sgg

DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOOD AND DRUG ADMINISTRATION

CENTER FOR BIOLOGICS EVALUATION AND RESEARCH

AND

PLASMA PROTEIN THERAPEUTICS ASSOCIATION

COMPARABILITY STUDIES FOR HUMAN PLASMA-DERIVED THERAPEUTICS

Thursday, May 30, 2002

8:00 a.m.

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

sgg

Doubletree Hotel
Rockville, Maryland

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

C O N T E N T S

C O N T E N T S (Continued)

P R O C E E D I N G S

Opening Remarks

DR. HEALEY: Good morning, ladies and gentlemen. I would like to start today's program. My name is Chris Healey, and I am the executive director for PPTA North America. On behalf of PPTA, I want to thank you all very much for coming. It is wonderful to see so many people here. I think some of you know that the program is so popular that we actually had to stop registration. So, consider yourselves fortunate, I guess, you were able to sneak in. But thanks very much.

Just a couple of things, I did want to, on behalf of PPTA, acknowledge how privileged we feel to have the opportunity to co-sponsor the program with the Food and Drug Administration. It has been a great process. I can tell you it has been one that has been a long time in the making. For over a year we have been working with the folks at FDA and within industry, putting together the program and the agenda to make sure that it is a really meaningful and informative series of presentations over the next day and a half.

In particular, I want to acknowledge Andrew Chang from FDA, who has spent a lot of hours and a lot of his personal time and resources helping design the program, and put it together, and making sure that it is appropriate and beneficial for all of us. Also, PPTA Craig Mendelsohn, who is our director of regulatory affairs, who has worked really hard on this over the past year or so, making sure also that the industry folks kind of had their act together, that we were working well with FDA and making sure the agenda came together. So, particular thanks to the two of them.

Also, in terms of administrative matters, Helms Briscoe is our meeting planner and I think they have done an outstanding job, and our AV resources, Capital AV have been a big help in making sure everything is set up. So, if you have complaints, feel free to go to them. If you have compliments, feel free to come to Craig, myself or Andrew. We will be welcoming those throughout the day.

With respect to questions, we anticipate that there are going to be a lot of questions. This is obviously a very important topic and, given the level of interest that has been expressed, we hope to have a really

dynamic program. We anticipate that we may be running tight in time with questions in particular, but we have index cards that are available so, please, throughout the day jot down any questions you have and we will have folks walking around the room, collecting those index cards. If we don't get to your particular question during the Q&A, there will be panel sessions where we can address additional questions. So, don't worry, we will get to all the questions that we possibly can.

With that, thank you very much once again and I am going to turn it over to Dr. Weinstein.

Agenda Overview

DR. WEINSTEIN: Thank you, Chris. Well, it is a pleasure to be here and I anticipate this meeting will be very productive.

The issue that we will be discussing primarily over the next day and a half will be how best to implement the provisions of our guidance concerning demonstration of comparability of human biological products, including therapeutic biotechnology-derived products. Henceforth we will refer to this as the guidance.

I think that as we go along here, we realize that understanding the provisions of this guidance on part of the FDA and on industry's part will help us to better formulate submissions; will understand each other's concerns and expectations; and eventually it should lead to reaching our overall goal, to provide consumers with safe, pure and potent products in the most expeditious manner.

It should also be remembered that a guidance is just that. A guidance contains recommendations; it does not contain requirements. So, I will emphasize over and over again that there is no substitute for good judgment, and this good judgment is based on good data and on experience and on knowledge. This will be, again, a continuing theme throughout this meeting.

CBER has a number of concerns that are rather routine. We learn at CBER school that there are a number of issues here that are rather common and that we can anticipate might affect the plasma derivatives if there is a manufacturing change. I should point out that this document was issued in 1996 with the intention of reducing the need for clinical trials if it could be demonstrated by other means that product, before a manufacturing change,

are comparable to the product produced after the change. Our problem is in defining exactly what "demonstrate product" and "comparability" actually means. These will be topics that we will be discussing throughout this entire meeting.

I should mention also that a copy of my slides I guess is being handed out now so you might not have to write everything down immediately. Anyway, you know, one can consider this as sort of a rogues' gallery of issues that we have concern about. This includes the production of neoantigenicity, new aggregation, polymerization and degradation after a manufacturing change, oxidation, deamination, altered glycosylation. We will be discussing to some extent during the course of this meeting introduction of vasoactive substances, particularly the pre-kallikrein activator; change in reactivity towards substrates or receptors; introduction or removal of so-called impurities that affect the product safety and efficacy; a change in molecular species distribution, for example, a change in the distribution of immune globulins; and the introduction or proteolytic enzymes.

The important point here is that the more information that a manufacturing can supply us about these particular issues up front in their submission, the more likely is it that we will be able to review the submission quickly and that we will be able to have reduction in the review cycles. I point these particular ones out because these are pretty much the hot topics. We will hear some other issues as we go along, but these are ones that are quite uppermost in our minds when we do a review.

The questions that we are confronted with following a manufacturing change: what is the clinical significance of an observed change in an analytical result? Now, often it occurs that one might see a slight shift in a peak on a chromatography readout, or you might see a slight change in the distribution of product species. The questions that a reviewer has and that you should have in anticipating what the reviewer question is, is what is the significance of that analytical result? How are you to evaluate it? And, it is very important that you assess your data; that you give us up front your assessment of what that change might be rather than just giving us raw

data where we pick out the anomaly and try to decide what it might mean.

The other issue is how confident can you be of product comparability if you observe no change in analytical results. The question there is whether your methodology is sensitive enough, and how can you assure us that, in fact, there isn't something that has crept into the new product?

We think, of course, that there needs to be an approved reciprocal communication about the expanding body of experience in paradigm situations that influence our regulatory decisions. What does that mean? The notion is that we have an increasing body of knowledge as we go along and try to manufacture. The PKA incident is one that, again, we will discuss in some detail. These are experiences that a particular company might have in the production of a product here, but it becomes the paradigm for our thinking about many other products. The point is that it is very important that we have an opportunity to share what these paradigm experiences are, and communicate them throughout the industry so that we will have a common reference point to go to, to be able to have better

submissions; that we are all working essentially off the same page.

So, what I presented in the previous slide was actually a number of issues that are our chief concern here. You will get a sense of our highest levels of concern here are, but again the idea of sharing is very important and, as one outcome of this meeting, I think we should work toward establishing some way of being able to share our experiences both at the FDA and from industry. Actually, this meeting is an example of that kind of sharing of information.

Now, to quickly go through some of the issues that will be presented during the rest of the one and a half days, we will have a discussion from industry and from CBER representatives on perspectives in general and comparability. We will discuss the review guidance in much greater detail. We will be talking about the characteristics of plasma-derived therapeutics as opposed to specified biotech products, can we use the same kinds of instrumentation of analysis on plasma proteins as we use on biotech products? We will talk about approaches to establishing comparability and we will have a rather

extensive review of latest analytical biotechniques that can detect changes in protein structure and product content.

We will also discuss the preclinical testing for product characterization and compatibility. The issue here is what can animal studies tell us? What are the weaknesses and strengths? We will also have a discussion of potentially a new system of detecting neoantigenicity that in the future may be used to give us a sense of a potential of a product to have neoantigen formation during the product change.

We will also talk about the design of clinical studies to demonstrate comparability. The issue there perhaps is should we always strive to establish comparability, or are there some instances where it would be actually a better strategy to consider a product as a new product. You may need less patients in the clinical trial under some circumstances than if you tried to establish that a product is comparable before and after a manufacturing change.

We will also discuss the classification of submissions and/or reports, CBEs, CBE-30s or prior approval

supplements. What should be the criteria for characterization? What are the advantages and disadvantages of a given classification? And, we will give examples of successful and unsuccessful applications of this regulation.

Here is a little quote, "though success is more pleasing, failure is often no less instructive." This is Francis Bacon, in 1590. I actually got this quote from John Finlayson.

Regarding case studies, we will be talking about successful use of a comparability protocol and a comparability study that reduced clinical trial requirements. We will also then have a very extensive discussion of two situations where we had the pre-kallikrein activator as an issue that affected product quality and, again, the notion here that small changes in manufacturing can have a large effect on product quality.

Finally, on the second day we will be talking about comparing fractionation intermediates. We will be discussing the need for fractionation intermediates in the marketplaces, the issue of reduced supplies of plasma, and we will be discussing the criteria for comparing the

intermediates. We will also review the draft guidance entitled, Cooperative Manufacturing Arrangements for Licensed Biologics, and have a case study of the acceptance of a fractionation intermediate.

This is a very ambitious agenda. Chris already alluded to the idea that we do want participation. I think that part of the success of this meeting will be the idea that we can exchange ideas freely. Please, do fill out these cards here and we will try to discuss as many of your questions as possible during our session later in the afternoon. Thank you for your attention.

The first speaker will be Michael Gross, who will give an industry perspective on comparability.

Industry Perspective: Comparability

DR. GROSS: Good morning, everyone and thank you for coming to this meeting that has taken us almost two years to put together. My name is Michael Gross, and I am responsible and proud to present the industry viewpoint on comparability as it relates to plasma derivatives.

Change can occur throughout the product life cycle of a biological product to, amongst other things, improve product quality, enhance compliance with

regulation, improve facilities and equipment, extend production capacity and improve efficiency. A plasma derivative change can occur in basic fractionation, purification, formulation, packaging, storage, testing.

Biologics are said to be defined by their manufacturing process, so making change in the process means making change in the product. This raises issues about adverse effects of change, and it also raises issues about the relevance of previously established clinical data to the product made by a variant process.

The comparability process is a tool with great potential for expanded application in the management of change in our industry. New chemical and biological bioprocess analytical methodologies provide tools to address difficulties encountered in the characterization of biologics, including plasma derivatives. Successful application of modern analytical methods encourages the management of change through comparability approaches.

Over the next day and a half we will discuss challenges, difficulties, success, failure and the future of managing change in the product of plasma derivatives and plasma protein products. We hope to come away from this

meeting with an improved understanding of how to best apply the comparability concept and what changes might need to be made in the future.

I will begin with an attempt to define a few key terms, raise a few questions and make some suggestions for areas that need improvement. The comparability concept concerns make scientifically sound judgments based on data, based on data comparisons; manufacturing and manufacturing history of change, perhaps as much as forty years of it; clinical experience; and it compares two pharmaceutical entities, perhaps a drub substance, a drug product, a key intermediate, one derived from a variant manufacturing process that was used to produce the other, to determine if they are sufficiently the same or similar enough that they can be considered--and I use the next word with some trepidation--equivalent in their effects. In particular, the chemical, biological and clinical effects.

Controlled change is good and the comparability concept was developed to provide a vehicle to accommodate the manufacturer's need to make changes and FDA's need to control change. Comparability was intended to provide post FDAMA regulatory relief, an approach to controlling change

that does not place the regulatory barrier so high as to discourage it.

The comparability protocol is codified in 21 CFR 601.12(e) and had to apply comparability as described in the guidance, 1996 CBER guidance entitled, FDA Guidance Concerning Demonstration of Comparability of Human Biological Products, Including Therapeutic Biotechnology-Derived Products. A comparability study is an experiment, a side by side comparison of properties of a drug substance, a drug product, a process intermediate made by an established process and a variant process. A comparability program is a collection of comparability studies that are well reasoned and well designed and are intended, in part, to eliminate the need for clinical data that provides evidence that a manufacturing change has not adversely impacted identity, purity, safety and potency of a biological product.

It is not a testing hierarchy. Rather, it is a program of studies designed with an understanding of the performance of a validated manufacturing process and the plasma protein product produced.

The design of a comparability program is driven by the nature of the change and its location in the process; the stage of development; the product characteristics; the potential for them to change and affect product purity; the type of product and its intended use; the physicochemical and biological properties of the product and their potential to produce related substances; the suitability and availability of analytical methods to asses the impact of change on these characteristics, and the relationship between product quality and biological activity and product safety and potency.

Comparability comparisons or chemical, physical and biological data for drug substance, drug product or intermediates will usually include routine release test and stability data. This may be supplemented with in-process tests at the manufacturing step most likely to be impacted by the process change. They may also include non-routine tests, including methods at early stage used to characterize the consistency of production, and they also require more advanced tests, newly developed tests if established methods are not sensitive enough.

Besides physical/chemical characterization studies, comparability studies will often include in vitro or in vivo bioassays in models and may include animal pharmacokinetic and pharmacodynamic studies and toxicity studies and, in some cases, even clinical data is required such as immunogenicity data, pharmacology data or maybe even safety data.

After data from a comparability program provides evidence to reasonably conclude that the manufacturing change has not adversely affected product characteristics, the change may be approved by FDA. When comparability, however, cannot be established because it appears that the product has been adversely impacted or perhaps the methods applied are not sensitive enough to establish that no change has occurred, then clinical studies may be required to show that the change is not associated with an adverse effect.

For older, long-established products routine release tests and methodologies may be based on classical bioanalytical methods, but for comparability studies more advanced methods are required and if they are not available clinical studies may be requested.

A comparability protocol is a prior approval supplement describing the comparability program. It is intended to facilitate the review and approval of a discrete change in facilities, equipment or manufacturing process by establishing agreement between FDA and the sponsor over the content of a comparability program and the acceptance criteria. It is intended to provide a route to reduced reporting category.

The biological licensing/reporting categories that are codified in 21 CFR 601.12 are the annual report, changes being affected supplement, changes being affected 30 supplement and the prior approval supplement. Typically, a comparability protocol would be used to downgrade reporting category for a prior approval supplement to a CBE-30 since that is where the biggest bang for the buck occurs.

Once FDA approves a prior approval supplement containing in part the comparability protocol and acceptance criteria, and when the protocol is exercised and preestablished criteria are met, the change may be approved on the bases of a CBE-30.

In a workshop concerning well characterized biologics held about seven years ago, FDA specified four types of biological products considered at the time to be well characterized, namely, therapeutic DNA plasmids, therapeutic synthetic peptides of less than 40 residues, monoclonal antibodies for in vivo use in proteins derived from recombinant DNA.

While the term "well characterized" is not rigorously defined in regulation or guidance, it was stated in the workshop that a well characterized biologic is one whose identity, purity, impurities, potency and quantity can be measured and controlled. The term also suggests to me having detailed knowledge of the mechanism of action, product process and clinical performance such that consistent and predictable manufacturing can be controlled. Well characterized seems to be related to the concept of comparability, although it seems that this relationship is not essential.

Plasma derivatives are not specified by FDA as being well characterized and comparability approaches are still allowed. The importance of well characterized designation in the regulation of change through

comparability approaches is something that will be clarified in this meeting.

What is the downside if plasma proteins are not considered to be well characterized? Can we still manage process change using comparability approaches? Plasma protein products are typically not as highly purified as products derived from biotechnology. They are frequently highly enriched concentrates of endogenous proteins. These may be considered to be less characterizable than a product derived from biotechnology. Is this a problem? Is this characteristic of proteins isolated from natural sources a stumbling block to considering plasma derivatives to be well characterized? If the answer is yes, then what is the effect of being considered to be less characterized? Are we waning in our level of understanding of a relationship between structure function of plasma-derived proteins and the impacts of change on structure function? Is it a problem that in fractionation multiple products are stripped from source plasma pools but in biotechnology the manufacturing process is directed at detection of a single product? If plasma derivatives are not considered to be well characterized and if that matters, then perhaps we can

suggest that they be considered to be substantially or approximately characterized.

The product characteristics of plasma derivatives are routinely measured and controlled using both classical and modern tools of protein chemistry, and some examples are shown here. Safe, potent and pure products have been made, for the most part, consistently over long periods of time. Today using modern methods, biologics can be characterized to a level not previously achievable providing improved opportunities and a stimulus to apply comparability approaches in the regulation of plasma derivatives. You will hear in one of our next talks how modern methods of bioanalytical characterization have been successfully applied to plasma derivatives.

We hope to dialogue over the next day and a half to better understand FDA's expectation about the characterizability of natural biologics, and the relationship of comparability, and what differences there might be between a well characterized biologic and one that is highly characterized but still may not qualify for this distinction.

Three important areas where comparability concepts are being applied are likely to be applied with greater frequency are in the development of new product presentations, such as new strengths, packaging presentations, the upgrading of facilities, manufacturing processes, equipment and in the exchange of fractionation intermediates. The nature of our industry today requires finding efficient ways for manufacturers to exchange fractionation starting materials, intermediates, under an appropriate level of control. Using a company's A cryoprecipitate to manufacture company's B Factor VIII, or company's C manufacture of immune globulins from company's D Fraction II plus III paste is not uncommon, and the practice is likely to increase over time.

