at CDER. As well, he was a member, as one might expect, of the exposure response guidance panel of CDER, and finally, as Larry has already mentioned, a member of the Clinical Pharmacology Subcommittee, in fact, a founding member of the Clinical Pharmacology Subcommittee.

Lewis' substantial influence on the science of drug development has, I think, been very well apparent and documented, and those of us who knew him will remember him for his passion for this whole subject, his intellectual curiosity, as Larry has mentioned; his warmth and engaging personality. He had a great impact on the people he trained and the people he collaborated with, even those of us or those of you who had more limited interactions with him.

He really established deep and lasting

relationships with his fellows, friends and a broad spectrum of scientific and business associates. He spawned several generations of quantitatively-oriented clinical pharmacologists worldwide, not only through his research but also for his commitment to research and training, which included a number of, I think, world-renowned courses in pharmacokinetics and in NONMEM and modeling, working in many cases with his friend Malcolm Rowland and his colleague, Les Bennett, at UCSF.

This is just a list of the many people that Lewis trained. You can glance up at this list. You probably see many people that you know on this list, people who are very influential and very important in the field of drug, clinical pharmacology and drug development. This picture was taken in 1992 at a 60th birthday celebration that was held for Lewis. You see him down there in the lower left-hand part of the slide. There were probably about 100 people. Kathy was very responsible for helping organize this meeting,

Kathy and Les Bennett, and I think it really represents the kind of loyalty and so forth that Lewis was able to generate.

Lewis served as president of the American Society for Clinical Pharmacology and Therapeutics. He authored more than 200 books and chapters; was on the list of most-cited authors in the area of pharmacology through ISI; had many honors and awards, including an honorary doctorate from Uppsala University; the Hunter Award that I mentioned, the Rawls Palmer Award that I mentioned from ASCPT and an honorary fellowship from the American College of Clinical Pharmacology.

Lewis lectured widely throughout the world as well as being involved in committees such as this one, and as Larry said, he certainly will be sorely missed. And I thought these two final pictures of Lewis really represented Lewis at his best: one in Amsterdam and one in Switzerland.

Thanks.

[Applause.]

DR. VENITZ: Thank you, Dr. Blaschke.

Our first agenda item as far as the scientific agenda is concerned is a discussion of surrogate markers, and Dr. Lesko is going to introduce the topic.

DR. LESKO: Thank you, Terry, very much for the thoughtful comments, and I'm sure Lewis is looking down smiling and saying I told you so.

I'm here at this point to introduce the last topic of this meeting, which we call the transition of biomarkers to surrogate endpoints. It's somewhat of a difficult introduction to make because of the broad nature of biomarkers and because of what's gone before, namely, a large number of discussions, many of them passionate, about the topic of biomarkers and surrogate endpoints.

My colleague, Don Stanski, urged me to be visionary, and being visionary is not something that comes naturally to me, so it's difficult to be visionary. So I looked for inspiration. And I looked for inspiration to the movie that I was watching on Sunday with my grandson, Nemo, and

there's a point in the movie where these two fish, who you probably recognize, come around the corner of a coral reef and come face to face with a menacing shark, and they say something like oh, no, not him again.

And I thought about that, and I called this the biomarker fear factor, because we've talked about biomarkers endlessly for the last 10 or 12 years, and one might be apt to say oh, no, not that again.

We've talked over the years in workshops and symposia on the validation of biomarkers as surrogate endpoints, and again, this is a topic that ignites a lot of discussion and a lot of debate, very much passionate debate, with the sides taking shape.

I happened to look in the Internet, using Google as a search, and I said I wonder what's going on in biomarker workshops these days. And I was able to pull up without a lot of trouble biomarker symposia that are taking place all over the world, from France to the Netherlands to South

America, and including Baltimore this weekend, where there's a biomarker workshop that precedes the ACPS meeting.

So a lot has gone before, and I'd like to begin with definitions. These are definitions that came from the FDA/NIH 1999 workshop, and you'll probably see these occasionally throughout our morning just to set the stage as to what we're talking about in biomarkers and biological markers and surrogate endpoints, and you can see that we're talking about characteristics that are measured or evaluated as indicators of a whole variety of things, from normal disease processes to pharmacological responses to drugs. And a surrogate endpoint is a subset of biomarkers that's intended to substitute for a clinical endpoint.

The problem that we have, I believe, with biomarkers is that the pace of biomarker discovery keeps increasing at a remarkable pace, with measurable improvements In the biomarker discovery area but not necessarily measurable improvements in predicting the success of drug development. There

was an article yesterday in the New England Journal of Medicine about the genetic basis for Parkinson's Disease, and this type of discovery is so ubiquitous these days that the genetic basis of this disease or that disease is sure to spur the discovery of biological markers that are going to play a major role in drug development and in patient monitoring.

But the past focus of biomarkers and maybe even the emphasis or overemphasis has been on biomarkers as surrogates, and despite the last 14 or 15 years of debate and discussion, there have been relatively few successes of biomarkers being, quote, validated as surrogate endpoints. We've had discussions of conditions that favor or not favor surrogate status for biomarker endpoint, things like the pathophysiology characteristics. We discussed these in our exposure response guidance that came out in April of 2003, and if you go back and read that now, it is not very explicit on either how you develop a surrogate endpoint or what the criteria is to specify one as such.

There's been a subtle resistance, I think, stemming from the past failures of biomarkers as surrogate endpoints to consider their development further. And in some ways, there's been a paralysis in development of this field related to the statistical rigor that's been associated with the biomarker to surrogate pathway.

Furthermore, much of the discussion of surrogates has been fragmented into individual therapeutic areas as opposed to an integrated overview of the entire process. And finally, there's been many workshops that I think have set unreasonable expectations for biomarkers and surrogates.

But putting surrogates aside, I think we need to refocus again and enhance the integration and use of biomarkers over the entire course of drug development as a natural path to the surrogate endpoint goal.

So with biomarkers, I think a lot has happened, but it does raise the question about how things can be improved. For example, have we been

settling for less in the biomarker area? We think biomarkers are extremely relevant to efficacy and safety, aside from them being surrogates or not. We don't need surrogate markers to gain the full impact of biomarkers. Just in the past couple of months, we've had many examples of this, and only using one of those, the Iressa story. EGFR mutations and tumor tissues have been reported to predict a response in eight of nine so-called responders.

Another question is can we more fully work up biomarkers from discovery to clinical outcomes than we currently do? One of the goals of biomarker development is to begin to reduce, over the course of time from discovery to clinical outcome, the uncertainty in what I'll call that gray zone between preclinical biomarker discovery and phase three clinical outcomes. By bridging those two areas, by bridging them in a clinical pharmacology/biostatistical context, it would seem that the process would more naturally lead to acceptable surrogate endpoints, instead of thinking

of it as a one-step process of going from a biomarker to a surrogate endpoint.

You're all familiar, I believe, with the critical path. It's a call to action. The critical path calls for a collaboration between academic, industry, patient groups to work with FDA to help identify opportunities, to modernize the tools for speeding and making drug development more efficient and more successful.

The biomarker vision is expressed in that document. It talks about adopting a new biomarker or surrogate endpoint for effectiveness that can drive clinical development, and it gives an example of the well-known case of CD4 and viral load that were used as surrogate markers for anti-HIV drug approvals in the early nineties and from that point forward.

It talks about the biomarker challenge: additional biomarkers, which we can think of as quantitative measures of biological effects that really link mechanism of action, i.e., preclinical biomarkers and clinical effectiveness or outcomes,

and additional surrogate endpoints are needed to guide product development.

So the document, I think, has laid out the problem. It's laid out a vision. It's laid out a challenge. And the question that we're here to sort of begin to discuss is what do we do next. And what we do next is very important, I think. We need a new construct. We need to break the pattern of the past. I think we need to go down a different path, with two objectives in mind.

The first objective: can we achieve a general, agreeable conceptual framework to continuously reduce the uncertainty associated with biomarkers over the course of the entire drug development process: what is that systematic path? Can we define it in a general way that is not disease-specific, that is not biomarker-specific but can be applied to many therapeutic areas?

We're seeing with genomics an increase in disease progression knowledge. We're seeing that there's benefit from systematically aggregating knowledge using modeling and simulation,

quantitative methods. We've seen that there are increasing ways of establishing the predictive nature of biomarkers. We talked about some of that yesterday when we visited the markers associated with predicting irinotecan toxicity. And there's a lot of initiatives that relate to the standards for biomarker performance. So taken together, these individual initiatives, I think, bode well for a general conceptual framework.

The second goal of this initiative would be to better articulate the standards or specifications to validate and accept biomarkers for their intended use, including surrogates for registration and any extension of those surrogates for additional applications, for example, in other drug classes. So it's a twofold goal that I think we want to strive for in the context of this initiative.

Now, we're not starting from scratch with this initiative. The agency has taken steps and intends to take many steps that move us along this path, and many of these are hinted at in the

critical path. We've already implemented the end of phase 2-A meeting, and we plan to have a guidance out in 2005. We've invested in resources at the FDA and are developing a new branch of pharmacometrics to focus on quantitative methods in the IND period.

We've begun to develop drug-disease state models, disease progression models in several therapeutic areas. We've articulated, and Dr. Stanski has articulated in front of the Science Board, a very clear stepwise framework for model-based drug development. We intend to conduct an inventory of surrogate markers and look at the evidence, whether it's epidemiological, pathophysiology, therapeutic or other supporting evidence, that allowed them to become surrogate markers, so that we can learn from our current situation.

We intend to establish an FDA working group on this topic, with the goal of moving those two objectives that I mentioned forward. The working group itself will explore the development

of a potential guidance on biomarkers. And we've initiated this discussion with the Clin Pharm Subcommittee today.

The critical path document and some of the presentations today will also reflect upon an express goal to develop a new form of FDA-industry-academic collaborations for critical path opportunities, and some of these are being discussed as we meet today.

From the industry side, steps taken or to be taken, I can't really speak to that. But there are many other examples of consortia of collaborations that have been successful. And I'm going to use one of them. There's another one I could have used; it's in the current issue of Nature Reviews Oncology that talks about a vision for the development of biomarkers in oncology drug development.

But this is one that comes from industry, and it was provided to me by Chris Webster, who is associated with the PhRMA Biomarker Working Group, and it was very appealing as a model for a

consortium, and it's the Semiconductor Research Corporation. Very briefly, this is a nonprofit, precompetitive academic-industry-government consortium, which is now about 20 years old.

You'll notice some parallels between this and drug development. It was formed in the 1982 time period because of a concern about decline in the semiconductor industry. It was geared towards, as an industry, reliance on huge payoffs from individual successes and isolated research across the industry in individual companies. There was a noted reduction in R&D funding with a limited success in new semiconductor technology and a shift towards short-term R&D as opposed to an investment in long-term successes. There was a talent crisis at the time, and there were many different technology challenges.

The consortium came together, with industry, academia and government, to really lead the industry's long-term research efforts, advance problem solving technology, integrate university research capability across the country and now

internationally and serve as a hub, as a catalyst for a large global network of collaborative sites that were charged with developing technology that would enable the semiconductor industry. They developed a central vision and implemented an action plan.

It wouldn't take a lot of imagination to see the parallels to what could be possibly the case for the biomarker situation, and whether we call it a biomarker consortium or a biomarker institute, it would have at its heart the same goals that this Semiconductor Research Corporation had.

So the goals for the Committee and the strategies to move forward today: we have no yes or no questions. We have no preconceived plan as to how we're going to move forward. We have some general ideas. And what we're here today to discuss is to hear your input on the science of biomarkers, the data that would be necessary, opportunities in this field, obstacles, whether they be culture, process, impediments, and also,

any thoughts you have on collaborations. What we're looking for is your input and help to define a new path forward for biomarkers and surrogate endpoints.

You're going to hear three presentations that I think will set the stage for the discussion. Dr. Woodcock will start off and frame the issues as one of the principle authors of the critical path and one of the visionaries for this field. We'll hear from Dr. Wagner an industry perspective, and Dr. Wagner will represent the PhRMA Biomarker Working Group, and he has, again, been working with the others on a very thoughtful position paper, and we'll hear some of the principles of that today; and then, finally, we'll hear an academic perspective from Dr. Blaschke, who has lived through over a decade of the biomarker surrogate endpoint progression, starting with the AIDS epidemic back in the early nineties and reflect on that and advise us on some thoughts about moving forward.

As I say, the discussion today, we'll be

listening to very carefully. What we hope to develop is a foundation for a national critical path opportunity, which the agency will begin identifying in terms of a priority near the end of this year. We realize that this project on biomarkers is going to be a very ambitious one. We're very optimistic. And of course, like any initiative that FDA undertakes, there's always that specter of progress dependent upon its funding, its sustained commitment and dedicated staff for such a project.

So we're not overpromising anything, but we would like to begin and move forward on this path, and I'll start by introducing Dr. Woodcock.

DR. WOODCOCK: Good morning, everyone. I'm really delighted to be able to be here and begin this discussion about moving the field of biomarkers in drug development forward.

I've named my talk a framework for biomarker and surrogate endpoint use in drug development, because that's really what we're, I think, discussing here, but obviously, it has much

broader implications if we're able to move this forward. And I'm going to address those as well.

First, I'm going to cover--Larry already went over the current definitions. I think there are some self-imposed limitations in the current definitions, and therefore, I'm going to present them again and talk about them. Second, I want to talk about overall the limitations, I think, of our current conceptual and developmental framework and the reasons which are multiple why we're not moving forward more rapidly in this area, and by we, I mean the biomedical research community overall. And finally, I want to talk about what potential we have for moving towards robust use of biomarkers in drug development and then toward regulatory acceptance of surrogate endpoints, which would follow on after the robust use in drug development.