Another new and important area for increased application of comparability concepts in the plasma derivatives area is in process, equipment, facilities changes that enable the rapid and economical deployment of new technologies aimed at achieving approved improved process, assuring pathogen safety, reducing product shortages and providing access to new processes, facilities and equipment that better conform to ideal models of GMP.

For comparability to be a useful tool, plasma derivative manufacturers need to develop comparability protocols with confidence that FDA's requirements are understood, and protracted negotiations over the content of a comparability protocol will not be routinely encountered. Predictability is very important. We recognize that a lot of our comparability depends on the specifics of a particular process change. Historically, everything has been case by case; maybe it is time to challenge this.

The comparability concept was developed to reduce regulatory burden, and the need is for industry and FDA to establish together, as best we can, rules, paradigms, guidelines, expectations, etc. to improve our ability to anticipate and plan to meet requirements. The comparability concept has the potential to be an important regulatory tool but for it to be useful, it must be predictable and provide realizable benefits.

The Plasma Protein Therapeutics Association membership hopes that the utility of comparability protocols can be improved through clarification and specification of requirements for plasma derivatives, and

we hope that this meeting will catalyze this effort. Thank you.

DR. WEINSTEIN: The next speaker will be Chris Joneckis, who will give the CBER perspective on comparability.

CBER Perspective: Comparability

DR. JONECKIS: Good morning. My name Chris Joneckis. I am the senior advisor for chemistry manufacturing and work for Dr. Zoon, which means I do a lot of interesting things and one of those things recently, for a period of time, has been comparability. Let me point out that this conference is very timely in that we are taking an internal look across CBER at what we have done regarding comparability over the past six years since the guidance, as it has been termed, has been issued. So, this conference is very useful and important to, I think, let us sit back and help us in our internal thinking as well.

Secondly, it is important I think also because there is an ICH initiative to development a concept or, rather, a guidance on comparability and, although the proposed scope will focus on biotechnology products, sometimes those guidances seem to have applicability in

other regions towards a broader variety of products. So, that is also very useful.

What I am going to talk about today is a little bit more of the broader perspective of comparability across CBER. Everyone is well familiar with the CBER mission statement, but the part I would like to focus on is the fact that the regulation of these products is founded on science and law to ensure that purity, potency, safety efficacy and especially availability of those products. Before that term "availability" was formally added to the mission statement, CBER had a long history of partnering and working with industry to facilitate the delivery of products and approved product changes. As you have heard, that in part facilitated the development of this guidance back in April of 1996.

It was clearly made possible by a lot of advancements in time of manufacturing methods, process control methods such as validation and other tests, and analytical tests to assess products, the drug substance, the intermediates and such. The guidance was devised for all biological products regulated by CBER and, as you have heard, the key product is to demonstrate product

comparability of a pre-change and a post-change product, and the whole issue was at this point that, depending upon several factors which some speakers have spoken about, one may allow a change without the necessity for preclinical and/or clinical testing.

Well, what is comparability? The closest we come to a CBER definition is taken from that guidance, and it is that FDA may determine that two products are comparable if the results of comparability testing demonstrate that the manufacturing change does not affect safety, identity, purity or potency, essentially a very broad and operational term. Others have proposed different definitions that I have seen in various forums.

I think the other way to define comparability is what it is and what it is not. The tests that are used for a comparability program do not really allow us to determine that the pre- and post-change products are identical in fact, although, for example, the analytical methods may be able to say that the pre- and post-change products are indistinguishable is the key point, I think, to make.

The second point is that you can also say what it is not. Well, clearly you can move into the range of it

being different. That also depends upon how one defines "different" and what the operational types of terms are used, the criteria that are applied, although with the caveat that you may have certain differences for example in analytical assessments as long as they don't translate into significant clinical safety and efficacy effects. So, comparability falls somewhere, in a sense, in that area. The concept of comparability, again, is across the life cycle of the product and that is how it is applied in the guidance and how it is applied at CBER. Of course, there is deference to the fact that there has to be some flexibility in where one is in this whole life cycle to allow for flexibility for process and product development, especially early on, but, again, in those situations you are always doing some type of clinical and perhaps preclinical testing. The real issues that come into play are more during the Phase III or post Phase III changes and post-approval changes where, again, the need may not be to do preclinical and/or clinical testing, and I think that is largely the focus of the issues today.

The other point is, as people have said, that the comparability protocol, which is basically an operational

mechanism which may allow one to get a reduced reporting burden on this post-approval change and may in certain cases allow for expedited product release, has been used and has been used quite successfully by a whole variety of products across CBER. The point to make though is that the concept of comparability clearly falls into this comparability protocol.

The elements, as other speakers have said, of the comparability concept, as well as various considerations in developing this comparability assessment program, can be sort of grouped into several areas: process, product, analytical, predictability and manufacturing changes. This is the way I have characterized them.

Just a few quick comments on that, the old dogma or the dogma that the process is the product may or may not still be applicable; may or may not be applicable depending on the product class that one is speaking about. Whether that is a debatable topic, I think it is still clear to say that the process clearly defines what that product is going to be regardless of the starting source or starting material that may be used. Again, the products across CBER are still heterogeneous. There is an inherent

heterogeneity across those various products. There are also considerations for certain molecular complexity and the influence of impurities which subsequent speakers will discuss.

Also, the analytical ability, where we are currently in the analytical capabilities--what are the capabilities to detect differences? What are the inherent limitations of all methods that we face not only in looking at the comparability issue but in approving new products? What is the ability to detect small differences in large molecules that may have profound consequences? Again, the predictability, as other speakers have said, is based largely on the knowledge, history and experience of your product and your process come from developmental studies. Lastly, the type and extent, the complexity, if you will, and where you are in the life cycle concept of the manufacturing change.

So, all these are looked at in considering how to determine the comparability assessment program. In many ways they are interrelated--again, what kind of process and the purity of that resulting product and the heterogeneity and complexity of that product to some extent determine

what analytical methods you may be able to apply, and the results that come from those analytical methods. The type and extent of this change will also influence what types of additional studies may be needed. Again, as has been identified in the guidance, the key component is what you can predict.

The next three slides is material I have already covered on this slide, so I will just skip them. This is the life cycle of CBER, as it has been called the world of products. When you look at comparability, comparability really has been applied to this upper quadrant of products, mostly because these products for the most part are still in the development cycle and there are very few licensed types of products. Plasma derivatives will be the subject of this conference so I won't discuss those but just a few comments on these other types of products.

Vaccines has had very limited use of post-approval types of comparabilities. In many ways, they are heterogeneous, naturally traditionally derived material and they have inherent concerns about being able to define the heterogeneity and the complexity of those constituents in the drug substance, including the active ingredient. So,

there has been rather limited use of the comparability post-approval, at least in terms of being able to not require some type of preclinical or clinical testing.

The largest group of experience we have had with comparability has been for the specified products, in particular these two highlighted here, the therapeutic DNA-derived products and the monoclonal antibodies from recombinant or naturally derived sources.

As other speakers have alluded to, what is the relationship among these various products? If one looks at a synthetic product which may include things such as chemical identity or a synthetic peptide versus the specified products as defined by the FDA versus some of the more traditional products, in general I think people have indicated that as one moves from left to right across the screen there is increasing impurity, increasing ability to be characterized and perhaps to some extent decreasing heterogeneity, although I should point out that I think there are, as I have tried to show here schematically, overlaps within these various types of product classes. So, in fact, there has been a naturally derived product that does have properties similar to that of the synthetic.

So, it may be easier to talk about individual examples or products within a class as opposed to the entire class altogether. I think that is reflected in how we approach comparability and is part of the overall assessment that one has to make in developing a comparability program.

I would like to spend the next few slides on examples that we have learned from specified products. Time does not allow me to go into substantial amount of detail, but I think it is important that we take a look at what we have learned from a well characterized group of products.

The reasons for manufacturing changes in these products is summarized here. I think it is generally applicable to say that these are also the types of changes that we see for various products across CBER. Again, there are the standard types of pre-approval process development. This is a big one, increasing product supply for many of our products and it has been accomplished through process optimization such as increasing yield, scale-ups or duplication of existing processes or additional manufacturing sites, either from a contract nature or from additional sites put on by the initial manufacturer.

Again, they may be driven by just process optimization, updating the process for economic concerns or newer technologies coming on line. Again, there are always compliance-driven elements, for example, the need to eliminate human and animal derived components, or inspectional issues, or general facilities improvement. So, I think those are sort of the types of changes that we have seen.

So, what can we say about specified products generally? Well, overall the scale-ups are generally less problematic in causing problems. It sort of makes sense when one stays with the same process and the same principles of those processes, they seem to be less problematic in types of concerns and issues that we have or have seen.

Changes in an early manufacturing stage, the cell bank or the cell culture fermentation are of great concern and do have great potential to impact the drug substance. There are several examples where we have some additional problems in terms of looking at comparability.

There have been examples of inadequate characterization of raw materials or components that can

affect products, and those can be naturally derived materials, biologically derived materials or actually even chemical entities. There are examples where if chemical entities have not been completely characterized or characterized to the right quality issues, they can profoundly change the whole molecule and cause products to not be comparable. There have been examples of that.

Formulation changes can affect products, especially for sub cu. and im. products. There are several examples where that has happened. And, there are several examples where there have been changes at site of manufacture that can affect the drug substance and drug product. Picking up the product and moving it to a new location with no or extremely minimal changes that one would not predict to have an impact have had changes. Examples have occurred during late product development, such as Phase III type of situations or even post-approval.

Overall, meeting drug product release specifications alone is not sufficient. Characterization studies must be performed, I think, for all the changes that we have had for specified products, as well as other products--the complete analytical characterization, the

complete package has to be performed, as well as perhaps relevant in-process control tests on appropriate intermediates.

Changes for within specifications may be important. There have been examples where even though specifications have been established for products and all of the resulting post-product changes have met specifications, there have been situations where one product or one product lot has been out of trend, if you will, and that has been shown not to be particularly comparable. What does that say? Does that say that perhaps specifications were not appropriately established to begin with? Appropriate characterization or process development studies were not performed? That is possible, but it is important to highlight that, again, within specifications changes can be important and they don't tell you everything.

What can be very sensitive measures is trending the acceptance criteria for multiple pre- and post-change lots, in addition to doing side by side characterization of, say, your typical three lots that people like to do. Looking at those trends has been very enlightening in terms

of seeing what types of changes have occurred and, again, combining that with what type of changes may be seen within specifications has been particularly illustrative. It is difficult, I realize, to do this especially as it requires a substantial post-change material. Of course, drug product and drug substance can change upon storage and you need to consider that in development of the post-approval comparability protocol.

Analytical testing has been the basis for establishing product comparability in specified products in some cases, actually in quite a few cases. The PK data may or may not be needed. Again, it is part of that overall algorithm and considerations that have to be looked at. For more substantial changes, depending on how one looks at that, PK studies are again particularly needed. PD studies really aren't done for specified products for many reasons. There are very few PD studies that we have had. Hence, it is really not utilized per se.

Clinical efficacy and safety data generally is less likely to be needed, although it has been requested in some cases. Again, assessing immunogenicity is of increasing importance.

To illustrate that point, I will talk about one specific example which I am sure is familiar to many people in the room, and that is with the product erythropoietin Eprex, which is made in a mammalian system. It is not marketed in the United States but is marketed and distributed in Europe, Canada and Australia. There was somewhat recently a manufacturing change that was made, including a protein free formulation that included removal of human serum albumin. As of April 30th, there have been over 116 cases of suspected red blood cell aplasia that have been reported to the FDA. Approximately 85 percent of those cases have been confirmed through bone marrow biopsy to, in fact, be red cell aplasia, and approximately 50 percent of the patients evaluated have had high antibody titers against erythropoietin. In other words, this is an example where making a change in a process that would not be predicted to cause this effect has resulted in the development of an antibody against the Eprex molecule and has resulted in removal of the endogenous epogen in the individuals in these cases. Those cases need red blood cell transfusions in order to continue to survive in many cases.

There are theoretical hypotheses as to why this has happened, but nothing has been conclusively determined at this point. There are reference articles that describe some of these earlier cases, and also a response from our epidemiology group and other groups within the agency that you can look at if you would like.

I bring this point sort of towards the end to make the point that these changes in comparability can have profound enduring consequences, and our ability to predict these kinds of changes from our knowledge and history is still limited.

So, there is no established formula for determining comparability testing requirements for specified products but I think that is the universal theme across all of our products. Despite best efforts to detect product differences and predict the impact of manufacturing changes, these surprises do continue to occur, again, echoing the theme made by the earlier speakers about where are we with our knowledge and experience database, if you will?

The algorithm that is described, if you will, in the comparability guidance has in many cases caught changes

for specified products that have not been seen through analytical testing. So, we have been fortunate in not being able to have any of those severe consequences from comparable products, such as the Eprex example, happening for products that are regulated by CBER.

In one of Henry James' novels, he writes experience is never limited and is never complete. We realize that we only see what you all show us; we don't make these products. In speaking with several colleagues, I know there are other experiences out there from changes that have been seen in untoward effects. So, I think that that be shared and discussed is very important because, again, it goes to expanding the general knowledge base that we all have.

The road ahead? Well, the changes that allow us to implement the 1996 guidance, those manufacturing methods, analytical methods and expanding our knowledge base continue to advance and change and I think they may be useful in future to affect what we can determine from analytical, for example, assessments. There are certainly new analytical methods coming along that provide more direct measures of structure and, of course, in

relationship to activity which is very important. We, at CBER, also spend a lot of time doing this. We have NMR facilities to look at the structural techniques. We are spending a lot of time doing proteonomics and other laboratory CHIP types of techniques to try to see what additional types of information can be provided from these methods, and the confidence that can be gained from those particular methods. The bottom line, however, is that approval of any product or product made from that manufacturing change must always ensure quality, safety and efficacy of that product.

Lastly, I would like to thank the individuals in the Office of Therapeutic Research and Review for some of the information provided regarding specified products, and the Office of Vaccines regarding the information they provided on the comparability experiences that they have had. Thank you.

DR. HAYES: Good morning. I am Tim Hayes, with the American Red Cross, and I am going to be co-moderating in the next session with Mark Weinstein. The next session is the comparability studies for human plasma-derived therapeutics, and we are going to be taking a closer look

at the individual components of that, product characterization preclinical studies, as well as clinical studies and how those work together.

It is my pleasure to introduce our first speaker, Dr. Andrew Chang. He is a special assistant to the director for the Division of Hematology, Office of Blood Research and Review. He is going to give us the FDA perspective on product characterization.

I would just like to add while the slides are coming up that Andrew and some of the other people in the audience have actually worked together with another FDA co-sponsored event, for well-characterized biological products conference, and have actually been working and gotten a very good start on this issue of appropriate application of characterization to these types of studies. Andrew?

Product Characterization

FDA Perspective

DR. CHANG: Thank you, Dr. Hayes. Thank you for the introduction.

Before I start my presentation, I would like to use this opportunity to thank the working committee for this workshop. Those include industry people and the

people in the FDA. For the industry people, we have Craig Mendelsohn, who is a co-chair for this workshop, and Michael Gross and Chris Healey and Jean Huxsoll. I apologize if I have missed anybody from industry side of the working committee. From the CBER side we have Chris Joneckis, Mark Weinstein, John Finkbohner, Timothy Lee and also Joe Wilczek and myself. I would like to thank the committee members for their hard work and effort to make this meeting a success.

Also, I have two housekeeping items. For people asking questions, please use the microphone. We are recording this workshop so we want to capture the questions. We may not have enough time for the questions that you are going to ask. You also can use index cards and, as Dr. Chris Healey mentioned, we will have a panel discussion later and use your index cards for your questions.

Another thing is, speakers, please, keep within your time frame. We want to finish this workshop on time. This also applies to our moderators. I would also like to thank all the speakers for volunteering to give their speech in this workshop.

My talk will focus on the product characterization for plasma derivatives. I have the following topics and I will very briefly go through the comparability policy that we have. The previous speaker has already covered that area extensively so I will keep that very brief. I want to focus on the experience that we have had in the past five years since the publication of the comparability guidance in 1996, associated with the plasma derivatives and some of the recombinant products that are found in the hematological product criteria. Lastly, I will go through some of the FDA's perspectives.

The FDA 1996 guidance on comparability actually is a policy that we published in terms of comparability. The policy has resulted from the desire to make improvements in the test methods and product production process within a single manufacturer--I emphasize single manufacturer, the same sponsor to make manufacturing changes for licensed products. FDA may determine that two products are comparable if the results of the comparability testing demonstrate that a manufacturing change does not affect safety, identity, purity or potency. This policy allows for changes in product characteristics if they have

no adverse effect. As Dr. Joneckis pointed out in one of his slides, the comparability is really between the identical and also differences. We have room to work with in terms of comparability as long as the differences will not affect the safety, identity, purity and potency.

In the guidance, as spelled out, there is a three steps approach. The first step is analytical functional approach, then preclinical and clinical. The one thing I want to emphasize here is that this is not simply a hierarchical system but, rather, a complementary one. For example, if you find some differences in your preclinical study that may trigger some additional in vitro analytical functional study. So, it is not that you have an individual study and then you leave that stage, it is, rather, complementary.

Now I am going to change my topic to the experience we have had in the plasma derivatives area. Some people may ask if the comparability concept been used for plasma derivatives, and the answer is yes. We have seen the comparability approach applied to the manufacturing changes for plasma derivatives. Then, some people may ask when does that complication approach start?

Did it start after the 1996 comparability guidance? Dr. John Finlayson has said, no, we have had that concept a long time ago, before the first biotech recombinant procedure was licensed in this country. So, the concept is there and has been used for the plasma derivatives.

Do we have any concerns? Yes, we do have many concerns, as Dr. Weinstein pointed out in his presentation. He listed some major concerns that we have in dealing with plasma derivatives, as well as all the biological products. Later I will point out some significant concerns for plasma derivatives.

I have tried to categorize the plasma derivatives that we have licensed in this country. These include coagulation factors, such as anti-hemophilia factors and von Willebrand factors. I have a list here and I am not going to go through the whole list. Then, we have another type of plasma derivatives such as albumin and TPF. We have a couple of NT protease inhibitors, such as alpha-1 antitrypsin and NT thrombin 3. We have a family of immunoglobulins that I have listed. That may not be the complete list for that type of product but I have listed these over here. We also have five recombinant coagulation

factors that have been licensed in this country. Those include the BeneFIX, ReFacto, Kogenate FS, Recombinate and the Novo Seven. We have one recombinant antihemophilic factor concentrate for further manufacturing.

What is our experience? How often are we seeing major manufacturing changes? With the help of the staff in the Division of Hematology, I have gathered some information just for your information. In the past five years we have received from 70 to 100 supplements for major changes. The previous speaker already gave you some indication of different categories of changes. Also, we have another session in the afternoon to further discuss the reporting category and requirements.

I am going to focus on the pre-approval supplements which are normally used for the major changes to a licensed product. The number for major supplements is actually quite steady since 1996 and also this is true for plasma derivatives and also for recombinant hematological biologics. At the top I list the number of the major supplements that require clinical data to support the change. You see that there are actually very few numbers of supplements that require clinical data.

We have six of the prior to approval supplements that require clinical data to support the change. Actually, three of them are efficacy supplements for a new indication. So, in terms of manufacturing changes we only have three major supplements for manufacturing change, and those changes include formulation changes and also alternate manufacturing processes. The percentage for those major manufacturing changes requiring clinical data is about 1.31 percent. As I said earlier, three of them are efficacy supplements for a new indication. So, the number is cut in half so it is, let's say, 0.7 percent.

If you look at the data in 2001 and then separate it to the different categories in terms of the changes, we have found that 13 of the major supplements involved a change for the process; 13 of them with assay, which could be a new assay or introducing a new standard for the analytical assays; 24 of them are one-time exceptions. I don't know whether you know this term. One-time exception supplement is that the manufacturer sends a supplement to the agency requesting to release one or a few lots that had some manufacturing differences from what they have normally have as permitted by the license. We have other 35

supplements involving other categories, such as stability issues and labeling changes. Again, this data was collected for the plasma derivatives and also the hematological recombinant products.

What is our approach, regulatory approach for major manufacturing changes? It has been handled on a case by case situation. The factors that influence that approach really depend on the following three areas, one is product. We have different products and we have different knowledge on the different products, and we know the risk involved with the product are different. The second element is type of manufacturing changes and, lastly, risk analysis and assessment that play a role in the regulatory approaches.

Instead of giving you some specific examples, I decided to give you two types of examples that will cover each end of the major manufacturing changes. The first series of examples includes type of changes such as a new facility with no change in in-process control, no change in specification, demonstration of individual comparability. Another type of change is a new assay standard for quality

control, lot release. The third one is a one-time exception.

What is the review mechanism for this type of change? We review this type of change as a prior approval supplement for which, under PDUFA 2, we have four-month review time for that. This normally includes data for individual biochemical, biophysical characterization and for some of supplements, such as new facility, we also conduct a pre-approval inspection.

Another type of example, which is at the other end of the major manufacturing changes, is the most difficult change, involved with a lot of changes. This type of change is like a new facility with an automated process, and also changes in specification for the drug substance and the drug product for which a company will demonstrate comparability.

Now, very often--actually, it is always the case that the sponsor voluntarily phases out the old process after agency approval for the new process. The review mechanism under PDUFA 2 is ten months review time because for this type of manufacturing change clinical data is required. This will include in vitro biochemical,

biophysical characterization, preclinical studies and some of the preclinical studies such as human pharmacokinetic data and some of the safety and efficacy clinical data. Pre-approval inspections are always required for this type of change. Very often the sponsor proposes a new proprietary name for the product manufactured with that alternate process.

Conclusions, comparability approaches apply to both plasma and the biotech-derived biologics. I said earlier that that has been used a long time ago. In our experience, clinical data has seldom been required to support manufacturing changes, however, major concerns remain.

What are those concerns? In addition to the concerns that Dr. Weinstein mentioned earlier, we have the following concerns, such as poorly defined study material, source plasma versus recovered plasma, different pool sizes can be used for manufacturing, and demographical and racial differences in the source material and, in addition, some of the manufacturing steps with different intermediates. That is another topic that we are going to cover tomorrow. Lack of robustness of the manufacturing process, minor

changes with major impact, and we will have some example case studies in this workshop to demonstrate minor changes with major impact on the product safety and efficacy.

We have a type of product with very low purity, and impurities may affect activity, immunogenicity or absorption. Often highly complex and heterogeneous proteins for plasma derivatives and also it is quite important that we have a history of viral transmission for this type of product.

So, what is the FDA perspective? Since my presentation is under the individual characterization I am going to focus on the individual characterization part. Analytical and functional testing, physicochemical, functional, biological, immunological studies have been used and performed to support the comparability. Sensitive tests measure all criteria functions of the product. We would expect that because of the complexity of the product, some of the proteins have more than one active site. The example for that is, for example, von Willebrand factor. Considerations for product related and process related impurities and contaminants; qualitative and quantitative

assessment, validated and qualified methods should be used for your study.

Another thing has been touched upon by Dr. Gross, and that is that we are not satisfied with your routine testing to support comparability studies. That routine testing includes in-process control testing, the final release testing to meet specifications and we ask for additional testing to support the comparability. What type of additional testing will be required for that particular manufacturing change? Well, it depends on the change involved and the product involved. So, there is no easy way to give you a formula to do that. You know best because you know the product and the process more than anybody else. So, you should be able to come up with a program to assess the comparability.

The approaches to establish comparability started out in the 1996 guidance. What I am going to do here is just point to some language in the guidance. Side by side comparison--manufacturers should provide to FDA extensive chemical, physical and bioactivity comparisons with side by side analysis of the old product and the qualification lots

of the new product. This is language in the guidance for side by side comparison.

Also, reference standards should be used in the comparability studies. This is also spelled out in the guidance. When available, fully characterized reference standards for a drug substance in the final container material should also be used. Lastly, comparison with historical data. That is especially important when the reviewer looks at some in-process control parameters and we are expecting the company to do some statistical analyses to demonstrate that a particular step has not been changed or is comparable.

I want to emphasize "know they process and they product." This term has been used in one of the previous workshops for process validation so I quoted that. Establish validated manufacturing experience and product history. That is a very important element. Conduct thorough drug product, drug substance characterization. Establish specifications, statistical trending, sensitive discrimination assays. As you all know, specifications should rely on your historical data as well as supportive clinical data. In order to have meaningful specifications,

your process should be well controlled. Those in-process control parameters and the range for those control should have good correlation with your final specifications.

It is difficult to generalize impact regarding a specific change to all classes of products. Again, we have three different categories that we are looking at when we deal with manufacturing changes. Those are product and type of manufacturing change, and also risk assessment.

Issues for consideration--these are issues that Dr. Hayes mentioned. We had a workshop, a 2002 symposium. In that symposium Dr. Hayes and I co-chaired a workshop to look at some manufacturing gaps that we have between the specified product and the plasma-derived product. We actually found that there are some gaps for the plasma derivatives. These include the source material, can or should study source material, in this case for plasma derivatives plasma be better characterized? Can or should the manufacturing process for human plasma biologic be validated to the same extent as that for a specified product? If a plasma protein is heterogeneous, to what extent should heterogeneity be characterized? How should impurities be assessed?

My last slide is with some recommendations. Control source study materials and characterize them if feasible. Establish robust, reproducible and validated processes. Establish sensitive and discriminating characterization and release testing. Document manufacturing history and experience. Establish normal variation of a product supported by statistical analysis and clinical experience. Conduct and submit developmental pilot studies for manufacturing changes. This is especially important for a comparability protocol. Consider applying new analytical technologies. Qualify impurities. Toxic impurities should be identified and controlled.

That will be all for my presentation. Thank you for your attention.

DR. HAYES: Thank you very much, Andrew. Our next speaker--we need to be moving along here because, again, we have a very ambitious schedule--is Ghiorghis Ghenbot, who was also a participant in that workshop at the WCBP, and he will be giving us the industry's perspective as far as product characterization is concerned as involved with comparability protocols.

Industry Perspective

DR. GHENBOT: Good morning. I will present on product characterization industry perspective. When I was preparing for this seminar or workshop, I had several specific questions to answer for myself. Somehow, they will be the same questions that you kind of ask yourself when you consider characterization of plasma derivatives or, for that matter, biotechnology derived products.

Here are the questions. Number one, why should we characterize a product? Two, what should be a characterization strategy be for a plasma derivative protein package include? Three, is there a procedure or a precedence for this activity? Do we have some experience from the past? If so, can this experience be used to characterize plasma-derived proteins? Finally, how is this characterization activity related to a comparability protocol and comparability activities? After all, that is exactly why we are here. We may vary in terms of answers to these questions but I hope we have some sort of agreement somewhere during this presentation.

I will try to give you my answers. Here they are. Why should we characterize a product? I think we

should characterize a product because particular characterization is an integral component of setting specifications, product specifications. This is very clearly stated in Q6B. As a matter of fact, the ICH guidelines for test procedures and acceptance criteria for biotechnology/biological products under Q6B state that specifications are one part of a product strategy designed to ensure product quality and consistency.

Secondly, we should characterize a product because proper characterization ensures the safety, purity and potency of the biological product.

Third, characterization helps in identifying the criteria parameters for defining product quality. Characterization data places comparability programs on a secure foundation. Again, back to a CBER publication of 1996, it states the following, manufacturers should provide to FDA extensive chemical, physical and bioactivity comparison with side by side analyses of the old product and the qualification lots of the new product. So, the directions are already stated in the documentation.

Also, proper characterization of plasma-derived proteins ensures the development of well-characterized

plasma therapeutics. The bottom line is we need to characterize it to make sure that things are fully under control.

Finally, I think product characterization is a sound business decision. In this case, the FDA has accepted manufacturing or controlled changes for well-characterized products again case by case, without clinical proof of product safety and efficacy in terms of money and dollar investments, I think this is a huge saving for any company which is involved in this type of activities.

Now, when do we characterize the product? This is not going to be an easy one. The bottom line is I cannot say that I can characterize a product at one time and that is it. No way. I think you should cover the lifetime of product. Therefore, I submit to you that we should start characterization activities during early stage of product development. Why is that? Because characterization efforts should proceed hand-in-hand with normal procedures of assay development, with particular emphasis on biological activity, so that one assures that the assay is a surrogate of the proposed physiological activity.

Second, we should also continue characterization of the product in Phase II when clinical data are evaluated. In this case, dose and efficacy targets are being devised. Process or formulation changes may be made, and physical chemical tests that ensure lot-to-lot consistency, safety issues surrounding impurity profiles, product heterogeneity and stability need to be further defined.

Also, we should continue characterization in Phase III, and that is because this is the stage where emphasis is placed on validation to show that the product meets specifications. In this case, product characterization is needed at this stage to assure and justify specifications. We have a chance to get rid of some of the assays that we have developed throughout Phase I and Phase II.

Also, we should continue characterization because at some point there is a need for definitive data. Finally, and very important, times are changing and times do change. And, almost every speaker before has alluded to this one, here.

Now, what should a characterization strategy for a plasma-derived protein therapeutic include? The elements that I think this characterization strategy package should include are the following: identity, quantity, purity, impurities, potency and safety. There is nothing new here. The difference is how should we form this package in such a way as to take advantage of the past in terms of biotechnology-derived products and now for our plasma-derived products? It was also very clearly stated that actually the experience is based not on biotechnology-drug products but actually on plasma-derived products. So, we are kind of coming back.

In terms of identity, I submit to you that one has to look for a highly specific test reflecting the unique aspects of the product structure. In other words, focus on what you need for that particular product and not go by a certain specific definition. In this case, one or more tests based on physicochemical, biological and/or immunochemical methods should be appropriate. In terms of identity, I think we can sum up the activity in terms of structural characterization as well as physicochemical characterization.

For the structural characterization aspect, there is basic information such as amino acid composition and N-terminal sequence and peptide mapping to look for some specific type of activities that we care about. Most importantly, these two, I submit to you, should be very carefully monitored for plasma-derived protein studies, post-translational modification in the form of carbohydrate composition, if there are any issues of sulphation and other things. In terms of carbohydrate composition, as I said before, we need to look at the structure as well as composition. Bear in mind that I am not focusing on sequence here; I am actually looking at the package of that profile and now it changes throughout the process and if there is any issue regarding sort of the total amounts.

To go back a little bit, I would submit to you I would be concerned about this issue if there is any concern about heterogeneity of that product.

Furthermore, in physicochemical characterization there are techniques available, old techniques as well as new techniques. There will be cases where we need to revisit the molecular size of the product that we have in hand. There is always that issue here. We have done it

twenty years ago, we have done it thirty years ago, but technology has changed and now we have newer elements or newer instruments so that we can fine-tune our studies in this case. It is very easy and it is possible to do it either by instruments such as MALDI-TDF as well as the ES-MS, right here.

In terms of the electrophoretic profile, we also have newer techniques, SEC. In terms of chromatographic patterns, we can also look at the various procedures such as reverse phase HPLC as well as affinity chromatography. As you know, in my presentation I am really focusing on an in vitro characterization approach because the other speakers are going to cover the in vivo approach.

I kind of like the bottom part of this slide, and that is because there is a possibility, in addition to the physicochemical characterization outlined above, to correlate activity with structure and perhaps, with some data in the future, with immunogenicity. I strongly believe that if there is some correlation between structure and function, the product should be well under control. In this case, you can look at the tertiary structure of the

molecule of interest using spectroscopy, calorimetry, analytical ultracentrifugation, as well as SPR technology.

Now, the following are few examples of a plasma-derived product. In the process of characterization some of the experience that we have here is very simple. This is HPLC. There is no big deal about it. But I think we also need to look at the reverse phase HPLC simply because if there is a possibility for breakdown products or related impurities, one can look at extended programs of this sort of approach.

The bottom part of this slide shows the electrophoretic pattern of certain plasma-derived products of the same product, by the way, and this one was also controlled by capillary zone electrophoresis.

Another example on identity is based on evaluation of the product, in this case using CD. You can look at far UV CD which is not that important in my experience, but I kind of believe that near UV CD has pretty good correlation in terms of structure and function.

More importantly, you could also use the same CD to look for the protein stability, and this issue becomes very important when you deal with formulation changes,

stability issues and shelf life. That is because we would like to know at what temperatures we can start to store material. One can monitor or simply make a certain baseline based on the CD profile of the transition temperature measurement for the protein, or you could do the same thing using differential scanning calorimetry. In this case, it is interesting to note that there is no difference whatsoever in terms of the temperature observed. Now, if that is the result we get every time, I think that would be very easy.

Moving on to the other part of the elements of product characterization, there are certain things to consider when we talk about quantitation of a biological product. There are a number of methods that every one of us is really kind of familiar with. Methods such as shown here have been around for a while, but in more recent days people have also shown that analytical ultracentrifugation can also be used to quantitate proteins.