Now, in the late nineties, NIH put together a definitions working group of which I was a member and some other folks in this room were to develop some terms and definitions about biomarkers and surrogate endpoints and to have an overall

conceptual model. There had been a lot of thinking that had gone into the field about how these interact. And this was an offshoot of the consensus conference that was held on this topic, and this was published in a paper.

The definition the working group had for biomarkers was that it is a characteristic that is objectively measured and evaluated as an indicator of normal biologic process, pathogenic process or pharmacologic response to a therapeutic intervention. And I don't have any quarrel with this definition, this one.

And this is ubiquitous, I think widely used and accepted, although there might be a few modifications you could make on this, but in our--in FDA's draft pharmacogenomics guidance that we published last year, in order to set up this structure for regulatory filing or not of pharmacogenomic information, we had to go further and define the pharmacogenomic tests as either possible, probable or known valid biomarkers, because this type of definition, then, determined

whether or not there would be a required regulatory filing under the law.

And these categories were sorted based on available scientific information on the marker and how much confidence you would have the marker actually represented some real outcome or real information. And we got a lot of comments on that to the docket for this guidance, saying that we needed more specificity on these categories and to define them more clearly, and we will very soon issue the final pharmacogenomics guidance, but I don't know if it's going to shed a whole lot more light on these biomarker definitions. As Larry said, that's something we need to take up in this larger context. So those are some of the extant definitions out there of biomarkers.

Now, the group put forth a definition of clinical endpoint, all right? And that is a characteristic or variable that reflects how a patient feels, functions or survives. And this kind of is the crux of the conceptual problem I had with this whole area. Note, you should note, and

this is my editorial comment, except for survival, all these outcome measures or variables involve some kind of intermediary measurement. It's really not possible to know how someone else functions or survives; we can only measure it in some--I mean, or feels.

And I think we can all agree with that. We have some kind of measurement that we interpose between that person and the numbers, and we somehow quantify how they feel based on some kind of measurement.

Now, you can disagree about this, and we should talk about this later, because this is very important, I think. But anyway, that's a clinical endpoint. And those are given in the scheme of things some kind of fundamental reality.

Now, surrogate endpoint, in contrast, is defined as a biomarker that's intended to substitute for these clinical endpoints. And the surrogate is expected to predict clinical benefit based on various scientific, you know, studies that have been done. And there is a feeling about a

surrogate, and this is something that we need to develop more. It actually was presented by Dr. Rowland at the biomarkers meeting, but there is an issue about how proximal or distal the surrogate is to the actual clinical outcome that you're trying to describe or quantitate and say a blood measure might be quite far away or might be very close, and that might be based on mechanistic pathway proximity or it might be based on a sort of clinical face validity, so there are a number of different axes on these surrogate endpoints, and I'm going to discuss that a little bit more in a minute.

This is the definition that was put forth by the working group, and there wasn't a lot of dispute about this definition. Now, as we all know, biomarkers are used in clinical medicine. They're not simply used in drug development. And that is kind of the larger issue here. They're used in diagnosis, as a tool for staging disease, an indicator of disease status and to predict and monitor clinical response.

And I see Rick Pazdur today, who's the head of our oncology group. He knows very well, often, the clinical use gets well ahead of the drug development use. And that's because the clinical use may be based on, you know, there's less--you can simply adopt a biomarker and use it without having an organized set of data and evidence that you base that adoption on. So sometimes, biomarkers will be taken up and used in clinical medicine, at the same time not being used for their corollary use in regulation or in drug development.

But because biomarkers are critical to clinical medicine, to the diagnostic tests of the future, there's more at stake here in this discussion, in this overall initiative that we're having than just efficient drug development, and this can't be stressed enough, especially to the outside stakeholders. Biomarkers really are the foundation of evidence-based medicine, because it is those types of tests that determine who should be treated, how they should be treated, and what they should be treated with.

And so, those quantitative measurements, diagnosis should go before treatment, and yet, for many of our treatments, we have very few discriminatory markers that we apply. Absent new markers, our advances in targeting therapy, either in the traditional ways, which would be according to drug metabolism and other standard markers, or in new ways will be limited, and to the extent that we can't or don't adopt these markers and use them in drug development, treatment will remain empirical.

So it's imperative for good medicine as well as cost-effective medicine that biomarker development be accelerated along with the development of new therapeutics.

Now, here, just to get people's minds around this, many of you in the room are experts in this, but many may not. According to the NIH definition I just talked about in biomarkers, these types of measurements would be considered biomarkers of different kinds. So it isn't just a blood test. It can be all sorts of imaging

technologies or bone densitometry, all sorts of things. Even an APGAR score is a kind of biomarker. It's a way of quantifying certain observations on a newborn.

Now, as opposed to use in medicine, biomarkers are also used in drug development in a decision making capacity to try to assess and evaluate the performance of candidate treatments. Where we have very good biomarkers, we can have extremely efficient drug development, because the performance of candidate therapies can be assessed in animal models. And by the time we get into humans, we have a very good idea of the performance, a very good predictive idea of the performance of the treatment.

The biomarkers can also be used to bridge animal and human pharmacology and pharmacologic effects of therapies by doing proof of mechanism. And again, I'm stressing here the early acquisition of information about the safety and effectiveness of the therapy and bridging the animal knowledge and the human knowledge.

There are safety biomarkers, and most of those are 50 years old. I will tell you that the markers we're using in the animal safety evaluation in general and the human safety evaluation are truly venerable, and they're tried and true, okay, but they do not incorporate modern knowledge there. They're largely empirically based, and they have reasonable predictive value for major organ system failure and not very good predictive value for mechanistic understanding of the safety problem or predicting more rare types of safety outcomes in the same organ system. So there are problems with that.

But the biomarkers, to the extent we have them, can be used to evaluate human safety and early development; hopefully predict safety performance of drugs early.

And right now, we use serum chemistries. We don't use cell surface protein expression very much. That would be a target for the drug intervention. Sometimes, that's used. Drug pharmacokinetics over the last 15 years due to Dr.

Sheiner's efforts and many others, many in this room, these types of measurements have become much more standardized within drug development and have tremendously contributed to our understanding of drugs.

Serum transaminases and other safety markers have been used forever. Genomic expression profiles are used very, very rarely right now, and imaging is, in specific fields, such as neuropsychiatric disorders is being used widely, the biomarker of imaging, but its utility is still not clear, I think is fair.

In later drug development, this is where the rubber really starts hitting the road as far as cost of patient and so forth, and the stakes start really rising. If you have good biomarkers to do your dose-response work and develop optimal regimens, it's extremely helpful before getting in phase three to have a very good idea. Safety markers to determine dose-response for toxicity, we aren't as good there and determine the role, if any, on differences in metabolism on the above

dose-response, and this isn't done as widely--is that fair, Larry, to say--as it probably would be optimal to do, for a variety of reasons.

Now, here's where we start getting some probability areas for dispute or discussion. Biomarkers used in later clinical development: I would--psychometric testing or psychometric scales or whatever are used as clinical outcome measures in trials of psychiatric disorders. I would argue to you that's as much of a surrogate as an HIV viral copy number.

It's just we're used to this, so we don't think of it as a surrogate. We've used it a lot, and we're comfortable with it. But we don't know that it represents a cure or a mitigation, necessarily, in an individual patient. A lot of work has gone on, and I think we have great confidence that the testing and outcome measures that are done for psychiatric diseases actually reflect efficacy of the drug and have tremendous utility in the approval of psychiatric drugs; however, I don't think people recognize that this

is as much of a surrogate as many other types of surrogate markers that have been discussed.

Pain scale is another thing: I mean, you can't feel another person's scale of pain. We have constructed different measures, metrics, and they have been run through the psychometric testing algorithm to look at their construct validity and so forth and so on, and we know their performance pretty well. But they are surrogates for actual pain.

Imaging can be done; culture status is obviously a very important marker, not necessarily a surrogate for antimicrobials; pulmonary function tests, serum chemistries, electrocardiogram. And I think what's striking about many of these is they are very traditional. They've been used in clinical medicine a very long time.

Now, what about surrogate endpoints that substitute for the clinical outcome measure? Well, obviously, there are surrogates for efficacy that can be used to assess whether a drug has clinically significant efficacy, and there are surrogates for

safety. And basically, our entire drug development program and the exposure of patients is, in some way, a surrogate for the real world safety, because that's what we're really concerned about is how the drug will perform when it's marketed and out there in the real world as far as safety goes, so the entire development program and the patient exposure experience and the way we look at that is used to predict safety.

Right now, known surrogate endpoints and points that are used include blood pressure, interocular pressure for glaucoma, hemoglobin A1C; as I've already said, psychometric testing; tumor shrinkage for cancer, and there's criteria, performance criteria around all of these. For rheumatoid arthritis, the clinical endpoints used in trials are the American College of Rheumatology criteria that were worked through by the rheumatologists with great effort, and then, pain scales are used for pain.

Now, what I want to turn to after giving sort of an introduction is what I consider

limitations of the current conceptual and developmental framework for biomarkers and surrogate markers. And the reason I want to do this is because I think we have to start there in rethinking, as Larry said, if we're going to put a consortium together, if we're going to try and work on new biomarkers, we all have to be on the same page conceptually about what we're trying to accomplish and what are the issues.

I think most people would agree that biomarkers represent a bridge in many cases between a mechanistic understanding that has been gained in preclinical development or in actual basic science, and what is largely now the empirical clinical evaluation, and the goal is to bring the mechanistic understanding more forward into the clinical evaluation to make it more predictable, both on safety and effectiveness. And the hypothesis is we can use biomarkers to do that if we understand their performance adequately.

Now, because of history, we didn't have the science in the past, and as regulators and the

regulatory system has been focused on empirical clinical testing. And there are tremendous limitations to that, but that is the best we have had. And that has really, though, we have that historical momentum that is continuing to skew our approach to the clinical, the human evaluation of drugs to sort of an all-empirical.

And what do I mean by that? Well, you just expose them, and you see what happens. You randomize people, and then, you count whatever you count at the end of the day, and that's basically empirical drug development, and that's one of the reasons it's so expensive and timely and risky, is because there's a tremendous amount of failure in this approach, and we don't gain as much knowledge. This is not a highly informative approach, either. And of course, the FDA is constantly criticized for drugs that are on the market postmarketing that we don't have as much information as would be desirable about those drugs.

I think all of us in this room know, nevertheless, how expensive, time consuming and

what incredible effort current clinical drug development is, but this is contrasted with the fact that at the end of it, we don't know that much. And we should have a discussion about this afterwards. That's true. We really don't know that much at the end of current drug development.

And as a result of this being skewed toward a more empirical approach, the early mechanistic clinical evaluation has often been lacking. And I think Larry can speak to that, our end of phase 2-A meetings are speaking to that. There really hasn't been that focus. And this isn't to blame anyone; the reason we haven't focused on that in the past is we have not had the tools to do this, and the question is is now the time where we are developing these tools, and should we put a lot of effort into this to develop those tools, and do we have enough scientific knowledge to actually make the process a lot more predictable? And I would say the answer is yes.

But I would say as a result of the history, the business model for biomarker

development is lacking. There was just an article in Biosentry magazine about this, I think, last week, about companies who have been trying to get into the biomarker business, and they say there's really not that much interest or a model for how they can move forward and develop these biomarkers and have them used in drug development. And we've heard this; I have heard this ubiquitously over the past six months as I have been going around talking to people about the critical path.

So a consequence of this that anyone can easily observe looking at the literature is there has been no rigorous pursuit of the evidence that would be needed to qualify a marker, really assemble the evidence on its performance or to assemble that evidence at a level where you get regulatory approval of that marker. That doesn't happen that much, and there are a tremendous number of markers out there, and we know very little about their performance in a rigorous way. And the exploration of their clinical relevance is generally ad hoc; it's pursued in an academic

manner.

However, I think there's an urgent need to overcome these obstacles I have just discussed. We have new opportunities to link biomarker development to the drug development process, particularly with a newer genomic proteomic imaging and other types of markers that have been developed and with the kind of quantitative modeling that we can now do.

This requires, though, a clear regulatory framework, a signal to be sent from the regulators, I think, of what kind of technical evaluation is required. And within our pharmacogenomics effort, we're getting a lot of questions. I think that's probably one of the major questions that is sent to us, which is what kind of information has to be sent to the agency at different stages of development?

But the need also is to develop some new business models that are viable, because someone has to develop these tests: either the drug developers, device developers, someone has to

develop these tests. They can't just be an academic tool if we're going to use them in this manner.

Now, I'd like to turn to surrogate endpoints. And I gave a definition previously about surrogate endpoints, how they stand in for clinical outcomes or clinical endpoints. As most of you know, the current model for use of a surrogate endpoint is based largely on cardiovascular and HIV experiences in the 1990s and sort of the analysis that went on around those experiences.

The cardiac arrhythmia suppression trial that was performed in, I think, sometime in the 1990s was done because of widespread use of antiarrhythmic agents to suppress the ventricular premature beats post-MI based on the hypothesis that that would decrease the incidence of sudden death in that population, because they're at risk for sudden death, and the surrogate there was the suppression of VPBs.

What happened when the arms of the trial

were unblinded is the mortality was increased in the treatment arms of this trial. And that was quite a shock to folks, probably akin to what happened when they unblinded or they looked at the postmenopausal estrogen treatment a year or so ago and found that myocardial infarction was increased in the treated arms.

This caused some--the cast outcome caused a lot of skepticism, particularly in the cardiovascular community, about our ability to rely on surrogates. This is despite the fact that there was a fair amount of evidence, I think, if you're sort of impartial about this, a fair amount of evidence that certain types of antiarrhythmic agents can cause sudden death as well as certain kind of antidepressant agents and everything that have certain electrocardiographic properties and so forth.