Why do I point to this particular one? Simply because with the same system you can also look at binding characteristics; you can look at fragmentation; you can look at product heterogeneity, although the system is

extremely expensive and issues of validation and things like that are of question. That is, it is purely investigational at this stage, at least as far as I know.

The other element of product characterization is the purity/impurity profile of the product. In this case, I would like to look at two issues. I would like to look at product-related impurities as well as process-related impurities.

In terms of product-related impurities, we can talk about truncated forms of the product. There is a possibility that some sorts of fragments of that product can be formed. I kind of think that it is very unlikely, especially with plasma-derived products, although there are very well documented cases where there can be some sort of heterogeneity in terms of the N-terminal sequence. I haven't seen that much in terms of C-terminal sequence. This perhaps is important when one has to consider the purity of the protein as well as the biological activity.

Also, and very often, chemically modified forms do occur. That is perhaps due to oxidation, most of the time, or perhaps due to deamination and maybe sometimes due to isomerization. I kind of believe that post-

translational modification for plasma-derived proteins in the dorm of modification of carbohydrate groups is very unlikely to take place.

The other part of this purity/impurity profile is looking into process-related impurities. Now, the previous speakers have mentioned the safety issues with plasma source material, and that is a huge chapter in itself and I have no interest in really focusing on that particular topic now but it is very important and it is hard to cover all the aspects of safety issues at this particular time because I am focusing on in vitro methods.

On the other hand, there are methods for looking for those agents that we suspect to be there. The only point or the only place that I would like to talk about is the bottom part of this slide. We are talking about downstream process and you can use LC methods to monitor these types of impurities. Again, it is not generally in plasma-derived proteins.

Other methods for procedures to look for structurally related impurities in the product that we have, what are the analytical studies? An easy way to look at this is probably is actually to break down the type of

changes that you see in the problem that you have at hand. You can very easily classify the type of structure in terms of those types of changes that are derived because of the chemical instability of the product of interest, or it is simply because of the physical instability of the product.

In terms of chemical instability, you can have some sort of breakdown or degradation, as well as oxidation, deamidation, disulfide exchange and glycosylation issues. The strategy that one can follow is a combination of seven different procedures, and there is no one particular method that one can say this is the method of choice to look for these problems but, in general, a combination of reverse phase HPLC and mass spectrometry, as well as capillary zone electrophoresis will do the job.

The final part of this part though is looking at the physical stability of the protein, which is expressed in the form of aggregation or denaturation. In terms of aggregation, as I alluded to before, you can look at systems analytical ultracentrifugation and simple procedures like SDS-PAGE. On this part, here, this is also a problem that can be monitored by this approach.

This is an example of looking at the purity of a plasma-derived product. In this case, there was a possibility that this protein could actually lose activity because of certain residues which can be oxidized, in particular, if you know the chemistry of the protein and if you know what residues are really related to the activity of your protein, you can kind of specifically look at those.

In our case, for this particular plasma protein, we knew that there was some sort of possibility that it could oxidize and lose its activity. If you look at this protein, we generated a peptide map of the protein, and this is before oxidation and here it is after oxidation. What you see here is only a snapshot of the entire peptide map. It is evident. Here you have an area which is basically two main peaks, and the bottom line here shows you there is something creeping up. In this case, what you had up here is actually starting to disappear. Actually, in this particular procedure we could monitor it in terms of activity as well as in terms of CD spectroscopy and reverse phase HPLC.

The other example in terms of looking at purity, as I said before, we have technologies such as FACE where you can actually look at the carbohydrate components of your product, in which case you label them; you glycosylate it and then you look at the labeled parts.

One very easy and cheap approach of looking at your is actually a combination of chromatography and light scattering detector, using the light scattering detector. In this case, we decided to look what would happen if you really heat your protein with an aggregating agent. Obviously, the mono peak, which is sitting here, completely goes away to an aggregate peak and you could actually devise this procedure.

Why do I care about this? Simply because I would like to know whether the components that we normally call dimers, trimers or aggregates adverse event discrete components of the original material and whether they can be monitored very carefully.

Finally, the other issue in terms of the package, I would like to consider the product potency. Again, this is nothing new for all of us but we do have our own different options that we develop. The only thing that I

would like to really focus on is this ligand and receptor type of binding assays, where we should try to kind of look at the in vitro activity and then try to correlate to the physiological activity.

In terms again of potency, we normally develop this potency or activity measurement under very, very controlled buffer conditions, salt concentration, temperature and excipients. That really takes most of the time when you start development.

But what happens as time goes on? We keep on using the same assay, assuming that it should be good throughout the development process. Oftentimes that is not the case. So, we need to go back and check this approach here. We can also control this by including reference standards, reference standards that we can have either in-house or reference standards from somewhere else.

These activities should also be correlated with some additional information in terms of stress testing, stability-indicating assays and shelf life. That would actually enlarge or narrow our specifications in terms of product potency. It is very hard if we stick to the first assay, the first time we develop it, and we want this assay

to tell us exactly what we want in Phase II, in Phase III and even after clinical material. It is going to be very hard. So, it is really important to go back and say my experience shows so much variation in the activity. In shelf life there is so much variation in activity. So, go back and modify the activity and then come up with new values.

Now conclusions, I would like to submit to you that product characterization is not an end in itself but, rather, a means of identifying the critical parameters required for defining product quality.

Secondly, and very important and every one of the speakers has really hit this point here, recent advances in analytical biotechnology allow one to characterize biologicals to the levels that were previously unattainable. One needs to take advantage of these developments.

I think the issue of biotechnology-derived products as well as plasma-derived products should really flourish. I really focus on protein therapeutics rather than classifying them this way or that way as far as this particular approach is concerned.

Third, these advances apply equally to methods of purification and process control and, as such, the information in product characterization package should reflect product development history, clinical and licensure experience.

Also, such data, and that is why we are here today, is invaluable in properly managing post-licensure comparability activities.

Finally, the concept of "well-characterized" biotechnology products provides a golden frame of reference. If you look at the elements of the characterization strategy that I put forth for plasma-derived products, it is exactly the same topic that has been covered and well characterized. There is a debate as to whether we have to call this process well characterized, substantially characterized or well understood. In the end, there will be no issue whatsoever if we have to focus on the product itself and the approach that we use.

In this particular case, there is one more slide. If we have to look at the elements that we put together before and look at the agency's publication for well-characterized proteins, and then we put against this

plasma-derived proteins, what would be my conclusion? I think my conclusion would be the definition that was given for monoclonal antibodies in what we call well-characterized therapeutic biotechnology products, the monoclonal antibodies were defined at that time as those proteins, the identity of which would be determined by reverse physicochemical, immunochemical characterization without fully knowing its chemical structure I think the biotechnology products, the concept of well-characterized products and the plasma-derived products meet in this particular definition of well-characterized proteins.

In addition, for these monoclonal antibodies one needs to know the purity and impurities. We have covered that aspect, that the purity be identified and the impurities also be quantified.

Thank you for your attention. If you have any questions, I will be very happy to answer them.

Q & A

DR. HAYES: So, we have the opportunity for some questions now. Again, as was mentioned earlier, we need to have everybody that is orally asking some questions come to one of the microphones. We have a portable microphone that

will be going around. Additionally, we are going to try to keep this to ten minutes and cut our coffee break short. So, I would like to have the different speakers that have been here this morning be available for the questions. We have some questions from the audience that we can begin with. Do we have anybody who wants to voice one at the moment?

DR. VAN GEODEREN: I am Hans Ven Geoderen, ZLB Bioplasma. We actually still have pending a submission with CBER where we requested a change--and this is about IGIV--where we tried to change our release test for anti-measles antibodies. We tried to replace the current release test for that by an enzyme immunoassay, ELISA. While doing so, we got into several discussions with CBER and the last one we had was one where we were requested to set up a comparability exercise between the two antigen mixtures that are in the two kits, so the existing test and the ELISA, a comparison comparability testing of the two antigen mixtures. So, this is not about changing the method of manufacture; it is about changing the release test.

Given the fact that this information about kits and which antigens are used is proprietary information which is held by the manufacturer of the kits, to me it is really a sort of stretching it a little bit. Can you comment on this? Have other companies had the same type of experiences with trying to change release tests? And, do you think this is appropriate? I will ask the question to Dr. Chang.

DR. CHANG: I am not exactly sure whether I got your question, but Dr. Basil Golding is probably the proper person to answer that. Let me just give you what I think. For release testing, the agency believes that any change associated with release testing is very, very critical and you have to demonstrate the assay sensitivity and it should be very close, if it is not better. Dr. Timothy Lee will have one example this afternoon to demonstrate how critical small changes in release testing could influence a major impact. Again, I am not sure what exactly--can you rephrase your question?

DR. VAN GEODEREN: Well, we have a set of release tests and specifications. We were trying to get rid of one of them. The tests were anti measles. I believe this test

makes use of monkey erythrocytes. So, we set up a side by side comparison of tens of lots of IGIV, testing with the old test and with the new test. We saw that the new test actually measured more anti-measles antibodies than the old test would do. Therefore, we proposed to raise the specification in order to compensate for that. But, still, we weren't allowed to do it because the concern was that perhaps the new test would measure antibodies that were not clinically relevant, that wouldn't work clinically. To me, that is amazing because, still, I think the clinical efficacy of a product is tested and is demonstrated in clinical studies, and the release testing that you do for these lots in essence should be consistency testing.

DR. GOLDING: I am Basil Golding and, unfortunately, I don't know the details of this case because I wasn't directly involved, but I don't think we can resolve your question here without having the reviewers who were directly involved and have them be able to express their concerns.

But maybe we can use this to just go over a few general principles. For looking at antibodies in a release test, one way is to do binding assays; another way is to do

a neutralization assay. Now, the information that you get from a neutralization assay is probably closer to in vivo efficacy than the binding assay. So, there are different assays and they mean different things. The actual measles assay is one of the very few assays that we ask for as a final release test for IGIV. Originally the test was one of the tests because measles was a much more problem and it was very important to have anti-measles activity in the product. But now the test is more used as a marker of the product and its consistency over time and its comparability to previous similar products. So, that is a critical test, and the changing of the test, as Dr. Chang indicated, is a critical issue.

Now to go into the details of your question I think is not reasonable because I don't have all the details at my fingertips to deal with that and, you know, we can have a conference and set up a time to discuss this in more detail with the reviewers.

DR. HAYES: We need to move on to another question.

DR. JONECKIS: Good morning again. This is a question for Dr. Chang. Thank you very much for your

presentation. It was very pertinent and I think raised a lot of important issues. I just wanted to tell the audience we will make sure to try and get copies of the revised presentation out to them before the end of the day. I am sure many of you were curious about that.

More specifically though, I saw that one of the items you listed was better characterization of starting material, particularly plasma and you listed a number of bullets. I am wondering if you could elaborate a little bit on that, particularly with respect to demographic and racial differences that you mentioned among donors and starting material, and if you could relate that to how that plays in the product characterization or finished product comparability, that would be helpful. Thank you.

DR. CHANG: First, Chris, thank you for the comment, that nice comment you made. The concept that we are talking about, comparability, if the study material started differently, then the downstream process when you compare them with the same process when you use a different study material, it is very likely that you will see some differences in the process downstream. Now, can we identify some critical elements that we can characterize to

provide better means later when you use a comparability study? When you compare a plasma-derived product to a recombinant product you know that the recombinant product, the study material, is a cell substrate which is well controlled. You know what you started with. But that is not the case for plasma-derivatives.

I mentioned several possibilities that contribute to variations, such as the source plasma was recovered plasma; pool size. I am not very clear whether or not for the same product, the same company, do they always start with the same pool size. My understanding is that they may not. See, the different pool size, with so many donors, has variations. In normal production, if we control that it may not be a problem in terms of safety and efficacy, but when you talk about comparability that becomes problematic because you have a different study material. You have the same process. You are expecting some differences. Then the question is whether that difference is significant. Does that difference contribute to the safety and efficacy? So, those are the concerns.

DR. WEINSTEIN: I would also like to make a comment here. I guess, you know, there are studies in the

literature that at least have purported to show that starting plasma that has perhaps undergone some degree of proteolysis in the manufacture of Factor VIII may end up with a product that, after say solvent detergent treatment and heating, resulted in new antigenicity of the Factor VIII product. It has been alleged that it was the quality of the starting material that resulted in a product that had epitopes on it that were seen by the patients as being new entities resulting in a new antigenic site. You know, the issue has been raised whether or not one should monitor starting plasma for degrees of degradation. The issue is that potentially if you start out with a product that is more degraded in some way, the process which would normally give you a product that would not have a new antigen may result in a final product that does have some abnormality.

What I am trying to say is the starting material certainly can have an effect on the end product here, and one should have specifications for that starting material that will lead to a final product that you can have confident will have specified properties.

DR. HAYES: I hate to do it but we need to cut off the questioning at this point. We have used our ten

minutes. I invite everybody to talk with the speakers during the coffee break, which we are going to reduce. Well, we can't be back in ten minutes; let's have a compromise and we will start the session, say, five minutes late and try to be back here at 10:15.

[Brief recess]

DR. WEINSTEIN: We would like to resume now. Our next speaker will be Michael Saunders, from Baxter, who will talk about the industry perspective on preclinical studies.

Preclinical Studies

Industry Perspective

DR. SAUNDERS: Thank you, Dr. Weinstein and Dr. Hayes. I would also like to say thank you and congratulations to the meeting organizers and to the earlier speakers. I think the state has been well set for further discussion of the comparability issues.

This morning I have the opportunity to share with you some views on preclinical studies for comparability. I also submit the following as a subtitle to my presentation because it truly is a clinician's perspective on the use and reliance on preclinical testing for product

characterization and comparability. My background is in clinical research and for the past several years my primary responsibility has been leading the clinical development of the second generation recombinant hemoglobin program for Baxter. This is added on to many more years of previous experience in drug development. So, you will see interspersed throughout my presentation references to recombinant product development to illustrate certain points.

As an outline for my presentation and for purposes of definition and establishing some common ground, I will talk a little bit about background and standards; suggest an approach for comparability preclinical studies; and share some examples to illustrate the points that I am trying to make. I will make a few comments on preclinical testing as it relates to clinical trials, and then make a few concluding remarks.

If you haven't already guessed, I should tell you that I am fairly methodical in nature so, to being with some basics, we know that as demands grow so do manufacturing facilities and capabilities. It is also recognized that process development, enhancements and

changes are inevitable as you gain experience with a process, and as you have additional desires for improvements in economic and efficiency in the process, those changes will occur.

The third point that I am trying to make here is that we are constantly faced with the confounder of biologic variability in our evaluations and their results in our testing and their interpretation. It is also acknowledged that resource and time is required as testing increases as we go up the hierarchy of studies, and I do believe that it is a hierarchy that might also be the complexity of the studies.

For standardization purposes and to try and submit our own idea on what maybe the definition of comparability should be, our group in Boulder, Colorado, the Hemoglobin Therapeutics Group, came up with this definition where we say that a newly manufactured biological product does not have any detrimental differences in identity, strength, quality, purity, potency, safety of effectiveness relative to the pre-change product. The product characteristics can differ between the two products as long as there is no adverse effect on

the above parameters. I recognize that might be a bit provocative, but so it goes.

There are also several important statements I would like to reiterate and emphasize in the 1996 guidelines. I think you have seen this already, but I think it is important to point it out again and paraphrase. There may be manufacturing changes which occur in the development of new biological products but which do not necessarily require the need to conduct clinical trials as long as comparability test data demonstrate to FDA that the product, after manufacturing change, is safe, pure, potent and effective.

A couple of other points, FDA recognizes that it is an important consideration for product comparability is whether or not it is anticipated that any of these manufacturing changes will translate into significant changes in clinical safety or efficacy. FDA has also stated that depending on the type of in vitro assays and animal studies and the quality of the data, extensive clinical data demonstrating equivalence may not be necessary. I am hammering home a point that you will see develop as my presentation goes that by prospectively and

carefully creating the preclinical development, you can actually avoid the need for comparability clinical trials.

The types of process changes are multiple and their significance varies greatly. We have categorized these into the major, moderate and minor categories. Without going through each one of these in detail, I would just point out that changes in process, purification, manufacturing and formulation would be considered major. Duplicate processes changes in equipment and testing site would be moderate process changes. Minor changes would include things like analytical modifications, tightening specifications, etc. The most important point here is that the assessment of this categorization impacts the extent of testing necessary to demonstrate product characteristics and comparability.

To begin with, the approach to comparability testing with preclinical studies involves also several categories. Broadly, these are analytical testing, bioassays and then, importantly, preclinical and animal studies. Analytical testing, primarily chemical and physical assays to characterize the structure, identity, consistency, purity, stability and solution properties, and

may also have some bearing on safety parameters. The bioassays are functional tests to evaluate the activity, potency, integrity, mechanism of action and may also have some predictability for human biological effects. Finally, the animal studies looking at pharmacokinetics, pharmacodynamics and toxicity, it is important to recognize, and I think everyone agrees, that we may not need to repeat all of the toxicity studies with a process change.

The characterization of these studies involves tests which provide an accurate description of the product, as you have heard from the speakers earlier this morning, and that they are consistent, reliable and predictive of the effects in humans. The important point of emphasis is that there needs to be investigation and evaluation of models to select those which provide the highest degrees of selectivity, sensitivity, specificity, precision and predictability.

Focusing on the preclinical animal studies, over the last several years we have seen a lot of technology advances, with enhancements in analytical procedures, non-invasive hemodynamic monitoring such as impedance

cardiography, as well as in the immunogenicity area where models using transgenic animals, knockout models, human gene replacement models, genetically manipulated models have significantly changed the landscape, and you will be hearing more about that from later speakers.

With respect to the pharmacokinetic studies, we primarily want to look at the time to maximum concentration, the maximum concentration, the integrated area under the curve, the half-life, also all of this done with an understanding of the appropriateness of the use of parallel or crossover design trials.

In pharmacodynamic studies we want to evaluate the pharmacologic and physiologic effects, as well as the overall effectiveness, potency and mechanism of action.

In toxicity studies for comparability, this is built upon the safety profile of the existing product, and it may be significantly influenced by the presence of a narrow therapeutic range or specific safety concerns. Experience with the product and the process change should also lead to an anticipated toxicity profile, as well as recognition of potential toxic impurities, and special consideration for immunogenicity issues.

So, the overall approach can be summarized by saying it is based upon product characterization and anticipated effects of process change which can predict the expected identity, purity, safety and effectiveness as they impact on the post-change product. As I have emphasized, it is important to select the optimal and most appropriate models for the preclinical comparability testing, with the goal to do enough to establish comparability but not necessarily to create any obstacles which unnecessarily delay the development of process changes of important new products.

So, with reference to the previously given definition, the ultimate goal is to demonstrate comparability as no clinically significant adverse experiences between the pre- and post-change product.

Also, we recommend prospective determination of a clinically significant change which generally involves development of a decision tree. Possible examples might be threshold of greater than one grade severity of an adverse effect that is observed; a greater than two standard deviation change of a laboratory result. It is also

recognized that often this needs to be done on a case by case basis.

Let me go to a few examples to try to illustrate some of the points I am raising. Here is one from the literature of a platelet antagonist, XV459, fibrinogen receptor antagonist for glycoprotein 2B3A antagonist, which is in development. It is a potent selective product. The LC-MS assay has been used for quantitation of levels in guinea pig plasma. They have also noticed that there has been similar binding in guinea pigs and humans, as well similar dissociation rate constants, and 50 percent inhibitory concentrations for platelet aggregation assays, therefore, one might conclude that the guinea pig is an appropriate model of PK and distribution studies of this product and, therefore, could be a reasonable model for comparability studies.

The next major area of examples that I would like to explore with you has to do with the area I am most familiar with, and that is the group I have been working with for the past several years in hemoglobin therapeutics. While I will be talking a lot about the recombinant product that we currently have in development, I should also point

out that much of this experience is really built upon a human-derived product, DCLHb that you may be familiar with. It was really the major point of emphasis of the Baxter Hemoglobin Therapeutics group over the past decade.

So, in principle and in action, our group has advocated a strong relationship between preclinical and clinical study development. We have done thorough and exhaustive explorations and evaluations of preclinical models to select those which are optimal. With that in mind, we designed the preclinical studies with clinical input and with an intent to mimic the planned clinical trial settings. The clinical trials are then built upon those results coming out of those preclinical trials, and it helps to direct subsequent study designs and the endpoint selection, with the hope that we can develop correlations between the outcomes of the preclinical studies and the clinical trials to try to make a match. We believe that this groundwork and collaboration allows for predictive comparability assessments.

In order to be able to make some sense of these examples, I need to give you a little bit of background with respect to the recombinant hemoglobin development and

some of the issues that we have encountered. Recombinant-based hemoglobin is to be administered intravenously, and should effectively transport oxygen to tissues. It has the potential of allowing reduction or avoidance of transfusions in the surgical setting, and may also have the potential of being a unique resuscitation fluid in trauma settings.

One of the key issues we encountered with respect to the hemoglobin development was a finding of an effect on nitric oxide. Nitric oxide is an important clinical messenger in the body that, among other things, causes smooth muscle relaxation. We know that the heme moiety of the hemoglobin binds nitric oxide, and this binding leads to unopposed smooth muscle contraction which produces several clinical effects. Among those are arterial smooth muscle contraction resulting in hypertension.

Now, let me step back a moment and characterize the two groups that we are talking about, first and second generation hemoglobin. The first generation hemoglobin we define as those hemoglobins which have a reactivity rate with nitric oxide comparable to native human hemoglobin, as opposed to the second generation hemoglobin where we have

modified the molecule and modified and significantly reduced nitric oxide reactivity.

To bring this to some clinical bearing, we did encounter clinical safety concerns with the first generation hemoglobins that Baxter had in development. As a consequence, we stopped the development of those first generation hemoglobins three and a half years ago--four years ago now--and focused the development on second generation hemoglobin. We identified nine features of the first generation hemoglobin which we desired to change, and we conducted a plethora of biology studies to fully characterize the second generation hemoglobin for an investigational drug status, as well as to distinguish it clearly from the first generation hemoglobin. We utilized cumulative experience, literature and consultations to identify the most appropriate models which were optimal for the biologic evaluations.

With respect to the background that I have just presented, and referring back to the blood pressure effects, we noted that there was increased blood pressure seen with the first generation hemoglobin in man, as well as in animal studies. It represented a safety concern in a

number of patients, and the second hemoglobin was targeted to overcome the blood pressure among other effects by reducing the nitric oxide reactivity.

We put together a ran hemodynamics model where we found that the blood pressure response corresponded to nitric oxide kinetics. I will demonstrate that to you in a moment. We anticipate that it will correspond to the human blood pressure response as well. We believe this is a very predictive model for the blood pressure effect, correlating with the nitric oxide kinetics, therefore, any process changes which should occur with the second generation hemoglobin development that might affect nitric oxide reactivity could be assessed by this model, and we anticipate being able to predict what the blood pressure effects might be.

Here is the demonstration with a series of hemoglobin molecules which were synthesized having a variety of nitric oxide binding kinetics. Here is the binding constant for nitric oxide on the lower axis. On the upper axis is the blood pressure response. You see a linear correlation, that the reduction in nitric oxide reactivity corresponded with a reduction in the blood

pressure response with the increase in mean arterial blood pressure.

The second example that I would like to share with you with respect to the hemoglobin is related to the pharmacokinetics and half-life. Second generation recombinant hemoglobin is a polymerized product to increase the size with the intent to increase half-life. We found that pharmacokinetic determinations in rat model did make a correlation with molecular size of the product.

So, subsequent potential process changes which might affect molecular size distribution could be effectively assessed by this rate PK model, and we are looking forward to demonstrating that these correlate with the human results.

I promised that I would give you a few relevant comments toward clinical trials, and with apologies to Dylan Thomas, I want to make the point that a decision to go into a clinical trial should not be taken capriciously. There are bioethical considerations for either doing or not doing clinical trials. Subjecting patients or volunteers to procedures in an unnecessary or avoidable clinical trial can represent ethical issues.

There are certainly resource and time considerations that go into performing a clinical trial and, therefore, we would advocate reserving clinical trials for those instances where the preclinical investigations fail. When a clinical trial is truly deemed to be necessary, we would utilize those preclinical results as a guide to focus the clinical trial design and endpoints.

So, the answer for the decision to conduct a comparability clinical trial may not be to do a clinical trial. It may, in fact, be to do a better preclinical profile ahead of time. We would urge that we enhance the preclinical testing for selection of the best predictive models, and that should be done in concert with FDA input and collaboration to exhaust all of the preclinical study alternatives.

Finally, I would leave you with the following thought, summarized in this statement: Preclinical testing should be performed to identify the most appropriate sensitive and predictive models for product characterization and for evaluating the effects of process and manufacturing changes in order to avoid the conduct of unnecessary clinical comparability trials. Thank you.

DR. HAYES: I have one housekeeping announcement to make. Because this meeting is being transcribed, you can be looking for that on the CBER web site in the next few weeks.

Our next speaker is Basil Golding, M.D. Dr. Golding will be discussing the relevance of animal modeling in predicting immunogenicity.

Relevance of Animal Modeling in Predicting Immunogenicity

DR. GOLDING: I am going to be talking about the relevance of immunogenicity in animal models and testing for immunogenicity for products.

Just an outline of my talk, I call it a road map--defining the problem. Animal models, are they useful? Differences between the animal and human immune system, and new approaches for testing immunogenicity of protein.

When a product is given, such as a plasma-derived product, and antibodies develop--the main problem we are talking about in terms of immune response is antigenicity; in terms of outcome are antibodies. So, antibodies can interfere with the product's safety and efficacy. So, you may get inhibition of product function and this could be either due to binding of the antibody to the functional

site of the protein, or you can get binding of the antibody to a non-functional site but, because there is increased clearance of the protein and alteration of pharmacokinetics, that will, in effect, reduce the function.

Antibodies can be formed which cross-react with self causing autoimmune reactions. The example given by Christian Eckers probably falls in this category. Then, immune complexes, regarding the EPO effect and the effect of antibodies with EPO, are associated with aplasia. So, immune complexes, in other words, antibodies binding to the product forming a complex, can cause adverse reactions, and these are the type of reactions that can be seen as arthritis or even kidney disease and are similar to serum sickness. In rare instances, but of important clinical effect, is the fact that some proteins that we use can induce IgE antibodies and this can cause allergic reactions and even anaphylaxis. The example that comes to mind is selective IgA deficient individuals who are given immune globulins who can develop anti-IgA and this can result in anaphylaxis.

Animal models have been used in the past. I don't want to give the impression that I don't think there is a place for animal models but they do have limited usefulness because immune responses may be different from humans. In particular, I am going to go into some detail later of this, there are differences at the level of the MHC Class II genes, the TCR repertoire and the antibody repertoire.

This is for the future, and already strides have been made in this direction. It has been possible to take mice and introduce human genes into these mice, and humanize these mice so that they now express some of these human genes and, therefore, their immune systems are beginning to look like human immune systems. So, mice have been made which express three of the four IgG subclasses, and it has been shown that you can generate antibody responses in these mice. So, this may be an important model to use for testing immunogenicity of proteins that are going to be given to humans. In addition, mice have been engineered that express human MHC Class II genes, and I will explain the importance of this in a second.

So, the classical testing of a protein use to see if it had gained some immunogenicity during production was to use rabbits or mice, and depending on your testing system and the product, you may use these animals that have targeted gene deficiencies. For example, if you wanted to test a new Factor VIII, you may want to use mice that are Factor VIII deficient.

The basic protocol was to immunize with a native protein, to immunize with the innovator protein and to ask the question whether the innovator protein generated antibodies that were different from the native protein as judged by absorption profiles to the native protein, and also to ask whether the innovator protein generated antibodies that inhibited action of the protein.

Now, I think there still is a place for this and, at a minimum, even though the immune system of these animals is very different from humans and can in no way predict what type of responses you are going to get in human, the fact that you do get a positive result when you are using an innovator protein would suggest that there is something different in structure in the new protein compared to the old protein.

Just to go through some immunological concepts because what I am going to do in the next few slides is introduce you to some possible testing that could be done in the future, making use of human cells and human responses rather than relying on animal responses, and I must warn you that when I presented it to my colleagues I saw body language that told me that, hey, you're going to be way over, but I think I made a particular effort to change the slide and to try and explain this in a way that would also be digestible by non-immunologists. We will see if I succeeded.

The first point to make is that antibody responses to proteins require T cell help, and that this help is related to MHC Class II expression and T cell receptor repertoire. I will explain this in a moment. Because these genes are different in different animal species and are, in fact, different from one human to another and any outbred species, you cannot predict from a response in one species that there is going to be a response to the same protein in another species. The same goes for humans. One human may respond and another human may not respond.

This is a slide that is very critical to the whole presentation of the next few slides. I indicated to you that there is a critical cell, called the T-helper cell, that is required to respond in order to provide--it is called a T-helper cell because it helps, among other things, antibody production. So, this T-helper cell has on its surface a T cell receptor. This T cell receptor recognizes antigen in the context of MHC antigens. So, there are two cells here that are interacting. One is a T-helper cell and the other cell is an antigen presenting cell. The antigen presenting cell has taken up the antigen from the outside, has processed it, and has expressed a small peptide, usually eight or nine amino acids, in a groove of the MHC Class II and the T cell receptor is seeing this complex.

Now, there is a tremendous diversity in the system to allow us to respond to all the multiple types of antigens we see in the environment. So, the T cell receptor consists of an alpha chain and a beta chain, and each chain is along a gene, the T cell receptor gene, which is very analogous to the immune globulin gene and has many genes. So, it has variable regions, V regions and J

regions, and these can combine in various combinations in the alpha chain and the beta chain to give an alpha/beta chain receptor. The diversity here is mind boggling. In the human it is approximately three million different T cell receptors that could be formed by this molecular rearrangement.

These genes are different in different humans and in different species. So, this T cell receptor repertoire is very different depending on whether you are a human or a rabbit or a mouse. In addition to that, the MHC Class II genes represent in the human at least 100 alleles per gene. Again, you have tremendous diversity over here.

Because of this tremendous diversity, it is very difficult to predict whether a human response is going to occur if you see a response in a mouse or a rabbit. But the question is how could we use our knowledge of this system to devise a better way of screening proteins to see if they are going to be immunogenic or not?

I should just point out that in immune responses the T-helper cell plays a central role, and when it is stimulated by the antigen presenting cell that I showed on a previous slide you get various things going on between

the T-helper cells and the B cells, and then you get antibody production. So for proteins the T-helper cell is critical. Without first activating the T-helper cell you are not going to get antibody.

This is the paper that I am going to be talking about, and I have only seen one paper to this effect but it is logical and I think is something that I would like to draw to your attention because I think more people should try and replicate these results, and maybe this is a mechanism for testing proteins that are going to be given to humans. The basic idea is that you could use naive human T cell responses to predict T cell epitopes in an antigen, in other words, to predict those peptides within an antigen that are going to stimulate the immune system, particularly T cell. If you want to look up the reference--I left out the data, this was probably in the year 2000.

So, this looks like a very complicated slide, and I actually lifted it from Dr. Estell's presentation to the FDA and it is on the FDA web site. Although it looks very complicated, it is actually very simple. You are taking peripheral blood from a patient, you take that blood and you purify dendritic cells, which are the professional

antigen presenting cells, and you also purify the helper cells. You then add the antigen. So, these dendritic cells will process the antigen, present it on the surface in association with MHC Class II and stimulate the CD4 T cells to divide provided, of course, that the T cells contain within their repertoire the T cell receptor that will recognize the peptide that is being presented by the dendritic cell. You can do this in microtiter plates and you can measure T cell stimulation by looking at thymidine uptake, which is a standard assay.

What you can also do once you have seen a response to the total protein, you can make overlapping peptides from the protein and you can actually determine which are the epitopes which stimulated the T cells. You can do this from a large number of human individuals so that you get a good representation of the human population in terms of the MHC Class II that is used and the T cell repertoire.

What they did, they used human peripheral blood mononuclear cells. They used an inactivated enzyme as the protein just for proof of concept, and they looked for proliferation of the T cells, and they called a response a

stimulation index greater than twice background, and a weak response was slightly above background, and no response was same as background.

This is just an example of the type of experiments that they did. I haven't got time to go over the whole paper, but when they took the native protein and used the system they found that over 50 percent of the T cells from normal individuals responded to the protein. So, obviously, this would be a problem if you wanted to give this protein as a treatment. You know, there was a weak response and no response in about 20 percent of the normal individuals.

What they did, they identified, using peptides, which part of the protein was actually inducing the T cell response. Then they genetically engineered the protein so that those peptides, based on all kinds of algorithms that relate to the binding of the peptides to the MHC and so on, they could change them by just doing point mutations, and then test the protein again in the in vitro system to see if now the protein had lost its immunogenicity.

What you see, for example, in this variant is that now most of the individuals are no longer responding

to the protein by making a single mutation in the protein. So, this approach not only allows you to detect whether a protein is going to induce an immune response, but it allows you to engineer the protein in subtle ways so that it will no longer be immunogenic. Obviously, this has to be don e very carefully and even single point mutations can have a drastic effect on protein folding and glycosylation, and all kinds of things. So, all those other concerns have to be looked at before making the changes.