Nevertheless, this cast outcome was a real shock. It kind of cast a pall over the adoption of surrogate area. And the whole discussion about this effort and everything can be seen in the

reference I have here by Bob Temple, who wrote up in the midnineties some of the experiences that FDA had encountered around surrogates.

Now, then, we had the HIV epidemic in the nineties, late eighties, nineties, and there was again discussion, there was discussion of the use of surrogate endpoints in this disorder; first, CD-4 counts, which were obviously not really on the mechanistic chain as much as some other endpoints, and as a result of this whole discussion, some rigorous statistical criteria for assessing the correlation of the candidate surrogate with the clinical outcome were published have a reference here they're called the Prentiss criteria, and it really called upon a surrogate to really encompass all the qualities of the clinical outcome, so you wouldn't learn any new information, basically, if you substituted the clinical outcome for the surrogate.

This is probably impossible, and no surrogate endpoint that is currently adopted has met these criteria. But again, this has caused

concern for people about what do you need to do, and is this a reasonable criterion? It is a good postulate of the problem, okay? And it frames the problem very well, and there are a lot of other articles which I could provide to people if you're interested by statisticians, discussing various performance characteristics of surrogates and the way you can be misled about surrogates.

But nevertheless, the outcome of this was that HIV RNA copy number was used as an early drug development tool. It's now used as a surrogate endpoint in trials, and it's used for clinical monitoring and antiviral therapy. There is a lack of complete correlation of this outcome measure with clinical outcomes, but my point is this does not compromise the utility of this measurement for its use in drug development or in monitoring patients. And the point is that all of our measurements are uncertain; there is some uncertainty and lack of full information associated with any measurement you might make on any person.

So there has been successful development

of antiretrovirals and control of HIV infections, despite the fact that this particular surrogate RNA copy number is not perfect and certainly misses certain parts of the outcome for any given drug.

But I want to move now to what I think is a more fundamental problem and has been a block in our discussion, and I alluded to this earlier, and as I said, people may disagree with my assessment of this, but as a clinician, I would say there is no gold standard in clinical outcome measurement. People always argue with this, and they say survival. Survival is an absolute.

And I will tell you if you look at the data, say, of John Wendburg and folks who developed that about what people would choose, would they choose longer life? Would they choose better quality of life? There are many people who would prefer to live a shorter amount of time if their longer life would--if they would have to trade off a very poor quality of life for that prolonged life.

So any measurement does not always capture

all the domains of interest for a patient; even survival. Now, I realize that's a strong statement, but obviously, if you survive sepsis or MI or something, you're left with no sequelae, you'd much rather be alive, and in those cases, that's a pretty good sequela.

But the generalizability of any single outcome measure can also be limited by the trial parameters. So we aren't really getting to full truth in a trial, even with a survival endpoint. As a rheumatologist, I'm very well aware of this because the rheumatologic diseases generally do not have a single dimension outcome, and capturing just one, capturing simply pain or function or whatever is not adequate for fully describing impact on the disease.

And therefore, many clinical outcomes and many diseases are multidimensional, and any single outcome measure we use may miss domains of interest. That doesn't mean we should throw up our hands. We should simply be aware of this, that there is no single gold standard that we're

comparing anything to.

In addition, and this is something the Prentiss criteria were talking about, because they were looking at survival, and survival can be diminished, obviously, by harm as well as prolonged by treatment effect, but in general, it's very difficult to capture both benefit and harm within a single measure. And we don't even attempt to do that within drug development. We're assembling information from a wide variety of sources, so that the concept of ultimate clinical outcome is very elusive. There's always a longer duration, say, for chronic disease. You could always follow people longer. The definition of what is ultimate is very unclear.

And so, I think we need to move away from the idea, and maybe I'm beating a dead horse here, that there's one single piece of knowledge that everything has to be correlated to. That's just really not how human beings and disease are. And knowledge about various dimensions can be acquired outside of a biomarker or surrogate measurement.

We don't have to put all our weight on a single surrogate measurement.

In addition, and this is becoming very important in this, I hope, new world of more individualization of therapy, the per patient view of outcomes is very different than population mean view of outcomes. If you are the person who experience an adverse effect from a drug, the surrogate means nothing to you, the efficacy surrogate, because something really bad happened to you.

And where we have the ability now to more individualize therapy through biomarkers, either through pharmacogenomic, genetic testing for metabolism, enzyme metabolism, where more sophisticated measures of determining who stands to benefit from a therapy or who is at high risk for an adverse event for a therapy, this becomes very important. So newer and older biomarkers do provide information at the individual level. And that's very important.

For the reasons I've just gone over, then,

I think our conceptual model should view drug development more as progressive reduction of uncertainty about the effects or, if you're the glass half full type, increasing the level of confidence about the correlation between treatment and outcomes, not a single, binary measurement of the drug is effective, it isn't effective; there are safety problems; there are not safety problems. We have to be dealing with, in other words, a multidimensional set of information, not a binary decision.

Now, I recognize that the regulatory decision has this binary quality about it. And I think what I'm telling you is that you should suppress the science into a binary box. That's not the right way to go about this. The regulators have to figure out when that evidence is enough separate from the way the evidence is developed and understood.

So no single measurement contributes all knowledge, and even if we get to the, you know, the star--what is that Star Trek, where they wave that

thing over people, and they get--they probably got a series of measurements. They weren't just doing one measurement with their magic wand. And population mean findings may not be valid for any given individual. And that's a very powerful statement, I think, as far as the fact that this anyone surrogate measure may not really predict for a given person a correct outcome.

So in the future, I think we need to move to more composite outcome measurements, more of a multidimensional understanding. And I realize--I mean, this is the Clinical Pharmacology Subcommittee; I'm preaching to the converted here. These folks have understood this for a very long time. However, we need to move this into the general understanding of drug development and therapeutics.

This means probably in general, as we move forward, we need to be looking at responder analyses and so forth and looking at the data in a more careful way rather than population mean analysis. And we also need to be moving towards

individualized therapy.

Now, we would expect, and this is kind of the quid pro quo here, with these evaluations, we also are going to have to see a larger treatment effect to provide some face validity here, if you follow me. But you would expect that if we were able to predict who is able to respond to drugs and sort out who is at risk for adverse effects. We should be seeing larger treatment effects, and in fact, we are for some of these therapies as they're moving forward.

A basic problem in a lot of drug development is the drugs don't work very well, because they are--a lot of people who are exposed don't stand to benefit from the drug and aren't going to benefit. But our empirical method of drug development causes these apparent, very small treatment effects.

Now, what should we do? What do I think we should do? And I think Larry laid out kind of the spectrum of probabilities or possibilities pretty well. What I would like to stress is

biomarkers have to be used to be accepted. We have lots of surrogate measures that we use in clinical trials and regulatory, I believe. I believe a lot of the things we use are surrogate markers. We just are so used to them, we don't think they're surrogate markers.

But what part of understanding the performance of these newer technologies is to use them, to see how they move with treatment or how they fail to move with successful intervention, to see how they perform in various populations and with a wide variety of drug interventions? With that kind of knowledge, that's the kind of robust knowledge we need, then, to have both regulatory acceptance and, then, wider acceptance in clinical medicine.

The barrier to this up to this point has been the add-on costs, and there have been many barriers, but a major barrier is the add-on cost in clinical trials. And I've talked to the imagers about this. Nobody wants to put an imaging arm in the trial if it's experimental, because it's going

to cost a lot of money, and not only could it not be used to support approval, but it might show things that are new and unknown. And there is concern that these biomarkers will, and they have, actually, segregate out the people who are most likely to respond and thus narrow the target population intended for that investigational drug. There's also concern that questions, new information would be found by these biomarkers; questions would be raised by the regulators, and that would slow the regulatory acceptance and approval of the therapy.

And, you know, we all have to get over this together, because otherwise, the use of biomarkers in trials will not occur. And that's what has to happen for us all to start understanding these.

Now, as Larry said, to bring all this about is going to require some kind of collaborative effort between government, academia and industry and probably not just the pharmaceutical industry but diagnostic side of

industry as well. And we're going to have to focus.

So I just said this: the diagnostic and imaging industry sector needs to be fully engaged in this effort. So it's going to require a lot of parties. And FDA must provide the regulatory framework and some reassurance as we move forward that individuals and firms are not going to be punished for this, so to speak. And the pharmacogenomic guidance that we published the draft last year is an example of that. It provides a space, an experimental space, where those tests can be done without the fear of all these regulatory consequences occurring and where the information can be shared.

Now, development of new biomarkers, you know, new biomarkers are going to revolutionize probably both the development and use of therapeutics and preventatives. But as I said, it requires commercial development of the biomarker technology. Academia's role, I think, is to identify these technologies, put them forward and

assist in their evaluation. But they have to be commercially developed, and we need regulatory pathways for the pair, the therapeutic intervention as well as the biomarker, and that's what we've tried to lay out for pharmacogenomics, but there are many other types of technologies that we also need to have the same pathway made available.

Now, for surrogate endpoints, I think we need further exploration and discussion of some of the ideas that I put forth today, and this is sort of the kickoff, but we're going to have to have more discussions of this. I could be dead wrong. I don't think so, but we need to talk about it. I think we need to get rid of the idea of validation, and Gerry Migliaccio is here, and we've gone through this in the last two years for the GMP initiative.

Validation is a term that, unfortunately, often conveys an idea of much more assurance and rigor than is actually attached to the activity, and we need to use more descriptive terms that everybody understands what is required or what the

activity actually is, so I think validation is a bad word to use in this context, because it doesn't convey any information.

And we may need to adopt new nomenclature overall around surrogates or perhaps refine the nomenclature. We need more emphasis on the fact that our understanding of disease and disease interventions is multidimensional. It's not a single dimension. And I think we need greater emphasis on safety biomarkers, because safety problems, obviously, are very prominent in the news. They're also a tremendous source of loss of compounds within drug development; maybe compounds that would be very good and for 99 percent of the people would actually benefit them and their disease.

So, we need to replace, I think, the idea of validation with something about degree of certainty or progressive reduction in uncertainty or some concept like that that is more graded. The problem with validation is it's, like, you're validated; you're not validated. It isn't like

that. And we have to recognize and remember that the usefulness of any surrogate will be disease, context, and to some extent, intervention-specific. And that's why one of the dimensions that needs to be investigated for any surrogate is generalizability across product classes, across patient populations, across stages of disease or what have you. That's why these have to be used in trials. We can't just have them out there in papers.

We need to develop a framework for understanding the usefulness of a surrogate as evidence, used as part of the evidence that's submitted to the FDA for approval of a drug or safety in a context-specific manner.

So in summary--it looks like I'm right on time here--there's an important public health need, I think we can all agree, but we need to get this message out, so that people understand why this is important. I don't think the general world understands what's at stake here. There's a need for the development of additional biomarkers to

target and monitor therapy.

To do this basically is going to require that they be used in clinical trials during development and postmarketing trials as well. The business model, in other words, who is going to pay for this, how this is going to happen, and the regulatory path for such markers is not clear to industry. And we need both clarification, in other words, what is the path forward, the technical, scientific path, as well as some probably stimulus is needed as well in the economic sense.

There have been definitions. Larry and I both alluded to those for these various terms. But I think further development of the model is needed to get it to a higher level of sophistication in order to increase the use and utility of markers in development and enable us all to talk to one another and know what we're talking about. I think this further development has to recognize the fact that single measurements will rarely capture all dimensions of the clinical outcome for any patient.

So I think that a multidimensional and

continuous model needs to replace the current model that we're using for clinical effect, and that's critical for the targeted therapy of the future, because this will be multifactorial as far as for any individual patient, whatever their metabolizer status or whatever it might be that the state of elaboration of various proteins, receptor proteins on their tumor cells, whatever it might be, these factors will influence their response to therapy, and many of these factors will not be binary themselves. You would not elaborate receptors on your tumor or not; it's going to be a gradation.

FDA is considering development of these concepts, as Larry said, as part of our critical path initiative, and this initiative, if we take this part up, would include a process for refining the general framework as well as individual projects on biomarker and surrogate marker endpoint development, because at the end of the day, the surrogates in particular are going to be, as I said, disease specific.

So I look forward to the discussion, and I

hope that this will lead to really something getting started in this area. Thank you.

DR. VENITZ: Thank you, Dr. Woodcock. Any quick questions or comments by the Committee members?

DR. SINGPURWALLA: Yes. I do have a comment. First is I find myself agreeing with much of what you say. Sometimes, I wonder if you're a doctor or an engineer.

[Laughter.]

DR. SINGPURWALLA: But the problems you've described are very isomorphic to the problem that engineers have found, and I'll give you two examples of what you said: one of your slides talked about validation, and you said that you shouldn't have something which is either validated or not. There's got to be some degree of uncertainty.

There is a body of knowledge called vague sets or imprecise sets where the boundary of the set is not well-defined, and you say there is a certain degree of membership in that set. It goes

under an ugly name called fuzzy sets, which the President sometimes uses.

[Laughter.]

DR. SINGPURWALLA: But I would strongly encourage you to look into that literature.

Now, as far as the markers are concerned, the problem again that you are facing is similar to what engineers face with, say, aircraft structures. The aircraft structure is degrading, and what they see is a crack. And they monitor the crack; they study the crack, and based on the growth of the crack, they predict the performance of the aircraft. So there is a large industry which looks at that. You may want to take advantage of that.

And the correct way to model these things is through stochastic processes, and these are bivariate stochastic processes, and that would be the direction in which you may want to go. One process is observed; the other process is unobserved. It's the unobserved process you're interested in, and the observed process gives you a clue. So at least I'm telling you that there is

some parallel paradigm that you may want to consider. I strongly encourage you to look into this.

Thank you.

DR. WOODCOCK: Thank you. I think what we found in our recently-completed GMP initiative is that bringing in the engineers and various other--multidisciplinary look to some of these problems we're facing provides tremendous power, because people have faced these problems in other fields.

DR. VENITZ: Any other comments or questions?

[No response.]

DR. VENITZ: Then, thank you again.

And our next speaker is going to be Dr. Wagner, and he's going to give us the industry perspective on surrogate markers.

DR. WAGNER: Great. So, thanks very much to the Committee for the invitation to and the opportunity to discuss a little bit of the industry perspective on biomarkers and surrogate endpoints.

And we've been giving quite a bit of thought to this. I represent PhRMA in this case, and in particular, the PhRMA Biomarker and Genomics Working Groups, and my colleagues Steve Williams at Pfizer and Chris Webster have been very large co-conspirators in this particular effort. And I represent, actually, a very large group that is noted at the very end of the slide.

So I want to step through a couple of different areas. I want to talk really about what our objectives and focus is right now, a little bit about biomarker nomenclature, which Dr. Woodcock and Dr. Lesko have already covered to some extent, and then talk about the idea of qualifying biomarkers as surrogate endpoints and the idea of it's not--very much along the lines of what Dr. Woodcock said, it's not really a binary process; it's actually a continuous process of increasing certainty and then end with some thoughts, some of our thoughts on collaboration.

So, there's not a laser pointer, I guess. That's okay. The landscape, I think we all agree,

is one that Dr. Woodcock already highlighted, that there's really a much more intense focus on biomarkers as aids for decision making in drug development and the regulatory evaluation of new drugs. And our objectives within the PhRMA Biomarker Working Group is really to work towards an improved framework for regulatory decision making, regulatory adoption of new biomarkers to work towards a refined nomenclature that will enhance the discussion and also to work on an optimized business model for biomarker research; again, something--these three things are really very important necessities in moving biomarker science and use in drug development along.

So our focus has been on the process, the process to select suitable biomarkers for potential regulatory purposes, to define what research is needed for qualification and regulatory use, to execute that research in a cost-effective manner and to review the results and agree on whether a particular biomarker meets the needs.

So I also would like to go back to the

FDA/NIH consensus conference in 1990--oh, thank you very much--and I won't dwell on this, but before that consensus conference, there really was, well, there was a lack of consensus. There was--biomarkers were--the term biomarkers were bandied about in a very casual way, and there was really no consensus on what folks were talking about. And the real seminal contribution to that FDA/NIH consensus conference was was this definition that Dr. Lesko and Dr. Woodcock already read--I won't repeat it--for biomarker and surrogate endpoint?

And it's really served as the groundwork for all the efforts that have come since then, because there really was a far-reaching agreement. We've done that; now, we can move on to some of the refinements that are really necessary to the next stage. And that's been part of the thinking over the last five years or since that consensus conference, and that's where we're going to go in the future.

But I think that we all agree that--or at

least Dr. Lesko and Dr. Woodcock agree that the biomarker and surrogate endpoint distinction is really not optimal for use of biomarkers in drug development, and there's a couple of guidances that really highlight that. One is, as has already been highlighted, that the exposure response guidance really makes a distinction based on the evidentiary status of biomarkers going from valid surrogates for clinical benefit to really remote from a clinical benefit endpoint.

And then, also, in the pharmacogenomic data submission draft guidance, there's really a further--that point is really drummed home even further, that there is a further distinction based on the evidentiary status of dividing biomarkers into probable valid biomarkers and known valid biomarkers, and that really leads into this idea of qualifying biomarkers in a way that makes them fit for the purpose that you intend to use them for.

I also don't like the term validation, maybe for not quite the same reasons as Dr. Woodcock, but I don't like the term clinical

validation, which is often used in the literature, because this process, I believe, has just as much to do with biology as it has to do with clinical outcomes. In the FDA/NIH consensus conference, the term evaluation was used for the process of qualifying biomarkers. That's probably okay, too, but we've settled on a term of qualification; it's really distinct from validation and captures, we believe, the idea of a graded process that leads to the right purpose for the use of the biomarker.

So we have sort of a simple working definition here, an evidentiary process that links a biomarker both with biology and with clinical endpoints. The purpose here, after all, is really to provide reliable biomarker data that's both scientific and clinically meaningful, and in the context that it's being used in.

In these remarks, my focus is very much on disease-related biomarkers that are intended as indicators in one way or another of clinical outcomes. There's, of course, a great deal of interest in all sorts of other biomarkers,

particularly pharmacodynamic biomarkers or mechanism-related biomarkers, but I think that the need for the regulatory scrutiny on those sorts of biomarkers is a little bit less than the disease-related biomarkers, because really, you know, the--how we approach the evidence to how hard a particular therapy is hitting a target is a little more clear-cut than some of the issues that relate to qualifying a disease-related biomarker. So my remarks are a bit more restricted to these disease-related biomarkers.

And then, the last point I want to make on this slide is that this fit for purpose biomarker qualification really is a graded with the accent on graded evidentiary process of linking the biomarker with biology and clinical endpoints, and it depends on the intended application. So this is the universe of biomarkers that came out of the consensus conference, biomarkers versus surrogate endpoints, and I think we can agree that it could be more useful to provide a little bit more granularity.

And one proposal that we've been exploring is to fill in this spectrum of biomarkers with graded levels of evidence, stretching from exploration through demonstration through characterization and finally through surrogacy. So, an exploration biomarker would be a biomarker which is really a research and development tool. A demonstration biomarker, then, would, in this proposal, would correspond to a probable valid biomarker, and a characterization biomarker would correspond to a known biomarker, and surrogacy has the same meaning: a surrogate endpoint, a biomarker that can substitute for a clinical endpoint.

So just to put a little bit more detail on there, it is not a lot of detail, because these really are draft concepts, but an exploration biomarker, then, again, is a research and development tool. It's not that there's no evidence. We wouldn't use a biomarker that had no evidence associated with it. There wouldn't be any sense in it. But the evidence is largely

restricted to in vitro or clinical evidence, and there really is no consistent information that links with clinical outcomes in humans.

A demonstration biomarker, then, one step up in evidence, again, corresponding to a probable valid biomarker is something with adequate preclinical sensitivity and specificity and some links to clinical outcomes but not really reproducibly demonstrated or reliably demonstrated or robustly demonstrated. A characterization biomarker, again, corresponds to a known valid biomarker, and this is one, again, that has the adequate preclinical data associated with it and is more reproducibly linked with outcomes through one or more adequately-controlled clinical studies.

And then, surrogacy is, again, has the same meaning as the NIH consensus conference, a biomarker that can substitute for a clinical endpoint. And the evidence, the details of how that biomarker becomes a surrogate endpoint are still very much a matter lacking in consensus. You know, some of the thoughts that we've talked about

are having an association and treatment effects across studies or times to events within studies; you know, there's other ways to couch the evidence that leads to surrogacy, and as I said, there's by no means any consensus there.

So just to give a little bit of a couple examples of where various biomarkers would fit in this kind of a scheme, exploration biomarkers really are only limited by the imagination and the state of the evidence that exists scientifically. There's numerous examples. A demonstration biomarker could be something like adiponectin, which is a P-par gamma agonist biomarker. Adiponectin levels increase at P-par gamma treatment, and they're associated with insulin sensitization, but the tie to insulin sensitization is far from perfect. There's also intriguing associations with cardiovascular outcomes with adiponectin, but the level of evidence is far from perfect.

So this is a biomarker that I would at least put into the demonstration bucket: do we

need it as a surrogate endpoint? I don't know; but it's a very intriguing biomarker, especially for P-par gamma agents, and in particular, because its response is very rapid as opposed to hemoglobin A1C and some of the more traditional surrogate endpoints in diabetes.

Now, a characterization biomarker that I listed here is HDL cholesterol, and there's really--there really is a great deal of clinical data associating HDL cholesterol with clinical outcomes, but there still is a lot of ambiguity about what some of those data mean. Some of those associations are still a little bit murky, and I think most folks would agree that it doesn't fit bar of a surrogate endpoint. And then, I listed LDL cholesterol as an example of surrogacy.

So we would say that there's a number of potential regulatory uses of qualified biomarkers in different categories. There's probably--you could make the argument that there may be less need for regulatory scrutiny of exploration and demonstration biomarkers, but we would contend that

there's at least some interest in focusing on how to move the biomarkers through an evidentiary scheme like this, and there's some potential roles of at least a demonstration biomarker, for example, as supporting evidence for primary clinical outcome.

A characterization biomarker, some of the regulatory uses that we would assert would include in dose finding and possibly in secondary and tertiary claims, and of course, surrogacy, as is already talked about, one of the examples of a surrogate endpoint would be in registration.

Now, there is a--this is a graded process of increasing levels of certainty, increasing levels of evidence. There's also really a life cycle for biomarkers. So not only is there a natural progression that you could imagine that goes from exploration to demonstration to characterization to surrogacy and then use in general medical use decision making; but as Dr. Woodcock pointed out, it also goes back here: things that are in general medical use. Not

everything goes through this data stream. It comes--many things are used in general medical use come back and only then become adopted into the use of in drug development.

Similarly, not all of these things work out, and we have to accept that as we study biomarkers, we're going to develop evidence that impugns their use. And I only put the arrows in this slide in these top two categories, but in fact, at any point, a biomarker can fall out of qualified use. And I think again, we have to accept that this is a risk of using biomarkers.

There's been much talk about the CAS study over the last 10 or so years in the biomarker field and about how that's really an issue, but I would submit that in drug development, we accept the risks of withdrawing drugs from the marketplace, and no one wants to have a drug withdrawal from the marketplace, but we seem to have a reluctance to accept the idea that something that we've agreed is a qualified surrogate endpoint, we're going to develop evidence that it's no longer a qualified

endpoint.

I would submit that it's a risky--the whole drug development process is a risky proposition, and we are going to develop in some cases evidence that surrogate endpoints aren't going to work out. And that is really a fact of life in biology and medicine.

The last thing I wanted to point out in terms of this line about qualification is that this really isn't the only example of a graded evidentiary process for qualifying biomarkers. A number of years ago, the NCI Early Detection Research Network had come up with this concept for phases of discovery and validation of cancer biomarkers, and they have five stages that go from preclinical exploration, where promising directions are identified, through retrospective longitudinal, where a biomarker detects a preclinical disease, and a screened positive rule can be defined all the way through cancer control, where the impact of screening and reducing the burden of disease on a population is quantified.

So this is a somewhat similar schema to the one that I presented, and I think that in general, this idea of a graded evidentiary scheme is a useful one. Of course, there was a number of issues here, and I list only some of them. There's many different schemes of biomarker nomenclature. There's many different uses of biomarkers, and I talked to some extent about that as it relates to ranging from hypothesis generation to regulatory decisions.

A particularly difficult issue with biomarkers is the different technology platforms for biomarker assays. So they range from immunologic assays to expression profiling to imaging to psychometric scales. It's very hard to talk in a uniform way about biomarkers in general when the range of the measurements is so wide. And also, as highlighted by Dr. Woodcock, there's the potential role for multiplexed biomarkers, but we really haven't gotten the scientific work done on how to put those into the right conceptual framework yet. It's really a very nascent field,

one that's rapidly developing but still very much in its infancy.

And I did talk a bit about the different strategies for qualification. And I didn't really talk very much about the assay validation side. But there's equally important issues about how the assays themselves are validated and then put into wider use.

And the last issue here is that there is an obvious need for collaboration in biomarker development. And that's what I wanted to spend the remainder of this talk on. So we would be the last to suggest that a collaboration model is the solution for all biomarkers. There's many, many uses of biomarkers that don't need any collaboration. But there are many instances: imaging is one example, where the scope of the project has become so large that a collaboration is really--it's really the only way to move it forward.

And there's many options for collaboration. I listed some of them here. The

PhRMA-FDA-NIH or other academic governmental collaboration, that's what we would really think of as the ideal new independent entity with FDA collaboration, PhRMA with FDA; without some of these other folks, PhRMA as a consortium or the status quo.

If we assume that a more wide-ranging collaboration is desirable, it really comes down to the question of how members of PhRMA can work with FDA, other governmental agencies, academics and develop qualified biomarkers in regulatory decision making. How can we do that?

Well, we believe that there are really two broad issues here. One of the issues is really deciding what biomarkers to pursue; making a development plan; executing the development plan; and maybe even at the onset, putting things into the right framework. And this is an issue, a group of issues that benefits from the widest possible cross-collaboration between groups.

The second group of issues is deciding what data would really be necessary for the

qualification of a particular biomarker or reviewing that data on a biomarker and advising regulators on its acceptance. And this is something that we view should be more independent of industry involvement.

So we would submit that one way to do this would be to have an executive consortium that would involve industry, both PhRMA and biotech, as well as diagnostics, devices, perhaps other areas; the government, in particular, the FDA, NIH, and academics.

Then, the other really important group would be a review and acceptance group, and this would primarily, in our view, fall on the shoulders of the FDA. How that would flesh out is something that could take various forms: a relevant review division for each biomarker if applicable; a new intercenter advisory group or a designated FDA advisory committee. If it were an FDA advisory committee, we really would recommend powering that committee appropriately so that the issues could really be worked on.

And then, in our proposal, form would follow function, and these separate groups would deal with each of these broad groups of issues, so that the executive consortium would deal with the group one issues, and the review and acceptance group would deal with the group two issues.

And then, going back to the executive consortium, the idea there is really not as the developer of all biomarkers; the biomarker science is a very, very large field, but to coordinate aspects of biomarker research, allowing a wide membership; ensuring that interested parties and specific biomarkers are connected and brokering syndicates, identifying gaps for qualification in biomarkers and really providing a forum, a one-stop shopping for sharing biomarker science and then acting as an expert interlocutor with regulatory agencies.