What they were also able to do, and this is getting back to the situation where you could actually take a mouse and humanize the mouse and use that to test your protein, so if you had a protein that you wanted to use in patients and you did those previous studies in human T cells, and you knew from the human responses in vitro which DR alleles were involved in the response--the high responders were people that expressed these alleles--you could then make a mouse that expresses the human HMC Class II, and some people have already done this using different human alleles, and test in vivo in the mouse whether that mouse responds at the T cell level, and whether that mouse makes antibodies to the protein.

In summary, this epitope mapping assay determines relevant priming epitopes. The epitopes can be modified to reduce immunogenetic potential. HLA-D restriction patterns direct transgenic animal mode choice, and modified enzyme variants are hypoimmunogenic in vitro and in vivo, as I showed you earlier.

So, general conclusions from my talk, animal models have limited usefulness but I want to repeat that I am not saying that animal models have no usefulness. One just has to be aware of the limitations of what you can learn from animal studies. I still think that if you take a rabbit and inject it with your new protein and you get an antibody response but you didn't get that response before, you need to be concerned that this is going to be immunogenic.

Promising new tests--and I am not saying I endorse this approach but what I trying to say that we need new approaches because what we now know about the human immune system and animal immune systems, we know that animal models are not an adequate model for testing or predicting human immune responses. And, this is an example of an approach that may be useful in the future. They use

primary human CD4 T cell responses. They also use transgenic mice which are expressing human HMC Class II genes. What I didn't go into but I did mention is that there are available transgenic mice which express human IgG subclass antibodies 1, 2 and 4.

Clinical studies are required until these methods can be validated. So, if you want to embark on this approach, methods have to be correlated to human immune responses. In the future it may be possible to engineer changes in products so that they are less immunogenic. Thank you.

DR. WEINSTEIN: Our next speaker will be Martin Green, from the FDA, who will talk about comparability studies for human therapeutics, preclinical and pharmacokinetic aspects.

Comparability Studies for Human Therapeutics, Preclinical and Pharmacokinetic Aspects

DR. GREEN: Good morning. I am Dave Green. I am from the clinical pharmacology and toxicology group. Before I start my talk, let me tell you a little bit about our role in this process because I think the talk reflects how we fit into comparability testing. We are in the

Office of Therapeutics but we act as a resource to the other offices, including the Office of Blood.

As far as comparability testing goes, we kind of see the exceptions rather than the rules. As you heard earlier, there are many changes that are proposed in manufacturing and in other circumstances that cause comparability to be an issue. Generally, we see very few of these studies and we see them only after there has been a question raised as to whether there are more simple physicochemical determinations or other easier methods have been looked at and found wanting in terms of resolving comparability and, therefore, more testing, usually some kind of animal testing is thought desirable either because the answer can't be obtained from the data that is available or there is an important question that needs more rigorous testing.

One other thing I wanted to point out is that oftentimes we are brought into the process when time is very important to sponsors, and within the constraints of their idea of developing their product and getting approval or providing their product to the market, they want absolute assurance oftentimes that the test that is agreed

upon with the agency will be acceptable to answer the question. That tends to drive the types of tests to the highest standard because there is no allowance in that setting to provide a chance for making an error, and that is that the study that has been perhaps less rigorous is inconclusive or doesn't provide the assurances necessary to allow FDA to conclude that the materials are comparable.

I would like to say that there is perhaps a different way of looking at the why for comparability, and that is, it is intended to ensure that there won't be a mistake that we will regret; that there is a continuity that the preceding information that was gained, either through the toxicology or clinical trials and also, secondarily, will preclude the introduction of new, unevaluated factors that will change the product's characteristics and safety profile.

One other issue that I want to mention in passing is that although we have had a lot of comments and a lot of sponsors want to dwell on the performance aspects of their products, there is the characteristics of identity and sometimes it turns out that at the end of all the analyses the best standard for establishing identity is the human

performance standard and that can be done through the pharmacokinetics or the determinants, but it is hard sometimes to put the pieces of the puzzle together to know whether identity has been assured so that its characteristics are conserved as the product is released for use.

One thing about comparability is that it varies in terms of rigor for which those questions need to be addressed. That has been mentioned by other speakers. As products progress in their development and as they come closer and closer to marketing, and after they are marketed, the consequences of making a mistake become more and more severe and the opportunity to address those mistakes diminishes as well. So, some product changes that might have some standards applied early in the process, there is an ability to correct for change as they go along. As the opportunity for correction decreases as we get closer to marketing and into the marketing situation, again, the standards have to be more rigorous oftentimes because the allowance for mistakes is much less and the consequences of making mistakes are much greater.

So, comparability studies from the perspective of clinical pharmacology and toxicology are often considered when the chemistry manufacturing controls are limited, or when they don't provide answers to the questions. Another type of general category for comparability testing is when the clinical study, and on this slide I mean efficacy determination, is either not feasible or inappropriate. In other words, comparability study provides simply another alternative to repeating the Phase III study which established the product's suitability for market in terms of efficacy and safety.

As other speakers have suggested or indicated, there is no absolute formula because of the various stages and complexities of these issues and also, again, because the nature and type of comparability testing varies with the phase and type of product.

Again, a different perspective on what is an appropriate comparability test is what is the issue that we want to resolve. What is the particular concern that we have? Is it one of safety? Is it one of identity? Is it new introduction of material? Also, how precisely do we have to answer this question? Will a more broad answer,

one that just tells us that there is nothing new about this product suffice, or do we have to get very close to understanding that this is essentially exactly the same as it was before?

Basically, there are two types of comparability issues that one has to look at. One is the broad exploratory issue in which we want to capture perhaps unanticipated, new or previously identified effects? In other words, oftentimes a toxicology study gets into this. That is, we don't know with this change has a consequence for safety so it is an unanticipated, perhaps new change that we haven't been able to put in context because we have no prior experience to know what the consequence is, or there might be a previously identified effect, such as the nitrite oxide example that we just heard about.

The other broad category besides the exploratory type is the verification type and many PK studies or pharmacokinetic studies fall into this area. So, there is a continuity that is going to be based on prior experience that is maybe historically controlled. Then, another type of verification, it may confirm a previously unrecognized attribute. That is, perhaps during the course of the

clinical study something arose that now tells that there is a characteristic of this molecule that we want to make sure is very important, and simply to verify it. But the two types of studies are basically broad-brush or very narrow-focused.

Aside from their intent, there are basically three characteristics or three types of studies, and I am sure you are all familiar with that. Those are the pharmacodynamic studies. Those tend to be not very common. They tend to be very blunt in terms of the precision; they provide estimates although sometimes there are important characteristics. An important aspect of taking a pharmacodynamic study is that its relationship to the clinical efficacy is often speculative. That is, we don't know for certain whether that is the case or not. In rare instances we do have a very good understanding of what the impact is in terms of pharmacodynamics on the clinical situation, and in those rare instances we have used pharmacodynamic studies to establish comparability.

A very popular aspect is the pharmacokinetic studies. Now, pharmacokinetics have a very important attribute to them, and that is for systemically

administered compounds we know that that is a critical aspect of getting that molecule to its sites of action. It necessarily must go through some pharmacokinetic phase. Whereas pharmacodynamically we are uncertain as to its relationship overall to the clinical effect, we know as imperfect as the pharmacokinetics may be, as the imprecision we may have in terms of measuring and correlating it to clinical effects, we know that it is a necessary component of that.

An intermediary between the more rare pharmacodynamics and the more popular pharmacokinetics is the toxicity study. Generally, that is, as I mentioned, oftentimes in the mode of exploratory studies ordinarily in the area of safety but toxicity studies are intermediate in terms of numbers.

For all these studies we need to consider some practical aspects. That is, how many animals we need including the number of human beings we need. The experimental design is parallel and crossover, as mentioned earlier. These boil down to certain characteristics that are generally recognized. It is important about making decisions about whether a change has occurred, the

reliability, robustness and variability, but generally the standard that we like, no matter what, is side by side comparisons.

An example of a toxicology study that can be informative is taken from the literature, by Blecker, in which they looked at 16 IVIG products and they were interested in characterizing the hypotensive effect which they attributed to dime content. Essentially, they showed about half of them had no hypotensive effect, which they categorized as less than 50 percent change in blood pressure, and 8 did have a hypotensive effect, which could be fairly large, 50 percent. These were done in rats, I believe.

They attributed the effect, as I mentioned, to dimers or the platelet activating factor. But if we have some suspicion that there was a difference from any number of factors, aggregation or dimer content, this would be a suitable way of addressing this issue, particularly if we had evidence in the clinical situation that hypotensive effects could be attributable to the product.

Again, triggers for asking for comparability studies using animals are that the analytical insight is

usually limited or not comprehensive, or the aggregate number of changes is more than we can understand in terms of its total impact. So, having a whole series of changes, even though each one is minor and even though each one individually may have insight in terms of whether it should be comparable or not, in aggregate they may have interacted in some way which is unpredictable and that, itself, would be a trigger for doing comparability studies.

Another general trigger is that we cannot determine whether a potential change is clinically significant or not. Either the physicochemical determination is insufficient in sensitivity, that is, it has its limitations, either instrumentation through chemistry does not address the feature of interest, or it can be also that there is a failing in understanding the proportionality between changes and clinical effects.

As has been mentioned by some of the speakers, some of the take-home lessons that we have about changes which are considered important are changes in degree of aggregation, changes in glycosylation, changes in charge, and I have an example that will follow, and also changes in the percentage of isoforms that exist within a product.

Another example taken from the literature is comparison of Factor IX being plasma-derived versus recombinant in which they found that there was about a 30 percent initial lowering of recovery, that correlates to something similar to area under the curve, for the recombinant Factor IX product.

Prior to doing the studies, maybe some people would have understood that some of the physicochemical determination revealed that there was likely to be a change, but I think in general many people would have suspected that there wouldn't have been a change. Not knowing the impact of this but realizing that it had a potential effect, it was important to do the study. So, it was realized that there were post-translational modifications but they were thought to be sufficiently similar. It was later realized that the N-glycosylation was more complex than they suspected and that there was a change in the degree of charge through phosphorylation which may have accounted for this difference in area under the curve.

Clinical studies, and by this I mean pharmacokinetic studies and limited pharmacodynamic studies

and not meaning here efficacy studies as would be found in Phase III, are commonly used to establish comparability. They certainly have the redeeming virtue that they are the best model. It is a model that everybody can agree will provide the answer to the question if they are powered right and if they have the right design.

These studies can involve patients or healthy subjects, and I will talk about some of the determinants whether they are conducted in one population or the other, and they may not be feasible or appropriate in some instances. That is, there may be lesser standards which equally provide the information that answers the question. Again, the problem oftentimes is that we don't have enough information to tell us which among the alternatives provides us an equally good answer, and we the need to get this reviewed or with the compresses time frame to get this to market, usually the agreement is on the most strict standard rather than entering into a process of exploration with the possibility of not coming up with an answer.

So, what are some of the reasons for not doing this in patients in a limited number of subjects who would be the best example of the population to which the product

would be given? Well, there may be too few patients available within a reasonable time to actually do the study. Or, the variability between patients may be too great. That is, the disease burden may be a factor in their pharmacokinetics and that may vary very greatly between subjects, and there may be an influence of concomitant medications. Even the variability within a patient may preclude doing a small clinical study for pharmacodynamic endpoints because either antibodies may be preexisting to the product or similar enough to be cross-reactive, or antibodies may develop as the product is given to these people. As I mentioned, the disease burden may be a factor sometimes in the disposition of these products and that may introduce a factor of heterogeneity, such that even though we might use a patient population which is on a continuum of disease, this population which has enough patients and has less variability does not really represent the population at risk.

So, if we choose to use healthy subjects because they give us reliable endpoints and they are more generally available, there are some considerations which prevent their use. Oftentimes it is a safety factor. That is, the

material itself might have inherent safety issues which don't make it a worthwhile risk for normal subjects, otherwise healthy. There could be the influences on them of future therapies. For example, if they were given a material such as monoclonal murine-derived monoclonal antibody and there was a high proportion of people developing antibodies to it, they would be essentially disfranchised should the need arise in the future for that kind of therapy. Lastly, there are some clinical routes which cannot be used in normal subjects.

With regard to animal studies, there are strengths and weaknesses and we will be going back and forth on those. The animal models that are established can be used. Even lower order animals such as rodents versus non-human primates can in some instances provide good answers in terms of pharmacokinetics if we know certain factors about the regulating underlying disposition of the pharmacokinetics. That is, it primarily reflects a non-specific uptake. The material primarily goes to organs such as the liver and its uptake is primarily dependent on blood flow rather than specific aspects such as amino 6 phosphate receptors.

Oftentimes we can understand this because we can go back and decide whether the animal is a good representative of the pharmacokinetics if the proper studies have been done because we have the clinical studies and we have the preceding animal toxicology, toxicokinetics or pharmacokinetic studies to compare them. Where the non-human animal is a good model for a human being, essentially it is dependent on the PK profile that is presented to us, and whether there are significant departures from the PK profile between those two groups, and whether those represent significant differences. Animal studies don't have a problem with recruitment. Instrumentation is possible. They do allow specialized designs and the monoclonal antibody document talks about a very elegant design, which has been very serviceable over the years, of dual labeling materials using the animal essentially as its own control in the same time.

Limitations are some of those just the other side of the advantages. Instrumentation introduces confounding factors, particularly anesthesia. Where there is complicated instrumentation or there are limitations on laboratory scheduling it can be an issue and, therefore,

you get heterogeneity or variability introduced of time. There are different technicians in different facilities and there can be a physical limitation of staff and facilities.

Animal studies can also have limitations for pharmacokinetics where the assay may not be directly taken from the human studies or there can be matrix effects. Using small animals, there is particularly the problem of limited blood sampling. Animals also have the disadvantage for human proteins, they may develop antibodies and, therefore, we are looking at some cases of steady state kinetics. This may be a factor and invalidate the results. But the overriding factor for using non-humans is that there is a degree of extrapolation; there is a leap of faith that has to be made between the results you get in non-human animals and that which you get in people. Sometimes that is a wide gap to go through.

This slide was mostly meant to talk about non-human primates. The non-human primates have the advantage oftentimes that we can understand a lot of the physiology so immunosuppression--we can do things to these non-human primates in a way which resembles human factors and have a lot of confidence about what its relationship is to people.

Sometimes we can standardize their disease burden if that is an important factor, but oftentimes we can look at some particular organs of interest. If we are looking at deposition uptake into the liver, into the brain, into the kidney, into the lung we have an ability to get at those tissues and look at primarily what are the consequences of the pharmacokinetics. Sometimes, but rarely, we have a chance to do that in clinical studies as well.

Limitations for using non-human primates, although there are relevant animal models, their availability is increasingly diminishing as we are going. In some species, such as the chimpanzee, which are a protected animal species, the utilization is extremely limited.

Another disadvantage of using some of these larger animals is that the test substance itself may be limited, and also that there can be unexpected reactions to some of the materials, binders or incipients or formulation factors which don't present problems to people but do present them to the animals.

Other people have commented on the availability of the guidance and I do not need to talk about that too

much, but I do want to talk about some particular points in the comparability document. Overall, I see it as an interrelated system of tests which are complementary, as other speakers have talked about. It specifically states that animal PK studies may be needed in the absence of demonstrated differences in analytical testing or functional assay. So, it may be needed regardless of whether the test showed there is an effect because they may not be comprehensive in terms of providing us assurance that there won't be important differences. Changes in the finial product formulation may need comparative pharmacokinetic studies. This is frequently a point of concern when sponsors do change formulations, particularly towards the end of development. Finally, final product formulations may need pharmacokinetic studies, and I think by that it is meant clinical studies pharmacokinetic studies. With that, I will conclude and thank you very much.

DR. HAYES: I want to thank the speakers for actually making up time in this session. So, we are actually back ahead of schedule. Our next speaker is Charles Maplethorpe, from the FDA, on the FDA's perspective

on the design of clinical studies to evaluate comparability.

FDA Perspective on the Design of Clinical Studies to Evaluate Comparability

DR. MAPLETHORPE: Good morning. My name is Charles Maplethorpe. I am from the Clinical Review Branch in the Office of Blood. I have been asked to say a few words about clinical studies to demonstrate comparability.

This is actually a very specific topic that would relate to whatever product found its way into this predicament of having to show comparability at the clinical efficacy level. So, instead of saying specifics about the trial design, I am going to talk about my point of view on the general process of showing comparability between products as a result of a manufacturing change.