Now, we recognize that there's a large number of issues, some of them very vexing, toward adoption of a collaboration approach. There's both incentives and disincentives to industry for

collaboration. We would submit that a major incentive would be regulatory predictability and process. The funding for such an enterprise is an issue, and it could take various different forms: intellectual property in this kind of a consortium idea is an issue, as is antitrust, and governance is a particular issue. The last thing that we would want to suggest is to create a new, difficult bureaucracy that makes things harder to do rather than easier to do.

So, again, I represent a large number of people that are working both within the PhRMA context and some outside of that. And in particular, I want to acknowledge the Biomarkers Working Group within PhRMA. It's been in existence for about a year as well as the Pharmacogenomics Working Group.

DR. VENITZ: Thank you, Dr. Wagner.

Any quick questions by Committee members?

Yes, Hartmut?

DR. DERENDORF: That was a very nice overview, and I like your proposal at the end, but

I'm a little skeptical if that really will be embraced by all companies. There's a lot of biological development going on in most companies right now. And you could look at it from the other side that it may be a competitive advantage to do that, and why would companies be interested in sharing that with competitors?

DR. WAGNER: That's in part--I agree with you. That's in part why I emphasize that not all biomarkers would really be ones that you would want to put in a collaboration effort. But there are many biomarker areas that really are basically--have grown too complicated and large and expensive for any one even big PhRMA company to tackle on their own, let alone having, you know, 20 of these companies all working at cross-purposes.

The folks that have been working on these biomarker efforts within PhRMA would submit that there's at least a subset of biomarkers that we could get general agreement that a collaboration model would benefit, but I agree that it's not something that is necessarily the case for all

biomarker research and development.

DR. WOODCOCK: Yes, I would submit, although I recognize all the work that's going on, that it has not necessarily been successful in bringing about either, in particular, more predictable drug development or regulatory adoption of these biomarkers. Therefore, when we published a critical path report, quite a few firms indicated that they would be willing to share in the precompetitive area, which is very much like that semiconductor example that was given. There may be different areas of precompetitive research where only a critical mass of effort will produce the, you know, the results that are needed.

DR. SADEE: Yes, I think such a broad approach is really necessary, and for those of us who do work on looking at biomarkers from a genomics point of view or, let's say, expression profiling or proteomics, what you find is that you begin with 20,000 transcripts or proteins, and you narrow it down to a few hundred, even a few dozen. And for each application, for, let's say, cancer,

chemotherapy outcomes, you can identify maybe a dozen genes or proteins that are predictive.

And the combination of those, you evaluate the best ones; what you end up with is a panel of biomarkers that each is just maybe slightly better than the other. There is no demarcation point. Some may be totally unrelated to the disease. And so, that's also coming to you, but it's not a binary thing. It's just a complete gradation. So you get a panel of biomarkers that just declines in validity. And so, if you want to validate it, you have to have a cutoff point someplace. But you do not know which ones are going to be most predictive in most clinical situations.

And so, I think that's really the reason why this biomarker field has exploded, and there are no singular solutions, and that's why we need this type of collaboration on a very broad basis.

DR. WAGNER: I agree. And you're also very much highlighting some of the issues surrounding the multiplexing of biomarkers, where one biomarker isn't worth its salt in a particular

prediction; a group of a dozen or so can be put together in a model where the aggregate is actually pretty good.

DR. DERENDORF: In your classification, I think one very important aspect is the differentiation between first in class or fifth in class, because obviously, with the first in class with an unknown mechanism, no clinical data, it's very difficult to validate a biomarker. It's impossible, as a matter of fact. And I think that is the challenge is that you can have so many different scenarios, it's very difficult to put them in a systematic one, two, three, four classification. I think we need to keep that flexibility an creativity in this field that we can really go any way that suits the particular case.

DR. WAGNER: Yes, I agree we certainly want to stay as flexible as possible, but your point also speaks to the idea that across classes, there is the possibility of biomarkers as well. And in diabetes, hemoglobin A1C is a gold standard example of a biomarker that is a surrogate endpoint

that's accepted across different classes of therapeutic agents, and there has really been acceptance that new agents that, that new molecular entities that are being--are first in class are compared on the same standards as agents that have been in existence for years.

DR. STANSKI: Okay; thank you, Dr. Woodcock mentioned two important pieces of this problem. One of them is individualizing and improving therapy for patients; a second piece is how do you pay for it, and how do you generate economic incentives? And if a consortium could be created whereby, with the right aggregation of expertise, which included engineers to help us learn to aggregate complex information and even using Dr. Sheiner's concepts of multidimensional response surfaces, because that's really what it involves, is that this group could then both foster the development of the research and at some point be able to make clear recommendations to funding agencies of what to pay for in terms of CMS or other agencies as to when some aggregation of

biomarkers has reached a critical point that allows improved therapy as demonstrated by clinical trials and has proper statistical validity and therefore can improve treatment; therefore, we're willing to pay for it. That could create an incentive to pool the intellectual capital, because ultimately, it's the funding gate that will allow the business model for this kind of work.

DR. WAGNER: I couldn't agree with you more about that particular point. The reason why the semiconductor effort was needed and why it was successful was they worked on standards that then could drive the expansion of their business. It's very much of an analogous situation here, where if there is agreement on regulatory standards both for--within drug development and in diagnostics, that would have a real role in substantiating a business model.

DR. VENITZ: Okay; thank you, Dr. Wagner.

Our last presenter for today is going to be Dr. Blaschke, who's going to give us the academic perspective.

DR. BLASCHKE: Thanks. Well, when Larry invited me to speak this morning, he suggested that one of the things that might be helpful would be to go into a little bit more depth on the issue of the surrogate endpoints for HIV. We can learn something from past experiences, and I think that there are some important lessons to be learned.

I will say that I am a surrogate. I'm a surrogate for Lewis Sheiner this morning, and some of the slides that you're going to see, in fact, will be Lewis' slides. I think he would have had a lot of important things to contribute to this discussion.

I think this is an important concept cartoon that if you can't read, I'll read it for you. It says it may very well bring about immortality, but it will take forever to test it. And that's a real problem with a lot of the drugs that we're using now for chronic diseases, and I'll give you a little bit of an academic perspective. I'll give you my perspective on the situation. I've been working in the HIV/AIDS area for about 15

years; I've been through a lot of the things that I'll show you on the next few slides, and there are a number of people in the audience who have also been involved in this that I'll acknowledge as I go through this review.

And we've seen this slide before. This is the challenge. We need more rapid clinical development. That was certainly true in the area of HIV, and you've seen this before. This was the example that was presented in the critical path document showing that the adoption of CD-4 cell counts and measures of viral load really led to a speedup in the approval of antiretroviral drugs, and this did result as a cooperative effort involving the FDA, a number of stakeholders, academic and industry, as I'll show you as I go on.

So what I want to spend the first part of this talk discussing is now surrogate endpoints were used for approval of antiretroviral drugs for HIV infection. And it's important to go through a little bit of the history of this, because it's not as simple as it would like to be. The first

approval, in fact, based on a surrogate marker occurred in 1992, with a drug called DDC, a nucleoside analogue, zalcitabine, from Hockman-LaRoche, and I've highlighted a couple of the features of a press release that came out at the time of that approval, which was on June 19, 1992; DDC was approved.

As noted in this release, it was the first drug approved since the FDA had announced its accelerated approval process, and as noted in red on the slide here, the process incorporates the use of surrogate endpoints to determine efficacy, and as you'll see later on, the process allowed for approval to be withdrawn if further review determines the therapy was to be ineffective, and John mentioned that point in his presentation.

So 1992 was really the first time that the HIV RNA and CD-4 cell count was used as a surrogate for approval of DDC. And what were the factors that really accelerated the acceptance of it? At this point, it was just the CD-4 cell count for approval of DDC. Well, obviously, it was the

urgent need for new therapied for this fatal illness, and one of the things in the position paper that PhRMA has generated is the environment here was risk-tolerant. We really didn't have alternative therapies for HIV. We knew it was an illness that was a fatal illness, and there was an urgent need for developing therapies.

There were strong patient advocacy groups, and most of us lived through that experience back in the early 1990s, late 1980s of these advocacy groups that were really pushing very hard for the development and the approval of new therapies. It led to Congressional interest in this, and importantly, it led to some changes in FDA regulations that allowed surrogate-based approval when a clinical endpoint was perhaps not what we were looking for.

I think very importantly, it also represented a willingness of the FDA to take risks by requiring a phase four commitment, and I would point out that Carl Peck, who I think is probably still in the audience, who was head of CDER at the

time, was also the acting head of the Division of Antiretroviral Drugs, and Carl was very forceful in promoting the approval of drugs based on surrogate endpoints, and you'll see a paper that I'll allude to in just a moment that I think represented a very important effort on the part of the Food and Drug Administration to look at surrogate endpoints.

And as I mentioned earlier, it really represented a collaboration among clinical scientists and statisticians from academia, industry, and the government, and it wasn't all that well-organized, as I'll try to show you. It happened, but it didn't happen in a terribly organized fashion, but it was a very important point in making this actually happen.

Now, this was the paper that I was alluding to by Stella Machado, Mitchell Gail and Susan Ellenberg. As you'll see from the affiliations, this is really a collaboration between the NCI as well as the FDA. Stella was somebody that Carl had really asked to lead this issue of using laboratory markers as surrogates for

those clinical endpoints in the evaluation of treatment of HIV infection. You'll see this was published in 1990, and as I said, the first approval based on these surrogate endpoints occurred in 1992. This was a very important effort and a very active, very busy effort to look at this whole question.

The next ARV class that was approved were the protease inhibitors, and they were approved in the mid-1990s, 1995. Saquinavir was first, followed shortly thereafter by ritonavir and indinavir about six months later, four to six months later. And this is an important, again, press release that occurred at the time of the approval of saquinavir that was provided by David Kessler, who said that the review of saquinavir is the fastest approval of any AIDS drug so far and demonstrates the FDA's flexibility in situations when saving time can mean saving lives. When it comes to AIDS and other life-threatening diseases, we have learned to take greater risks in exchange for greater potential health benefits. And I think

again, that's a very important concept that we have to remember, especially in something like HIV.

Carl has talked about this subsequent to that in presentations that he has made, and I think it's important to highlight what this meant for the development of these protease inhibitors that I just mentioned; that for saquinavir and indinavir and nelfinavir, you can see from the top line there that the development of these compounds really was very, very short compared to the usual development times: five, three and less than three years in clinical development; a relatively small number of clinical trials that were required prior to the submission of the NDA; relatively small numbers of patients in those trials, about 1,000 patients in each of the NDAs, and accelerated approval, as I mentioned before, that was based on a surrogate endpoint and a requirement for postapproval clinical confirmation. So it really did make a difference.

The result of using these surrogates for the antiretroviral drugs meant the rapid approval

of new drugs to treat HIV. We now have over 20 antiretroviral drugs on the market; most of them really have been proved in record time, both the pre-NDA time frame as well as the, obviously, the review time for these compounds has also been quite rapid and quite short.

It's provided, I think, incentives for companies to develop new drugs for HIV, because the pathway to approval is really fairly straightforward. It's now been embodied in an FDA guidance for antiretroviral drugs. And I would also say that, in fact, because these drugs are so efficacious in the treatment of HIV, approval now without the use of surrogates would, in fact, neither be feasible nor ethical. It would take years and tens of thousands of patients in order to demonstrate efficacy using clinical endpoints for HIV infection, so this has really been a remarkable achievement in terms of the development of surrogate markers.

But let's go back a little bit and look at the process that actually occurred in qualifying

the use of these two surrogates, that is, the HIV RNA, plasma, CD-4 cells and surrogates, because it really didn't occur in, as I say, in a nice, simple fashion.

Let me go back and talk about some general principles, and then, we'll illustrate how those principles were, in fact, applied in the use of the surrogate endpoints for HIV. First, is that a surrogate endpoints qualification has to begin with a hypothesis about the pathogenesis of the disease. It ends with the establishment of its applicability by using clinical trials, and what happens in the middle? The important thing is that we have to have basic and clinical studies of pathogenesis. We have to have markers that are discovered about disease progression. We have to collect data from both preclinical and early clinical studies. I assert that we need to develop mechanistic and semimechanistic models and avoid the use of only empirical models and, again, collaboration and sharing of information in order to qualify those biomarkers as surrogate endpoints is certainly what

occurred.

And I'll go through these components pretty quickly, because I think they're fairly well-known to everybody. We know that HIV is caused by an infectious agent. That needed to be discovered. It was discovered and was, I think, well-documented to be proven as the causative agent of AIDS, and of course, what we really needed to show was that suppression and prevention of HIV replication would really alter the course of the disease.

A lot of work was put into pathogenesis of HIV. We learned an enormous amount in a very short period of time about the nature of HIV replication and its interaction between HIV and the immune system. These were extensively studied in vitro, in animal models, and in vivo. This was largely an academic endeavor carried out within the NIH and at a number of different academic centers; really, a tremendous effort that occurred in order to make this happen, and it led to a detailed understanding of viral structure, replication mechanisms,

interaction of the virus with the CD-4 cells, involvement of co-receptors and so forth, and this was all extremely important in the development of therapies for HIV, and it was largely carried out that--development of antiretroviral drugs was largely carried out, as one would expect, within the pharmaceutical industry, although in this case, there was significant collaboration that occurred with the NIH and with academia, and I would note the role of the NCI in the development of zidovudine and in protease inhibitor development. So this really was a very collaborative effort in terms of pathogenesis as well as in drug discovery.

And then, we had the discovery of these biomarkers that I will call the biomarkers of disease progression, and these occurred, really, because of the efforts of multiple groups, again, mostly from the academic side who evaluated many possible biomarkers of the progression of HIV to AIDS. Along the way, there were a number of putative biomarkers that were evaluated. P24 antigen was one of the first; then came CD4 cell

counts and a number of other measures that were looked at very carefully to look at disease progression, and this occurred, really, because of the availability and the support of a number of cohort studies, and I've just listed half a dozen or so here.

There were many others, both large and small, that contributed enormously to the information on biomarkers and on disease progression, and that required these important steps that John also alluded to, which was the validation of biomarker assays such as the CD4 cell count, the HIV RNA assays, and then, the next important step which occurred essentially in parallel with many of these was the collection of that biomarker data from interventional clinical trials, and Janet alluded to that as well.

And then, subsequent to that was the creation of mechanistic or semimechanistic models, which incorporated those biomarkers to see what interventions might do to those biomarkers and ultimately then to the qualification of those

biomarkers as surrogate endpoints. And this was one of the very important studies that occurred relatively early on in terms of trying to understand mechanistic models for HIV infection, a study that was done by David Ho and Alan Perelson, published in Nature in 1995, looking at the rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection.

This was done in collaboration with Abbott Pharmaceuticals; John Leonard at Abbott Pharmaceuticals, and what these investigators were able to demonstrate was sort of this multicompartmental location of HIV replication, a very important observation, a very important finding in terms of understanding viral replication, and that, then, because this was an interventional study as well, then helped understand the role of antiretroviral drugs in the treatment of HIV infection.

But as I said, this really didn't occur in a nice, linear process. So you start looking at some of those dates that I've shown you; we

approved, or we, the FDA, approved the first drug in 1992, but in fact, a lot of this work with biomarker development and the evolution and qualification of those biomarkers into surrogate markers ultimately or surrogate endpoints, John, ultimately leading to a guidance on this approval of antiretroviral drugs really occurred in much, much later than that first approval in 1992. So just recognize that when you have a disease like HIV, where there's a lot of pressure to get things done, things will happen, and they often happen in a--as I say, a nonlinear fashion.

And I'm using this to, just, again, recognize that here in 1997, we have a nice review of the approach to the validation of markers for the use of HIV RNA in clinical trials that was done, again, a collaboration between academic, FDA and the NIH, and even more recently published in 2000, we have a surrogate marker collaborative group talking about a meta-analysis of the use of RNA and CD4s prognostic markers and surrogate endpoints in AIDS.

So there's still a lot of active work going on in this field to try to really understand, again, from a mechanistic point of view and a pathogenesis point of view how these markers can be used to help us better understand the therapies of HIV and, in fact, approval of drugs.

And I show this one slide not to--because you can read it but because I really want you to see what a large group of people were involved, for example, in this HIV surrogate marker collaborative group that published that paper that I just showed on the previous screen. So listed up here are actually 55 people as part of that collaborative group as both international representation from both industry and academia.

So these kinds of things really do require a lot of input, a lot of data, and a lot of the people involved in this were heavily involved in generating the data that's been used to develop these biomarkers and surrogates in HIV.

So, now, I'm going to turn around and put my Sheiner hat on, and I'm going to talk a little

bit from an academic perspective about the general principles of biomarker use and qualification. And this was, again, one of the slides that perhaps Lewis showed at one of these earlier meetings; I'm not sure, but basically, the principles here is that to establish causality, given an empirical association, by supporting pharmacological activity as a mechanism, not by ruling out other causes.

And so, the evidence that would support a pharmacologic action is that the response correlates with temporally-varying exposure; that causal path biomarkers change in a mechanistically compatible direction, rate and temporal sequence, and we saw that when we looked at viral RNA and CD-4 in the HIV area. And as Lewis pointed out, learning trials and analyses are well-suited to mechanistic interpretation of time-varying data, and independent causal evidence is still required. Causal evidence from the same randomized controlled trial doesn't rule out some sort of transience or interaction. So again, the key point that he was making there is that causal path biomarkers need to

change temporally in a mechanistically compatible direction, rate and sequence.

So what are causal path biomarkers? Well, that's illustrated on this cartoon here, and we begin with the pathology that influences the physiology and ultimately the disease progression. What is next incorporated into this concept is the idea that we have an intervention, and here, an area that both Lewis and I were interested in was not just to incorporate the drug but in fact incorporate drug exposure, which represented both pharmacokinetics as well as patient adherence in order to get better information, so the model that was used for the intervention represented, again, both individual differences in pharmacokinetics as well as patient adherence.

The pharmacokinetics, of course, lead to time bearing plasma concentrations, and then, what we're looking for are biomarkers that change as a result of the changes in exposure to the drug. And of course, what is important in terms of really then understanding whether a biomarker is, in fact,

something that we really want to continue to pursue in more detail is to look to determine whether we see the correct temporal sequence, which gives us some confidence that there is a mechanistic involvement of this biomarker in the physiology and ultimately in the clinical fact and in the disease itself.

So let me just go back and talk a moment about causal path biomarkers as opposed to biomarkers in general. So causal path biomarkers are those that serve as indicators of the state or activity of the mechanisms that connect the disease to the clinical manifestations. They have to be scientifically plausible based on our current understanding of the disease itself, and that was certainly true with HIV/AIDS.

As knowledge increases, the confidence in the validity of the biomarker will increase, especially when drugs in the same class or with the same indication affect the same biomarker, and I think this is an important principle. If we have a biomarker, if we have a disease, and we have a

biomarker that's influenced by drugs of different structures and different class, it really increases our confidence that this particular biomarker represents a causal path biomarker, one that's important in the disease and in the disease progress itself.

More biomarkers will be useful in developing models of drug action, and again, causal path biomarkers need not be surrogate markers when they're used for drug development decisions or as confirmatory evidence of efficacy. And I won't get off on that tangent for awhile; as you know, Lewis and Carl Peck have been very interested in the concept of using causal biomarkers as confirmatory evidence along with fewer clinical trials.

So the credibility of these causal path biomarkers does depend on the state of scientific knowledge of the disease mechanisms, consistency of the association with a clinically-approvable endpoint and the biomarker; proximity of the causal path of the clinical endpoint. Obviously, the closer that biomarker is to the endpoint of the

disease, that gives us more confidence in that biomarker and then multiple biomarkers changing in the correct temporal sequence, and again, this alludes to the concept of having perhaps multiple markers that may be important rather than just a single marker, and again, similarity of the biomarker exposure and the clinical exposure response when both are studied together, and all that came, as you saw, at the bottom of that slide from a workshop that was held by CDDS a couple of years ago, involving Carl and Lewis and Don Rubin as well.

And this next couple of slides and tables just was something that appeared in a paper that was published from that conference by Carl, Don Rubin and Lewis in Clinical Pharmacology and Therapeutics about a year ago, just a table of causal path biomarkers. Just highlight a few here that are really already either biomarkers or becoming close to being surrogate endpoints and a few others on this second part of the table of, again, biomarkers that might well be those that

could be qualified as surrogate endpoints.

So again, establishing pharmacological causality is really what we're trying to do here, and what it basically means is that if we start with an empirical association that we get from preclinical or clinical studies, we establish causality by directly supporting pharmacologic activity as the mechanism and not by ruling out other causes. It's more demanding, in fact, than empirical confirmation, and the evidence is this establishing the credibility of those causal path biomarkers.

Now, this is, again, a slide from Lewis that demonstrates that one can, in fact, gather information about biomarkers and causal biomarkers during phase two and phase three trials; in particular, of course, Lewis, as I mentioned earlier this morning, emphasized the learning elements of the phase three trials that can be carried out by looking, for example, as you see from the slide here, at those surrogate prognostic covariates, serial biomarkers. PK compliance,

again, is emphasized here and then the use of model-based analysis as part of the process of analyzing not only phase two trials but also phase three trials.

One of the important things which I think Lewis contributed was his concept of learning while confirming, and I think again, this is a concept which I hope we will see more of in the whole drug development process. The point that he wanted to make here was that when we look at confirmatory trials, which we usually think of as phase three trials, we're talking about random assignment, placebo controls, clinical endpoints, baseline covariates, homogeneous patients and so forth, and that's a typical outline of a design for a phase three trial.

However, if we add some additional measurements, pharmacokinetic measurements in phase three, compliance in phase three, but importantly for the purposes of this discussion, serial biomarkers or other covariates that we can look at, we may increase somewhat the work involved and the

number of patients involved, but in fact, what we gain is considerable. And then, if we add to that heterogeneous patients that we begin to look at, this individual patient therapy is, as Janet also mentioned, we begin to have some mechanism for looking at responders and non-responders rather than looking at a more homogeneous group.

And then, specifically, an area that Lewis and I have both been interested in is the use of multiple different doses and potentially even individual dose escalation trials to try to really understand the dose-response relationship. And the point, again, to be made from this slide is that we can do this in the context of a phase three trial. It may produce some increase in the effort involved in the trial; it may increase some of the time involved in carrying out those trials, but the kind of information that we gain from this sort of approach can really be quite valuable.

So the other point that I think needs to be made is the issue of when is a surrogate ready. And I've sort of alluded to that already in terms

of the HIV problem, but I think we're all comfortable with the idea that the empirical certainty is not highly necessary for drug development decisions.

In fact, we want pharmacologic activity, and we want mechanistic activity for those drug development decisions and for labeling, but I think the most important one that I want to focus on here is that when we have great potential benefit along with a high prior presumption of a positive risk-benefit ration and the excessive cost of objective evidence, those are really the kinds of areas in which we really need to go ahead and look at the use of alternatives to clinical outcomes in terms of evidence for approval.

And again, what Lewis talks about is that confirmatory really should also include learning. And this goes back even to an APS meeting back in 1998, in which Lewis described the sort of situation in which empiricism needed to be balanced with the use of causal models, drug regulation demand, certainty and information; causal models

are inevitably uncertain but highly informative, so when do we use this sort of model, and when do we use, in fact, surrogate markers at an early stage, when lesser certainty is permissible, as in labeling of the drug so that we can use modeling and simulation and so forth to improve our knowledge about labeling, but importantly, about safety and efficacy when there's great potential for benefit or high prior presumption, and basically, again, a plug for the use of modeling, that modeling certainly can yield high certainty when we have credible models and the correct performance of some of these tests under the null hypothesis, and that sort of gets into this other area that I mentioned earlier this morning about use of alternative statistical tests when one is analyzing trial data.

So, again, just from an academic perspective, what do I see as some of the next steps that we need to take? This is actually a slide that I took from Janet's presentation a couple of weeks ago at the ACCP meeting and what

she said at that meeting about what we need in biomarker development, data pooling, synthesis, analysis, identification of what's known and not known and gap analysis. We heard John talk about that, identifying what studies are needed to fill those gaps and then doing the work and not just standing on our heels.

And as a final comment, I think that basically, the public wants more therapies at reasonable prices. I think we've heard that over and over again, and the high cost of drug development is something that I think all of us believe could be improved by a number of approaches that are part of the critical path document, including the implementation of better surrogate marker data or surrogate endpoint data.

I don't think the regulatory issues are necessarily any longer a major impediment. I think the regulations are in place to approve drugs on a surrogate endpoint basis, so we don't need to have a lot of new legislation in order to make this happen.

I think what we're hearing this morning and what we're hearing in general is that the FDA is very willing to move forward with new surrogates, that we don't need to think that there's a resistance on the part of the FDA to do this.

Substantial collaboration among academia, industry, and regulatory bodies will be necessary, and I think John spoke to that very nicely. All I'd say about academia is that unlike the FDA and unlike the industry, we are not organized.

[Laughter.]

DR. BLASCHKE: And when I talk about academia, who knows what I mean?

[Laughter.]

DR. BLASCHKE: There are a lot of us out there. But I think that there are mechanisms for getting people to come together for this kind of important activity.

And I think what I've tried to illustrate is that this past history with antiretroviral drugs for HIV indicates that such collaboration can occur

and that it benefits all of the constituencies. And we've already heard that there are already a number of meaningful collaborations underway and that we really need to encourage and support these.

So I'll just finish with this: I think the goal that we all have is not just another proprietary bestseller but really to get through some major breakthroughs, and I think that this kind of approach that we're hearing about this morning can help along that path. And I'll stop here, and I think we'll be ready to open it up.

Thanks.

DR. VENITZ: Thank you, Dr. Blaschke.

Any quick questions before we take a break and start the--

DR. SINGPURWALLA: I have a comment.

DR. VENITZ: Go ahead.

DR. SINGPURWALLA: I enjoyed your mentioning of causality, but I wanted to draw your attention to the fact that there is a body of knowledge called probabilistic causality which your colleague at Stanford, Supes, specializes in. And

there are different interpretations. There is something called prima facie cause; genuine cause; and a spurious cause.

I'm wondering--and a lot of information on causality is rarely discussed in the literature, the philosophic literature. And I'm wondering if the drug community is looking at that particular angle, and if it's not, I'm recommending it.

DR. BLASCHKE: Well, I'd go back to the comment that Janet made to your earlier comment, and that is I think that bringing together people with different expertise and so forth really does add to the value, and if there's a reason for collaboration, it's just exactly that kind of reason, that we can't all know everything, and there are plenty of experts out there in various disciplines that I think we need to bring to bear on these questions.

And I don't know them all, and I think that's the kind of input that we need to have.

DR. DERENDORF: Very nice presentation. I agree with everything you said. I'd like to come

back to this definition or desire of a causal path biomarker. Clearly, that's the most desirable situation. But I don't think it should be a prerequisite for biomarkers. There are many examples where there is no causal or no apparent causal relationship. Think about developing of benzodiazepenes based on EEG as a surrogate or fentanyl derivatives, as Don has done.

So it doesn't necessarily have to be a causal path, and it can still be operative.

DR. BLASCHKE: Well, I think we start with empiricism. And what the academics can often contribute to this is to move that in the direction of understanding the mechanism or the scientific basis for the change, whatever it is, whether it's a change in receptor, et cetera. I certainly don't think it's a prerequisite, but it's something that I think we do strive for is to really understand how something works and why it works and the way it works.