It is appropriate that this meeting is taking place in 2002 because this is the 100 year anniversary of the Biologics Control Act which gave rise to CBER and, therefore, FDA. This is the first of the laws that provides the regulatory authority for the FDA to regulate drug development. It is interesting to read the law because you can learn a lot about the historical

perspective of regulation back then, and also to see the differences that evolved between biologics and other products such as devices or drugs.

One of the interesting things about this law, which is not very long, is that it says that products can be licensed if the sponsor can show that standards are used to ensure, "the continued safety, purity and efficacy of the product," with emphasis on the word "continued" demonstration. So, there seems to have been a recognition in the law that biologic products were, you know, a necessary safe and effective thing and this law actually came about when a St. Louis manufacturer, which was, of course, before FDA existed, chose to institute what he might have considered a change as being effected immediately, namely, he used the serum from a dying horse that made diphtheria antitoxin to make his final batch of the product from that horse. It seemed good enough to him but, unfortunately, the horse was dying of tetanus and a number of children in St. Louis died and, as a result of that, we have this law and CBER.

So, from the beginning biologics has been interested in manufacturing processes, and the original

regulatory system that was set up at that time has continued to this date, and that is that laboratory scientists are very intimately involved in the regulation of biologic products. Given this 100-year history and involvement of laboratory scientists in the regulation of these products, we have accumulated quite a lot of experience in knowing how these products can go wrong, and knowing what sort of things to look for and how to measure them. Therefore, we can tell quite a bit about the changes that have taken place in a product based on laboratory tests.

This law introduces the concept of potency without explaining it, at least through regulation. So, this is the regulation that defines potency. It says, the word potency is interpreted to mean the specific ability or capacity of the product, as indicated by appropriate laboratory tests or by adequately controlled clinical data obtained through administration of the product through the manner intended, to effect a given result.

So, if you look at this, you can see that there is an assumption here that it is possible to do some sort of laboratory test or an adequately controlled clinical

trial, and there seems to be an implication that you can do an in vitro test that reflects something about the clinical efficacy of the product. That may or may not be true, depending on the product and to what extent we understand what is going on in terms of the clinical efficacy of the product. It also shows that the potency test can be tied to the indication. It says, "given through the manner intended," or tied to the indication. So, one product can have a different potency assay depending on the indication.

This slide shows the general sequence of studies that one would go through. We have been talking about them all morning. It has the same resemblance that you would see for the studies in general product development for licensure, but there is a big difference here. That is, for the preclinical studies that are used in product development one of the major concerns is whether or not the product is safe enough to use in human beings. By the time you have licensed a product and are looking now at product comparability, you have the basic assumption that the product is safe. It is not the same as when you are doing product development.

So, what you are basically doing here is looking as closely as you can at each stage and asking the question can I stop? You know, if I have done my manufacturing change and I compared it to some of the very sensitive techniques we heard about earlier and found no differences, is it reasonable to stop and say that the product is comparable? If there are differences, can you be enlightened by doing some sort of animal model, efficacy maybe safety, or some sort of in vitro bioassay looking at activities? Then you ask yourself the question can you stop there.

Frequently, especially if there are differences, there can be a concern if the PK, if the human PK is the same and you may be asked to do clinical pharmacokinetics. We have heard that it is actually quite rare to be asked to do a clinical demonstration of efficacy after a manufacturing change, and if you find yourself in that situation it should essentially be that you failed somewhere earlier in this chain.

So, repeating the requirement for a clinical demonstrating of comparability could depend on several things. It could depend on the failure to demonstrate

comparability by preclinical or PK methods. It could depend on the clinical correlation and the on-site validation of the potency assay. In other words, how strongly do we believe that the potency assay truly does reflect clinical outcomes? The second part of that relates to what extent have you, the sponsor, demonstrated that you can perform this potency assay in a competent manner.

The third point is regulatory past experience with the product indication. We have a large number of products under IND. Some of them are very traditional products that go back the full 100 years of our history, and there is quite a lot of experience there and a given change--we have seen it before; we know what to expect. There are other products that are brand-new recombinant products or even traditional biologic products being used in new and different ways and in those cases we might have higher standards to apply if you make a manufacturing change. At any rate, the decision to require demonstration of clinical comparability is made at the preclinical review level after finding insufficient information to permit a declaration of comparability.

What standard do we use if you find yourself having to do a clinical trial for comparability? We use this equivalence range, which I think is derived from generic products, the standard that they use for showing the products are equivalent. The point estimate and the confidence intervals should fall within this range of 0.8 to 1.25 for a given parameter, usually an efficacy parameter, because the tolerance range would be strongly dependent on what the adverse event was.

When you make a manufacturing change and you want to tell the world that you have the same product, well, you have every intent that the true point estimate is, in fact, going to fall at 1.0. After the trial is over and you come to us with the results, sometimes the focus is on the other end of the range, the 0.8 to 1.25, and we have even been asked on occasion to stretch this. We are told, look, this is just an arbitrary interval that you set up here to show equivalence and the true range should be extended. Well, you know, this is actually quite a wide range here and we have licensed products on smaller differences than this. So, we would really want to hear a strong argument, if you came to us after you had done your clinical study and

failed, for why you would expect to be considered equivalent because you had already, presumably, failed quite a number of steps to get to the clinical trial and now we are being asked to change things again.

That is because this is sort of the nightmare scenario that we have in the back of our minds, and that is, what this slide is supposed to show is how bad things could get if you failed at the 0.8 or 1.25 levels after only four different manufacturing changes. This is where you would compare 2 to 1, 3 to 2 and 4 to 2. You can see that after only four changes you could have half the efficacy and twice as many adverse events. This is just a nightmare scenario.

So, what we can see is that a clinical trial for comparability is actually a very crude and expensive instrument to use to detect small product differences and if you see a difference, it is quite a reason to be concerned. But if your purpose is to ask the question do I see a difference between these two products, a clinical trial is not the best way to do it. These preclinical methods that we have been talking about are much better for doing that.

The problem when you get into the statistical design of such a trial is that an equivalence trial could end up being larger than the original trial that was done for product licensure. So, you really have to say what are we doing here? Is it worth it?

That brings us to the final slide, which is a question, does the failure to demonstrate clinical comparability rule out a demonstration of clinical safety and efficacy? The standards for licensure are actually quite low. All you have to show is that it is effective. Would a de novo demonstration of safety and efficacy be a better route to licensure? We have realized that there are all sorts of other business considerations that go into this idea of wanting to be the same product after you do your manufacturing change, but in your planning you should also consider the possibility that it is going to be a lot more trouble showing comparability and maybe, in fact, you should just go the new product route. I guess that is it.

DR. WEINSTEIN: We have time now for some questions from the audience.

Q & A

DR. GEIGERT: My name is John Geigert. This morning Dr. Chang told us, and it was iterated just now, that for comparability clinical studies are quite rare, one or two percent. Could anyone comment on his list of about 400 comparability or process changes over the last five years, what percentage also required preclinical? We know that obviously it would be 100 percent for the analytical and one percent for clinical. What percent would be for the preclinical?

DR. WEINSTEIN: I don't have that figure. We haven't analyzed the data to get that information. It would be interesting to know.

DR. CHANG: Well, the simple answer is that I don't have that. It is difficult to capture that information. I think that is something we should establish in the future in the agency, to capture that information. For clinical data supplement we can easily capture it because under PDUFA 2 we have ten months review time as compared to the PDUFA 2 four months review time without clinical data. So, you can do a search to get that information. But for the preclinical data, I am not aware that we have a system.

DR. GREEN: I guess the comment I would have it I think it is important to classify the preclinical data in terms of the type of studies that were done, for example, were they toxicology studies; were they potency studies? I guess there is also a question, which is probably really the one that you would like to have answered, is were they important to the approval or were they done by the sponsor and were considered something good to do but not necessarily other things where a physicochemical determination would have been sufficient in itself. I think it is unclear. I think there is a tendency to imply a standard because the sponsor may be successful in getting their applications approved when they provide, in some instances, a rich abundance of information that is way beyond the threshold and that somehow that becomes now what people perceive as the threshold when it actually isn't.

DR. LYNCH: Tom Lynch. I was also struck by Andrew's statistics but I wonder if they are not sort of self-fulfilling, i.e., that manufacturing changes that would require clinical data are simply just not made because they are too expensive.

DR. CHANG: I think that is probably true and industry people know this better I think, that they may not decide to do a comparability study if they find that significant differences are present between the old and new. As Dr. Maplethorpe pointed out, the comparability clinical trials are even more difficult in terms of design and expense for that. So, that is possible.

PARTICIPANT: Can I follow-up? That is actually a central issue here. I think everybody would agree that a rigorous statistical demonstration that two products are, for all intents and purposes, identical is a very, very difficult trial to do, with many more patients than the original trial to establish safety and efficacy of a new product. So, that logically puts a major manufacturing change, where an issue like any of the list that Dr. Weinstein showed at the beginning, right back into the new product category and seems to undermine the whole idea of comparability. So, it seems like that is the dilemma. I mean, if you are going to handle all major manufacturing changes because you can't answer questions about immunogenicity or bioavailability necessary in an

analytical or preclinical model, you are going to handle everything as a brand-new product, would you not?

DR. WEINSTEIN: The fact of the matter is that I don't know that we have that many brand-new products. We have a number of BLAs but those are relatively few so I think that most of the changes that we have seen are being handled at the preclinical or analytical stages here. You could say that this is actually beneficial because the data that we have has convinced us that we haven't had to go to the end stage of the clinical trial. We have received sufficient information in many submissions to convince us that, in fact, the change was not that great and that products were comparable as determined at the earlier stage.

DR. MAPLETHORPE: I would just like to say that if you have a licensed product and you have a manufacturing change and you, in fact, have made differences, measurable differences in either the safety or efficacy profile I don't think we, at FDA, see any need to preserve the product identity. If that is a new product, so be it. We don't feel the same need that I think the sponsor might feel to preserve the product in the new formulation.

DR. GREEN: One thing I believe is a perception is that there are going to be many, many changes which don't register in people's mind because they are basically ordinary, but when we do ask for new studies it is usually unanticipated by sponsors and really causes lots of resources to be expended, both time-wise and in terms of patients, and it usually impacts much more strongly. So, I think even though there may be a very few number of them, they stand out very strongly against a background of all the other changes which basically go unnoticed.

Usually a lot of the changes that are done for very good reasons, to produce a better product, and I don't think we disagree that we want a better product. The question is whether aspects, for example, of the labeling in terms of adverse event reporting and the therapeutic benefit retain the same characteristics. Sometimes it is very difficult because when sponsors make a change which makes it essentially, in the view of the FDA, a different product, they want to continue with the identity of the product and they also sometimes present the complication of wanting both products on the market at the same time, which means that post-marketing adverse event reporting now gets

confounded, and even switching from one to the other where a patient has been stabilized can now be a very complicated affair and there is no provision for managing that with the physician or the patient.

DR. ZEID: Actually, you bring up a very good point, which is what I was going to ask. In the absence of demonstrating comparability by just sheer analytical and pharmacokinetic, and there is a question mark over whether additional clinical data are needed, I would submit that a lot of these products are individually titrated for the patient depending upon their state of disease status or other comorbid factors. Sometimes where there is this question mark the innovator may be obliged to alter their labeling or modify their labeling to reflect these changes. So, it gives them some flexibility through their labeling without having to go to the next step which is clinical testing.

My real question was in a classic scenario where and IGIV or recombinant Factor VIII product is tested, usually a sponsor will do, just for their own benefit, a system-wide analysis of where their product stands in structural characterization to the field of players that

are out there. If I have made a significant manufacturing change and now I see some distinct differences between my pre- and post-product but I see comparable activity or comparable profiles in products that are done by other innovators, isn't it possible that I could make that stretch of logic that the changes I am observing in my post-modified product are not representing major safety and efficacy issues because I can point to other examples of other products with similar profiles?

DR. MAPLETHORPE: Well, I mean, we all have a sense of IGIV as being a fairly good, you know, long-term product that we feel quite comfortable with, but if you take your general concept and apply it to other products, let's say like antithymosite globulins which I also happen to regulate and have quite a few under IND, for that product category, they vary all over the map in terms of individual characteristics, and if a manufacturer came in and said, oh, guess what, I'm going to reduce the RBC absorption by half because this other product has a titer down there, since the manufacturing processes are completely different I don't know how that is going to affect the clinical outcomes in that product category.

So, you are asking a specific question about a specific product and you have the assumption that we know enough that we can be comfortable with the outcome. I would have to look at that individual case before I could give an opinion, but in general I don't think you can give a general answer that would apply across all biologic categories.

PARTICIPANT: I have a question for Dr. Saunders and Dr. Joneckis and I would hope they could both comment on it, and I would just like to hear more about how can preclinical pharmacodynamic studies contribute to the establishment of comparability.

DR. SAUNDERS: I think the example that I used was the hemodynamic model, and we are looking at that as being a critical factor for defining the recombinant hemoglobin. I used that as an illustration. So, when we see significant changes in blood pressure that are beyond a range, and whatever that range is has to be predetermined, a prospective determination, we use that as a potential model for comparability. We are saying that that is an essential feature of what we have engineered into the molecule. Again, this is a somewhat artificial situation

or a different situation because we are talking about a recombinant product as opposed to a plasma protein and I am not sure, as Dr. Maplethorpe said, that I could generalize to all of the possibilities but I am just saying that in certain instances where you identify a characteristic of a product that you have developed and specifically targeted, and you can establish with experience what a range ought to be and if you make some changes and maybe with respect to the hemoglobin it affects the nitric oxide binding kinetics, you may end up with a change that falls outside the range. So, theoretically to be comparable, it stays within that range.

DR. GREEN: Let me attempt to answer for Chris. Your question is whether PD markers have ever been used for comparability and how could you determine whether the case would be. I think the flow of logic is PD is very important where we understand its relevance to the clinic, and where it provides sometimes a better estimate than other estimates such as pharmacokinetics. Typically it is reverse, but there are rare instances, as I mentioned in my talk, where PD is a better endpoint and a preferable endpoint. The one example, the hypothetical example I gave

to you is a neurotoxin which might be used for dystonia and so forth, and where we understand exactly what it does with a great deal of assurance.

Now, it is kind of artificial in a way because if you give something essentially to paralyze a muscle it is essentially a local injection so there is no PK to collect. So, it is another reason to look at the PD part, but we understand its relationship to the clinical efficacy and its performance, that is, its muscle paralysis is a clinically relevant endpoint and it can be measured very precisely in terms of a number of factors. It can also have the advantage that if given in an isolated way, just like the monoclonal antibody where you can use dual labeling and each animal serves as its own control contemporaneously, you might be able to use it on different similar muscles because of the bilateral symmetry of animals to determine whether the product is basically comparable. So, that is one instance of a PD marker.

DR. GOLDING: Can I just add to that? In terms of the immune globulins it is very difficult, in my view, to do any PK studies in animals because if you take a human antibody and stick it in another animal the PK profile is

very different from what you would see in the human and everything is shortened. The same thing would occur with products that are human in nature and cause antibody responses in animals. The antibodies themselves that were made in animals may not be made in humans would obviously interfere with the PK. So, there are I think severe restrictions in some of our ability to use animals to study PK for some of our products.

PARTICIPANT: Just to kind of follow-up on that, a general question, we had a couple of examples of where classic toxicology was very relevant but those examples were drawn from bioengineered products, recombinant products, things like that. Can the panel comment on the value of more traditional toxicology for the older plasma derivatives that are used generally as replacement therapies, like the Factor VIIIs or the IGIVs? Would one who is planning marketing a new albumin, for example, do a toxicology study in two rodents? I mean, would that have value?

DR. MAPLETHORPE: I think we saw one example for IGIV that I presented where hypotension was the toxicology endpoint. We could split hairs between whether that is a

safety pharmacology study or whether that falls within the purview of traditional toxicology studies, but the toxicology endpoint was informative perhaps about the amount of dimers that are involved. But if you had something like an albumin and you were going to change it and wanted to know if they were the same, some animal toxicology studies might be useful if they were focused on particular endpoints that you might be concerned about, for example, deposition in various tissues because you are worried about association of the molecules, one to the other, and you want to know the content in lung tissue.

So, this can be to determine the characteristics of it, but it might be one that would tell you that the disposition, in a general sense, falls within more classical toxicology studies. But oftentimes I think that toxicology studies in the traditional sense are intended where there are formulation factors which are now put in there whose toxicology is really unknown. It may be that there is sufficient toxicology known about these formulation factors independent of it, but maybe there is something about the combination of the two. Perhaps the product in some ways acts as a carrier, or it somehow

modifies its expression. So there, again, can be some reason to considering toxicology studies.