DR. DERENDORF: I think it has to be reproducible and predictive. I think--

DR. BLASCHKE: Ultimately, absolutely.

DR. SINGPURWALLA: I think your point is very well taken, and that's why I'm drawing attention to Supes' book on causality, where he does cite spurious cause as an empirically observed phenomenon which may not be the real cause, but that's the best you can do. So again--

DR. BLASCHKE: Point taken. I agree.

DR. VENITZ: Okay; then, let's take our break. We'll reconvene at 11:00 and start a general discussion of the topic.

[Recess.]

DR. VENITZ: Okay; before we start the Committee discussion, I would like to ask Dr. Lesko to kind of give us our charge, what kind of feedback you would like to get by the Committee.

DR. LESKO: Okay; thank you, and I'll try my best to lay out some structure for the discussion.

A couple of--I mean, we've heard some very interesting presentations this morning that I think lay the groundwork and help us tee up what amounts

to a new initiative in the world of biomarkers and surrogate endpoints. Some of the thoughts I had with regard to the Committee discussion would be knowing what you know from the presentations, what are your thoughts on what FDA can do to assure that we gain some momentum behind this project and move it forward?

Let me continue with a few others that we can keep on the table: what does the Committee think industry can do to facilitate the proposal that we've tried to lay out collectively here this morning? And finally, what can academia do?

Another issue would be what didn't you hear today in the area of the biomarkers? What was missing from the presentations that may be on your mind with regard to advancing this field in the way that we've talked about?

Dr. Blaschke in his presentation mentioned, in a sense, a means to an end, but the means to the end was not a linear process in the area of AIDS. It was a process that at the end worked out. But the question would be, and maybe

some discussion can occur around this, is that the way it's going to be? Is that the way it has to be? Or can there be a more systematic way, if we were to think of the problem of the AIDS again and then think about how that could be moved forward? Is it possible in the current environment to do that in a systematic way?

We didn't talk about this too much in the presentations, but there was the list of biomarkers that was in one of the slide sets that came from the CDDS workshop on biomarkers, and there were many biomarkers there listed side-by-side with clinical outcomes. And one of the thought I had is does the Committee have any specific ideas on what we would now call biomarkers that would be in close proximity either in a causal way or even in an empirical way to a clinical outcome, and what could be done to close the gap between the biomarker and the surrogate endpoint in terms of predicting clinical outcome?

A couple of examples of what I mean: one example would be bone mineral density; that is, a

causal path biomarker for fractures and reduction in fracture rate. Bone mineral density is used as an approvable endpoint for a claim of prevention of osteoarthritis, but it is not used as an endpoint for an indication of fracture rate reduction. So there's a gap there. What kind of data would be needed to move biomarkers in specific therapeutic areas to further along towards the surrogate area, and how could those sort of gaps be identified in terms of what we know and how we might get the additional data?

A couple other examples: gastric acid, a causal state biomarker; can it be advanced with additional data, data mining, new research to become a surrogate endpoint for additional clinical approvals. Third example, just to stimulate some thinking, H pylori eradication and its usefulness in terms of duodenal ulcer recurrence and things of that sort.

So anyway, I'll just pause here. I think there's a couple of things on the table that maybe we can get some discussion going, and there is no

boundaries on the discussion. There's a lot of possibilities, but I just wanted to throw out a few things for the group to think about and to kick around.

DR. VENITZ: Okay; any comments by the group?

Jeff?

DR. BARRETT: Larry, I wanted to address, you know, the point about the systematic approach relative to maybe the convoluted path. One of the things that struck me, and we talked about this briefly, was a lot of the emphasis is focused on the early stage discovery processes involving biomarker identification and evolution through the development process, but it strikes me that another area of focus could be from the back end as far as working with thought weeders relative to the basis for an approval.

I think we seldom are in areas where it's completely unknown what is going to constitute the basis for an approval. So from the standpoint of looking at those study designs, criteria both

statistical and clinical that constitute the basis for an approval, what would those decision makers at that stage like to see at the earlier stage to show some level of association between a marker to be named and that basis for an approval.

So, you know, perhaps there could be a meeting in the middle of the biomarkers that get advanced at early stages relative to what is ultimately going to potentially be a surrogate marker. So that was one thing that struck me. And the other thing that I thought was an interesting point was acceptance criteria on making generalizations. We talk about empiricism a lot as perhaps being a dirty word here, but I think the exploratory nature of the biomarkers has to be there at the early stages, and it's very rare that a company will invest in studying a biomarker without some justification or rationale, so I simply feel that for the most part, that is in place, but there has to be some criteria by which we make those generalizations, when, it's okay, when it's not.

So that kind of acceptance criteria on generalizations will help you, I think, differentiate compound-specific mechanism-related biomarkers versus things that may be associated with a class.

And then, I think the other point I wanted to make was just to be able to differentiate between the measurement detection issues relative to the response measurement issues associated with observational and exploratory versus a confirmatory test. Those pieces, I think, really need to be compartmentalized and focused on if we're going to move forward.

DR. VENITZ: Comment that I had in my mind the crux, as far as it relates to coming up with surrogate markers is this mix of using empiric evidence and mechanistic evidence in the right mix to convince ourselves that we have either lots of empirical evidence on the Prentiss criteria, which means it's going to be very difficult to actually do that short of doing clinical outcome studies; at the same time, what is the level of evidence that

you need mechanistically to convince ourselves that those biomarkers are related to the causal pathophysiology in the disease?

So I think one of the things to focus on, in my mind, at least, would be what evidence, what burden of evidence do we put on mechanistic information? Just like we classify right now in clinical treatment, therapeutic treatments, the evidence to support individual treatments? Let's come up with criteria to assess what mechanistic evidence do we need to argue that a biomarker is more likely than not related to the causal path? I don't think we have had that discussion, and it may be a matter of just going through a couple of examples.

We had a similar discussion last year when we talked about the pediatric decision tree, where one of the key questions is is the disease similar? Well, what evidence do you need to support the contention that the disease is similar in pediatric and in adults?

And you're getting back to the same issue:

short of doing empiric studies, which means it's very expensive and very long-term doing it, what mechanistic studies, at what level, in vitro, in vivo, animals, what have you, do you need to support that hypothesis? So I think we really need to think about how we evaluate mechanistic evidence to support transition from biomarkers to surrogate markers no matter what the ultimate qualification would be like.

DR. STANSKI: Yes, I think that's a very good point. Obviously, at some level, this is going to be marker and intervention specific. However, we could, I think, much more exploration of the general principles on the mechanistic side could be done to provide a general framework, and I think that's what we were talking about earlier, that perhaps we can engage in a discussion about the general framework for doing this; maybe using examples is a good idea. What do you actually mean? And what level of evidence is acceptable that something is on the causal chain?

There are so many variables that probably

even elucidating those variables would be helpful. I was talking to Rick Pazdur at the break, and we talked about, you know, for the serious and life-threatening illnesses, because we have the accelerated approval mechanism that was spoken about earlier, then, the tolerable degree of uncertainty is greater. You accept greater uncertainty, because you can pull the drug back, and you're expecting those confirmatory studies.

I think depending on your priors, the priors that you have are extremely important in this analysis. And, you know, whatever we did or did not know about HIV, we were pretty sure it was an infectious disease, and we have a very good model about eradication or, you know, suppression or microbes or viruses and the relationship to disease progression in many infectious diseases. And so, we had very strong priors about that doing that would be successful in helping control HIV disease.

And that's very different in each kind of disease area we're talking about. But a general

discussion of that would be helpful.

Now, getting to the other end, which was just raised by the previous comment, on the acceptance end, the regulatory acceptance end, I think we also need to write specific guidance, because a surrogate doesn't stand alone. It has to be embedded within a trial design. There have to be quantitative limits on what success means as far as the duration of the trial, the kind of observations, the analytic validation that has to go on for the particular measurement and so forth and so on. So there are a lot of specific, condition-specific things that could be talked about at a disease-specific area as well.

DR. VENITZ: Wolfgang?

DR. SADEE: I think that maybe a compilation of a few examples would be useful in where it's becoming very clear what we need to do and others that are not so clear. And so, one example would be the growth factor receptors and tarsin kinases that are increasingly targets for cancer chemotherapy.

And so, you already have--you know about the mechanism, the expression or the mutations in these target genes are important. In many cases, y you can inhibit these target genes, and nothing is happening. And so, it becomes exceedingly important to define the criteria by which we go forward, and that's a whole class of compounds that comes to the fore, and I think that would be a very useful mechanism to set up a rational approach from the beginning, because we are only looking at the tip of the icebergs in terms of the types of compounds coming along the line and which ones will be useful, and with EGFR inhibitors, only 15 percent responds, and that's correlated to certain mutations.

But maybe not always. And so, that's one class that requires a clear set of guidelines that one can use in order to take maximal advantage of this over the next five years.

DR. VENITZ: Another comment relates to the fact that you are advocating to find more safety markers, which I think we all would agree

with, but a lot of safety issues are not necessarily related to the primary mechanism of action of the drug. So I think most of our discussion so far has really focused around the mechanism of the drug and the pathophysiology of the disease, which may or may not be related to any safety issues.

So I think there should be a separate initiative, if you like, to look at potential safety markers for hepatotoxicity, and things that are very difficult to, at this stage at least, to predict. So maybe we can get away from the true and tried serum transaminases. So safety markers to me is a different domain to look at, because it does not relate to the mechanism of action of the drug. It may or may not relate to the pathophysiology of the disease.

DR. STANSKI: Yes, we agree with that, and in fact, safety biomarkers, safety markers in general have had a different evidentiary threshold completely than what we're talking about for evidence of clinical benefit. So it is really a

different game entirely and probably can be pursued separately but probably is equally important.

DR. WATKINS: Just to expand on that, you could imagine a treatment for osteoporosis that you could show was effective in 20 people with the right genotype, with the right surrogate marker. But until the issue of safety and particularly idiosyncratic reactions is solved, even if the FDA were willing to allow that to go to some postmarketing surveillance, you know, aftermarketing, the medical-legal environment in the United States, I think, would be a powerful argument for the company to go ahead and study thousands of people for a long period of time anyway.

So all the advantage of the efficacy surrogate markers would be lost until there is some kind of an understanding or progress made in safety biomarkers.

DR. MCLEOD: Sticking on the theme of safety, safety does represent an area that all three of the stakeholders that were mentioned have

commonality. And it's probably the only area where there is commonality across all the companies. I mean, if you're interested in cancer, you may not care about bone disease and vice versa. There are some large companies that try to do everything, but many do not.

And so, it may be as a proof of principle for pushing this concept forward that that would be the right framework, if nothing else to try to standardize things, because it's starting to happen to a bit. We, in this Committee, have spent some time on surrogate safety markers like QT prolongation, et cetera. And there's some--but there's also a lot of those areas that are very different from company to company, and maybe they want to stay that way. But it is one area of commonality.

On the efficacy side, people usually care about a small number of things, and that's going to make it very hard to get people on the same page, even just programmatically.

DR. VENITZ: Hartmut?

DR. DERENDORF: Well, I'm not so sure if it's really a difference, at least not conceptually. I think what we're trying to do with the biomarkers, we're trying to find something that is easy to measure to replace it with something that's hard to measure and do it in a faster way to predict what we would get if we do the hard thing.

So a good example for a safety biomarker that fits in that mold is cortisone suppression for inhaled corticosteroids is a great predictor for long-term osteoporosis or growth retardation in children, studies that would take years to do; you can do it in a single dose study and have a pretty good idea how that product will perform in long-term use. So I think conceptually, it's the same thing. The issues, obviously, are different.

DR. GIACOMINI: Yes, I just want to amplify on the safety biomarkers, I think it's a really good model for bringing together a consortium of people from academia, FDA and industry. First of all, if it's a rare adverse event, it requires large populations, large

clinical populations. I think Paul is participating in the drug-induced hepatotoxicity NIH-sponsored network, right? And that's one that requires a lot of people together, but this could bring together industry, academia, and all of that around safety biomarkers, so I just want to second that.

I also want to say on the efficacy biomarkers, one thing I think that FDA could do is bring together people from different disease-related or treatment-related groups to talk about the issues in those particular treatment-related groups, because I do feel that the biomarkers in each group may be very different, and it would be more conceptual to think about them in group-by-group, disease-by-disease.

DR. VENITZ: Other comments?

DR. LESKO: Yes, just to throw out another thought, and it actually somewhat relates to our discussion yesterday of predictive tests in the context of irinotecan. At some point in time, we're going to have to come face-to-face with the

statistical issues that revolve around the biomarker and the predictiveness of it. And yesterday, when we were talking about a pharmacogenetic test, we were talking about the probabilistic nature of the test and attributes of the test that convey its ability to predict something. We talked about sensitivity, specificity, predictive values, likelihood ratios, et cetera.

And there seemed to be some common ground, or at least we could probably, with more discussion, reach a common ground on the performance of a test that would be generally acceptable. So it gets me around to the question: is an approach or a framework that has been used for the predictiveness of diagnostic screening or other types of tests appropriate for biomarkers? Or is the statistical sort of framework for what we're talking about in place already, or are there needs for new statistical models to deal with this problem?

Dr. Woodcock mentioned the Prentiss

criteria. That was one model. But do we need to be thinking about new statistical approaches, new ways of expressing predictiveness of biomarkers, or are we sort of satisfied with where we are on that, and that may be for Marie and David.

DR. DAVIDIAN: Well, there is a lot of work in the statistical literature; there has been, in fact, recently, as we speak, in trying to sort of refine the--the Prentiss criteria are, let's face it, very stringent criteria, but they do lay out the, I think, what's the key issue for a surrogate, which is that you want the effect of the treatment on the surrogate to--the effect of the treatment on the clinical endpoint to be seen when you paw the treatment, you know, through the surrogate.

So, I mean, I think that is the key issue there. Now, how you go about quantifying that and characterizing that, I think, is what you're talking about. How do you actually do that? And there's been various proposals that are out there to do so. I think to try to get a perfect

surrogate is impossible, as has already been mentioned.

But I think in the context of this sort of discussion here and bringing in mechanistic considerations and so on, I think there would be additional work to be done, and I think bringing statisticians in from that point of view would be a good thing. I mean, most of the work in the statistical literature now, in fact, all of it is totally empirical. It's trying to come up with empirical models and ways of characterizing surrogacy and based totally empirically.

So I think that's where the new work can be done.

DR. JUSKO: The discussions this morning were extremely good and very informative, and as a member of this Committee, I very much encourage all of the participants to continue evolving this area. One thing that is admirable about what companies do is when they screen drugs, they often use receptor systems and animal studies, and eventually, they get to a study commonly called proof of concept, a

phase 2-A type of study, where they then may try to utilize a vast array of potential biomarkers to see whether or not the drug has any activity that's in concert with its basic mechanism of action that they understand it to be. And then, many more studies are pursued after that.

One thing that's frustrating to me in academia is this huge vault of information accumulated by companies in diverse areas, including all of these kinds of biomarkers that they've measured. The FDA may be aware of part of it, but there's probably an immense amount of information that's lost to the general scientific public that could be better harvested if there was some concerted activity through this type of organization that's being proposed here.

So I just want to voice that degree of frustration and encouragement towards collecting some of this information in a more systematic manner.

DR. SINGPURWALLA: I was going to respond to your question. I think I've already said a few

things, and I'm just going to repeat them.

You talked about modeling and simulation in one of your slides, MNS. That's the kind of stuff you hear at the Pentagon all the time, and that's good.

[Laughter.]

DR. SINGPURWALLA: I think one of the things that you may consider in this context of markers is the stochastic process models. You don't want to look at them in a very traditional statistical framework. You want to look at it in a dynamic way. Markers evolve dynamically; diseases evolve dynamically. They're correlated and what kind of inference you should do and what kind of confirmatory studies are needed is something that needs to be researched and worked.

I also hear the word mechanistic models, mechanistic considerations. I would hope that you're looking carefully into Bayesian methods, which combine both the knowledge of medicine and whatever have you with empirical evidence and try to put the two together.

And lastly, I would suggest that when you have these panels of people looking at various things, I would encourage you to go out of the normal umbrella and look into other disciplines. And I just don't have in mind engineers. I strongly suggest you look into the philosophers. They write a lot on causality; in fact, there are a lot of books on causality written by philosophers.

I think also, you should look at ethicists and people who look at moral issues. So I think you should expand your umbrella of expertise to include some other cultures and characters.

DR. BLASCHKE: I want to come back to a question that you raised, Jurgen, and also a point that Marie made. And that is maybe one of the principles of surrogate endpoints and part of this qualification process is that you have an advantage, in fact, if there are multiple drugs to treat the same condition. If you're getting the same effect when you're using drugs, working through what are believed or hypothesized to be different mechanisms, yet at some point, their

effect on a surrogate is consistent and also then consistent with a clinical outcome, it gives you a lot more confidence that this surrogate is, in fact, not an epiphenomenon of some sort but, in fact, is a causal path marker that could be used as a surrogate endpoint.

So perhaps when we're trying to think of sort of general principles and so forth of things that make a biomarker more likely to qualify as a surrogate endpoint, I think the fact that it could--and that could even work with new chemical entities. I mean, even if it's a first in class. I mean, somebody mentioned earlier that maybe it's hard for a first in class compound to be approved on the basis of a surrogate endpoint, but in fact, no. If that surrogate has been proven for several other drug classes, it may even be a stronger evidence that this new drug about which maybe has a new mechanism is ultimately working through that same pathway to produce the beneficial effect in the disease.

DR. STANSKI: Bill Jusko mentioned the

sequestering of information. I'd like to ask people who work within the pharmaceutical industry to what degree is this precompetitive knowledge and prevention of sharing to do patent issues and competitive advantage something that can be overcome? Or is that just a reality of a for-profit industry, or for the sake of moving this concept forward and having more efficient drug development, how can that barrier be broken?

DR. VENITZ: Would anybody care to comment, or was this a rhetorical question?

DR. STANSKI: Well, someone in the industry must think of this and to be able to respond to it, I'd hope.

DR. VENITZ: Go ahead. Can you introduce yourself?

MR. WEBSTER: I'm Chris Webster. I'm director of regulatory strategy and intelligence from Millennium, and I'm speaking for myself here. I'm not speaking for the industry, but perhaps my views are, because I've been involved in some of the working groups, may be useful to you at this

point.

Obviously, everybody is very aware of the topicality of this issue relating to the publication of clinical trials, and there has been, as you know, an initiative published by PhRMA to put up clinical trial data in a public place for patients and physicians and others to see it.

I think what you're talking about here is something more far-reaching than that, and it's not, I think, a--you know, this is not the first time I think the industry has become aware of it. I'll refer you, for example, to the comments of Dr. Kalif at the Science Board last April, where he again touched on this point, and so I think we are aware of it.

I think that it's probably not impossible to be done, but I think that there would need to be some kind of really high level working group to really look at very sensitive and difficult issues related to intellectual property and ways in which information could be perhaps shared in an anonymous way, in a generic way so that it wasn't identified

with particular companies or particular drugs but perhaps could be useful for the purposes of scientific research.

And perhaps some degree of parallel to that is the creation of voluntary data submissions for pharmacogenomic data which, of course, was published by the agency just about a year ago now, and so perhaps, that might be to some extent a model for this.

I think it's very difficult, though; I don't want to project any illusions about this that it would be easy, but I think perhaps it's a conversation which the industry might be ready to have. Thank you.

DR. LESKO: Yes, Chris, while you're there, you did mention the voluntary genomic data submission pathway that the agency created, which was kind of a groundbreaker in many ways, and I know you were part of that with the working group and the workshop. SO, really, my question is do you see a difference between a similar pathway for nongenomic biomarkers as we set up for that

particular reason? We set it up for genomic biomarkers, but is there any reason why it couldn't be utilized for getting some of the information that's sequestered in some of these areas to submit to a group separate and apart as we've set up the interdisciplinary pharmacogenomic review group to do the evaluation of these and begin to synthesize, really, a greater association with the clinical outcomes and so on.

MR. WEBSTER: Yes, I think that's why I suggested it could be a model, and personally, I, myself, don't think that there is a qualitative difference there. But I think that in the sense that genomics is a new science, a new technology; its application to drug development versus drug discovery is something that is perhaps newer; and also, the fact that there was kind of this safe harbor concept around the submission of data, all of those were, I think, if you like, material facts.

Now, as I say, I think it perhaps is a model which we could explore, and if, perhaps, in

the context of this morning's discussion, the agency were to create some parallel to the IPRG but which allowed companies to come in and discuss a broader context of biomarker research with the agency, and if that was part of the entire, if you like, game plan, then, I think that might be a lever to move this forward.

DR. VENITZ: Wolfgang?

DR. SADEE: There are actually companies out there that make their business to compile vast amounts of data of that very nature, for instance, Iconics. And you not only have array data; you have 500 assays available for the 500 common drugs used, and so, that's a business model by itself. And I would strongly suggest that we get this type of folks involved in the process, because they have already integrated much of the information one would like to use, actually.

DR. BARRETT: Larry, I wanted to come back to your initial question about the statistics. In the discussion yesterday, when we got to look at some parameters associated with sensitivity,

specificity, and predictive value, my comment to your question was I don't think I've seen enough of that across different therapeutic areas to where you could make an assessment of that, and they seem to be very reasonable and applied metrics.

The question I had is, you know, it would seem to be a good example where you could use some modeling and simulation to look at what would those metrics look like if you had good association or bad association, if you had a high prevalence rate or low prevalence rate, as well as if the pharmacokinetics were predictive of the biomarker or not.

It would seem to be that you could look at the performance of these characteristics almost independent of their application to define whether or not they were reasonable to look at. But to answer your question, I don't think we've seen enough of it in a standardized manner, which is, again, part of the problem of having enough of a data set to look at across therapeutic areas.

DR. DERENDORF: I liked the proposal that

we've heard many times this morning on collaboration between industry, FDA and academia. But I think there is a big problem coming our way, and that is that we are not training enough scientists in this field. There is a shrinkage of clinical pharmacology programs, pharmacometrics programs, a lack of funding in academia, and this will be a problem. And I think industry really--I feel it's in their own interest to maybe help academia a little bit in establishing systems, how we can provide the training. It's going to be a problem otherwise.

DR. WATKINS: Sorry, just to bounce around a little bit, but in the issue of getting companies to cooperate and sharing data, I'm aware of one initiative which is the International Life Sciences Initiative that's been going on for several years where participating companies are submitting preclinical toxicity data and safety data in man in a blinded fashion, creating a database to look at, you know, markers of predictivity from animals into man, so that there's at least one precedent for

that.

The other thing I thought I would just mention is what Cathy brought up, which is the drug-induced liver injury network as a potential for collaboration with industry and the agency. This is funded by the NIH and the NIDDK in particular. And these five centers, which cover about 12.5 million lives, are prospectively enrolling into the study people who have clinically significant toxicity due to any drug. And in addition, they're getting genomic DNA and immortalizing lymphocytes and getting serum and liver wherever possible; we're also creating a--and I'm chair of the steering committee--creating a registry, and the people agree to be contacted up to 20 years to undergo genotype/phenotype correlation studies in focused clinical centers so that, you know, that seems to me a very nice potential model for industry to participate; obviously, we'll be finding out things about their drugs before they know them, and I'm sure we'd be open to any kind of collaboration that could come

down the pipe.

DR. D'ARGENIO: Yes, this comment also has to do with databases and biomarkers. One of the real challenges in developing these causal paths that are mechanistic-based biomarkers is understanding them and disease progression. And that is a real challenge, but there certainly are data out there on just general models of disease progression, at least one would think, in the postmarket area, and those data would help inform, you know, the relevance of biomarkers to follow disease progression.

DR. CAPPARELLI: I think the last two comments also focus back on the issue of looking at the surrogate marker going backwards as well. You know, one of the issues, even with the disease state, this is a dynamic issue. You know, looking at HIV as the example, working in pediatrics, the surrogates don't work exactly the same.

And so, I think there will be sort of an evolutionary process of understanding the relationship, and that is a huge data mining and

iterative process of working that forward, so, you know, the concept of looking at some key areas, especially ones where the clinical endpoint takes so long to develop, and we may have good mechanistic reasons to think we have something that occurs rapidly that we can measure.

And that was the other aspect of HIV, that the whole research really showed that it wasn't such a static disease that takes a long time, and we can see the effect of drugs very rapidly, and that time differential was, I think, extremely important in bringing that forward from an industry and academic standpoint to utilize these tools.

DR. VENITZ: Any other comments, perhaps on the recommendations that Dr. Wagner talked about with respect to setting up committee structures to manage the process?

[No response.]

DR. VENITZ: Any other comments?

[No response.]

DR. VENITZ: Then, I guess, I'm looking at you, Larry, as the final comment.

DR. LESKO: So, I guess that means it brings us to the end of the road--

DR. VENITZ: Right.

DR. LESKO: --for this meeting, and the closure is stated as a summary of recommendations, and before I do that, I'd like to not be remiss in acknowledging the people that helped put this committee meeting together, and I'm specifically referring to Hilda Scharen, who's sitting next to Dr. Venitz; Karen Summers, who was behind me for most of the meeting, I guess keeping me in line; I'm not sure why, and Bob King, who has been helpful in getting all these materials out to the Committee and my colleague to the left, Peter Lee, who did a lot of the coordination of it.

We didn't make it easy for this crowd this time around. We really imposed upon their administrative support, and I really appreciate their flexibility in meeting deadlines and going the extra mile to get everyone who participated cleared appropriately and within the laws.

As far as the summary of recommendations

goes, I suppose the summary is really captured by the voting that the Committee did on the yes and no questions that we posed yesterday in particular, and there really isn't much more to comment on those questions, because I think they did speak for themselves, although the discussion in between the various questions were very useful to us in illuminating the vagaries that we're dealing with in some of these areas, in particular, the area of transporters and multiple inhibitors.

What was particularly useful to us was what I said yesterday: voting aside, the value of this meeting, the added value of this meeting is really in the areas that surround the discussion of the issues. And the discussions in this Committee meeting were very helpful to us in helping shape our way of thinking about pharmacogenetics, drug interactions and biomarkers, and I think that's why we came here together.

I really enjoyed this meeting. It was quite an interesting intellectual debate. The members, even late last night until 6:00, were

fully engaged. I did miss the after-meeting discussion last night, but I'm sure it was also very intellectual, but you were willing to work hard and late night, and I want to express my thanks on my behalf, and as Dr. Woodcock had to leave to go downtown, she asked me to express her appreciation to the hard work that the Committee did on her behalf as well.

Well, I think this meeting, we really teed up some new issues and some challenging topics, some of which, of course, haven't been resolved. We didn't expect that: transporters, the biomarkers, the surrogate endpoints, and I hope all of you really look forward to further meetings, where we hope to discuss these issues in more details as our thoughts come together and as more data become available.

So in closing, I would like to express my thanks, thanks on behalf of the Clinical Pharmacology team that worked to bring the topics to you. Of course, all of the presenters and to all of you for your time and public service and

providing us the intellectual firepower that we need to resolve these issues. So have safe travels home; thank you, and I'll turn it back to the chair.

DR. VENITZ: I agree. I thank everybody for participating; wish everybody a safe trip home, and the meeting is adjourned.

[Whereupon, at 11:42 a.m., the meeting was concluded.]

- - -