I think oftentimes my experience has been that they are more times done by sponsors in anticipation of whatever constraints they have than they are as a formal request by the agency. The number of times that I have personally requested toxicology studies and the times that they are presented by sponsors, my requests are a minority of them.

We have thought about toxicology studies for immunogenicity, not so much to know whether things are immunogenic but in rare instances where we want to know whether there is a quantitative shift in the likelihood in rate of development. So, even though there are vastly different immune profiles in terms of the underlying immunology, we still felt in some rare instances that there was some idea of knowing whether the rate rise of antibody development, even across species, was greater for one product than for another because there was a safety issue. I think that is an important perspective. Again, what are the concerns that we have now? What are the concerns we need to preclude? How accurately do we need to preclude

them to make a reasonable and prudent judgment that this is a worthwhile thing to do, particularly where we don't have opportunities to study it under clinical investigations? So, when people give us a change in their whatever post-marketing, we are not anticipating clinical trials but we still have, as the sponsor does as well, a responsibility to ensure that we have thought about in the past carries over in the future, and it is not considered a good thing just to not address it in a way that establishes the fact rather than hoping for the best.

PARTICIPANT: If you have multiple indications for the same product, if you prove clinical efficacy in one case must you do all other clinical indications?

DR. GOLDING: Sometimes yes. It depends on what those indications are and what is the mechanism. For example, immune globulins are used for a variety of conditions, used to replace antibodies in people who are primary immunodeficient, but they are also used for Kawasaki's disease, and they are also used for immune thrombocytopenic puerpera. So, in many of these conditions the precise mechanism of action of the antibody is not known and, of course, we are dealing with a product that is

very heterogeneous. If it works in immune deficiency, in the past we have asked each manufacturer to do an efficacy study for their own product to show efficacy for the other indications if they wanted that indication on the label. But if you are dealing with a product that is a single molecular entity, I think you may be able to make some extrapolations.

DR. MAPLETHORPE: I agree with that. Your question is very product specific. For example, hepatitis B immune globulin--there is a set of "needle stick" indications and if you get one you would get them all. But other products that are being used for different indications, as I said, could have completely different potency assays. So, it is not a given that just because you get one you would get the other.

PARTICIPANT: Actually, I have two questions. The first question is do any of you have any experience with a product where the two products showed equivalence and both stayed on the market, either under one brand name or two different brand names? They are equivalent, so both products were allowed to stay on the market?

DR. MAPLETHORPE: Both made by one manufacturer?

PARTICIPANT: Right. And, what would drive that? What would drive the decision to keep both of those products on the market as opposed to pulling one off if they were equivalent, both safety and efficacy?

DR. CHANG: Well, one point I made during my presentation was that a manufacturer volunteered to phase out their old processes. But there is a transition period where both products--I should say the product made before and after changes are on the market for a transition period. I think there is a good reason to permit that practice.

Both sides, the industry and also the agency, do our best to avoid any product that is not effective or not safe on the market. But some of the safety issues are learned later in the post-marketing phase when more patients use that product. So, to give a transition period is not a bad thing. The company also needs that time to make that transition, especially for those products that are in short supply. So, basically the answer is yes.

PARTICIPANT: Thank you. One other question, in addition to a pure equivalence design of a study, could you comment on the use of, say, a non-inferiority design or a

not worse than design in order to show equivalence between two products?

DR. MAPLETHORPE: I think you would have to ask that of one of our statisticians, but usually a sponsor, if they are making a manufacturing change, wants to say that they have the same product. You know, they want to say that they are comparable; we are going to use the same name, the same ads, etc. That would fit the paradigm that I showed. If they did a trial to show that they were better --

PARTICIPANT: Not worse than, not better than because that would denote a different product, or would it? If you showed that you were better than, would that be considered a new product or something that was equivalent?

DR. MAPLETHORPE: I think a lot of that would be up to the sponsor. I think if they showed that suddenly they had a product that didn't, you know, transmit some viral disease they would want to change the name.

PARTICIPANT: Yes, that is obvious.

DR. GREEN: One comment about non-inferiority is that I think there is less experience in the realm of comparability, although there is more experience in the

realm of approval. I think one aspect of it is if something is no worse than or there is that issue of the lower limit but also the upper limit of exposure, and that would be that there were no adverse events that were associated with something that basically we didn't care how much people got. I think that often is not the case because of safety concerns and also because of just product availability in some cases. I mean, there are other problems about packaging, how much is in the container that is used. So, although I think it is an approach that people might consider, it is not one that I know of that has been used.

PARTICIPANT: Thank you.

DR. GEIGERT: As we have been talking about these plasma proteins, it seems to me there are two classes. There are, of course, those that are plasma-derived and then those that are bioengineered or specified or genetic engineered. In terms of comparability, does the FDA treat them in any way different? I haven't heard anything yet today that says that in terms of comparability there are different criteria, different concerns, different issues that would be addressed whether it was derived directly

from the plasmid or whether it actually went through a bioengineering or genetic engineering approach. Could anyone comment on that from the FDA, if there is a difference that I haven't picked up yet? And I apologize if I haven't.

DR. CHANG: Again, we come back to a case by case situation. Let's say a specified product, we may treat that differently if it is a different product based upon the knowledge that we have and the industry has of that product. If you are asking whether there are significant change differences between plasma derivatives and a recombinant product, I will say that we have different concerns, as I listed in my presentation. Those concerns will make some inference about the judgment. If you say that is a difference, yes, I agree that is a difference between the biotech product and plasma derivatives. In some cases we may have even more concern about the biotech product if, for instance, the cell substrate in the host cell protein is a foreign protein as compared to some of the proteins derived from a plasma derivative of human origin.

So, we do have different concerns and in terms of an approach, in terms of policy, this morning we heard many, many times that we have the 1996 guidance on comparability. That comparability did not make a distinction between the biotech and naturally derived product. So, we have the same policy as used before for all biological products.

DR. GROSS: This is Mike Gross. In the spirit of looking to constantly validate the comparability concept, has FDA experienced situations where perhaps the comparability concept has led to a scenario that was undesirable, perhaps adverse events that were not anticipated on the basis of either physicochemical characterization or animal models but perhaps not clinical studies?

DR. GOLDING: In the case studies, I am going to be presenting some information related to that, and I don't think this is the time to go into any detail, where manufacturing changes were made and they were not anticipated to cause any effect and, in fact, they were thought would improve the product and, yet, were associated with adverse events, if that is what you were asking.

DR. GROSS: I am familiar with some of them but I am saying where a comparability approach was taken, perhaps a comparability protocol was developed, agreed upon with FDA, engaged and then ultimately perhaps some things were discovered that were not anticipated.

DR. GOLDING: Yes, I think it is quite a common occurrence. I am not sure I am at liberty to tell you those stories.

DR. MAPLETHORPE: I am not sure if we have answered this or not, it says is there any application for comparability studies/protocols for products from two different manufacturers/sponsors? That is to say, can a manufacturer of a new product compare with another manufacturer's product?

It looks like the question is can a new product be licensed by doing an equivalence study to a licensed product, and the answer is yes. In other words, do you have to do a placebo-controlled study? Well, there are many situations where you can do a placebo-controlled study so I guess the answer would be yes.

DR. JONECKIS: Let me try to clarify that a little bit. The comparability document in 1996, as clearly

indicated, is for changes within a manufacturer and that is how it is applied. Now, maybe I will interpret that question somewhat differently to say that if you want to make some type of claim, a comparative claim, for example, regarding other competitor's product that would have to be designed into your clinical trial in terms of evaluating any type of claim that one would want to make. But the comparability concept, the guidance, the comparability protocol is not meant to show changes or do comparisons between manufacturers; just within a manufacturing scenario. So, that is really the focus.

DR. MAPLETHORPE: Well, my remarks pertain to a demonstration of clinical equivalence whereas, obviously, the comparability document conveys the notion that you can show it preclinically, which I was not implying.

DR. GREEN: I think maybe the intent of the question, which has the same answer, was that if, for example, you had three or four manufacturers making a very similar product and one had two percent aggregates, can you use one manufacturer's standard as the standard for everybody and say that, well, now we are at one percent, let's say, so we should be okay. We went from one and a

half percent to one percent. You already have a product on the market that has two percent, therefore, here is a generic standard, and the answer is no.

DR. HAYES: This is for Andrew or Chris. Can the comparability for a manufacturing change be demonstrated at pilot scale, and is approval possible on the basis of this data?

DR. CHANG: Well, as Dr. Joneckis has pointed out, the comparability study is actually embedded in the whole product development, or actually Dr. Joneckis pointed out after Phase I. The comparability study can be used from the Phase I clinical phase to post-approval. Now, if you said can the comparability approach be used for the pilot facility, I am not exactly sure what that question implies. Let me say that if you have a pilot facility, you make the material for a clinical trial and you have another facility which is a production facility for marketing, the comparability approach can be used here to demonstrate your testing article used for a clinical trial manufactured at a pilot scale is comparable to the commercial product, and the answer is, yes, you can use a comparability approach to demonstrate that your

commercial product is comparable to the testing article used in the clinical trial manufactured at a pilot scale. Did I answer the question? l

DR. SEAVER: That wasn't my question. My interpretation of the question is different. You are manufacturing at full scale. You want to do a process change that you are going to implement at full scale. Can you demonstrate comparability only using lots made at pilot scale?

DR. JONECKIS: That is the question, Sally. I actually happened to speak to this individual during the break, and the answer is, no, you can use that as supporting data but you still have to evaluate material coming out of the actual proposed commercial scale.

DR. HAYES: We have some other questions here that came from the audience. How many lots, multiple lots would be for trending acceptance criteria for pre- and, most importantly, post-change lots?

DR. JONECKIS: I am not sure what the question means but if it refers to a remark I made during my statement about trending of lots made pre- and post-change can be a sensitive indicator of many things, there is no

formal requirement that that be done so there is no particular number. Clearly, the more manufacturing experience one has with the pre- and the post-material being made from the pre-change and the post-change product gives one additional confidence and you can look at various trending statistical models to see what type of confidence you can get out of those values. So, there is no formal requirement that that be done and there are no formal numbers. When it has been done, at least in what I have seen submitted, it has usually been that for the pre-change material all the relevant available lots have been trended. For the post-change material, it is whatever has been made. It is typically greater than three but it has been, again, for the relevant post-change material. I have seen eight, ten, things of that nature to get some kind of sense as to what is going on.

DR. GEIGERT: Maybe a follow-up, that would be true, Chris, I would assume for a market-approved product but for clinical, because we are covering the whole realm of early phase as well as all the way through to post-approval, you may have one of the post-approval for clinical. So, just we don't get an expectation that

everyone at Phase I is doing nine runs every time they make a process change.

My other question is it has been five years now since this 1996 document came out. Is FDA more comfortable now with this policy, less comfortable? We have had inferences that maybe our approach of comparability is creating some concern that we are not catching everything. I am trying to get a read from the FDA here where do you think it is going? Is it going to get tighter? Is it going to get loosened more because you have more experience? Or, are your experiences giving you more concern?

DR. GREEN: Well, I think one easy answer is that the issue of immunogenicity has certainly taken more prominence and has become a more focused concern. I think that with some of the other issues, such as pharmacokinetics, both the FDA and industry have had some experience now for a while, those issues are now a part of how people think about the problem and know the type of approaches that they need to look at. But I think immunogenicity is a new aspect that both sides are grappling with, and in the future other aspects may come

about. I think you have to see this against the background of the trend in biologic products in general, which is that they are now getting a more broad patient population. If you look at the entire spectrum of the diseases that they are treating, they are more chronic in nature and they are entering into areas where they are not only used for life-threatening, serious morbidity as they have in the past. Also, the nature of the industry is changing as people are trying to improve yields and compress the time for their own development as many of these issues provide innovations which aren't anticipated at this point.

DR. GOLDING: Can I just add that I would underscore what Dr. Green has said, and what has struck me--as you know, I am looking at it from the perspective of heterogeneous products derived from source plasma so the starting material is very complex; the potential of contaminants being in the final product are much greater. But then when you look at the bioengineered products that are recombinant proteins that are so well characterized and you have all these fancy biochemical and physical techniques to look at them, yet, you hear story after story of very subtle changes that are missed by all these

techniques and are associated with immunogenicity or other problems, and if you are seeing it in that area, then in our area where we are dealing with heterogeneous proteins with a much more complex manufacturing set up, I think it should make us be a little bit more cautious about what is going on and whether we can use analytical methods to characterize plasma-derived products to the point where we can say, well, this is enough; we now have looked at the physicochemical characteristics and we now know everything we need to know about this product and it is safe and effective.

DR. JONECKIS: Let me just make one other additional comment. I sort of alluded to in the beginning that we are taking a loot at comparability across the Center and across our products and, in fact, we are because we have had approximately six years of experience. Are people more comfortable? Well, I think it depends upon whom you ask. For some of our product classes, I think the experience has been relatively minimal and to some extent I think people are not terribly comfortable. I think we have had a lot of experience with other type of product classes.

Although we have had some very successful stories with comparability and no real nightmares, fortunately, with a product getting on the market that has caused some untoward side effects, as I tried to say in my talk, we are surprised about certain things and we continue to be surprised. So, we are going to continue to discuss internally where we are going with this. In part, I think we are also being challenged by many things, some of which have been reflected, the need to meet market supply and demand. Despite some extensive changes that we have seen, especially post-approval, people are starting to push the envelope more, I would say, in terms of those types of changes and a lot of this has precipitated, as our experience has precipitated, the need to sit down and say, okay, where are we going with this policy? What is our experience? What can we conclude and not conclude? We will hopefully come up with some type of guidance at some point in time. I am not exactly sure in what form it will be, but that is a discussion that is still ongoing and is, as I said, across the offices and at the Center level with all the various offices participating in these discussions.

DR. GREEN: I think the biggest grievance that I have heard from sponsors about the comparability document is their disagreement as to what they think a change implies in terms of testing. Usually they feel that it requires very little and, in some instances, we feel it requires quite a bit more. I think at the heart of this there is certainly the interpretation of the words of the comparability document, but I think more essential is what was alluded to, the base of the experience. I think this is a problem for us communicating to sponsors because a lot of the findings that we have we can't make public.

So, we gather a great deal of experience which look like isolated experiences but are really systematic throughout the process. Oftentimes they are surprising because the unusual things which, unfortunately, are shared by many processes give rise to effects that people don't anticipate as having profound effects and they do. Yet, they occur with a single sponsor and a single product in that period of time and there is no way we can say, at least directly say, okay, this change necessitates this effect.

DR. MAPLETHORPE: One of the things that surprises me in the cases where differences are only detected at the clinical level is the extent to which, when you go back and look at the file and look at the preclinical studies, the finding was already there and it was either dismissed, not noticed and, for whatever reason, you know, they went full steam ahead and then the project stops clinically.

So, that is one of the messages that we have tried to put out and I think Dr. Saunders has tried to put out, that you are not losing time or money when you spend that on preclinical studies if you trust your data.

DR. SAUNDERS: I agree with that.

DR. HAYES: I believe that concludes our question and answer session at this point in time. Please join me in thanking the speakers from this morning. We also encourage you to continue these discussions over lunch, which is across the hallway. We are to reconvene at 1:30 for the next session.

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

- - -

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

DR. FINKBOHNER: All right, good afternoon, everyone. I think everybody is settling in. I would like to welcome you to the afternoon session of the conference today. A couple of quick housekeeping announcements, first I will introduce myself, I am John Finkbohner. I am the Branch Chief in the Division of Manufacturing and Product Quality at CBER and I will be giving a presentation a little bit later in the session today. The co-moderator for this session is Christopher Healey, sitting to my right.

I have been asked to announce that the conference is being tape recorded and transcripts will be available through the CBER web site. It is an awfully long web site address to read to you, so you can check the main CBER web site or you could also ask for additional information from the Freedom of Information staff. The phone number is 301-827-2000. The transcripts will be available in approximately 15 working days.

Also, since this is being tape recorded, I would like to make just a general announcement that under 21 CFR

10.85(k) we have a nice regulation that says that verbal opinions expressed by FDA employees are not binding upon the agency--

[Laughter]

--so what you hear us talking about today will be our best judgment based upon our experience.

DR. HEALEY: John, that takes effect after the announcement, right? Everything up to this point--

[Laughter]

DR. FINKBOHNER: I will have to defer that to our general counsel for interpretation! Also, one last request, if you have questions, if you could give us your name and affiliation as you are posing questions to the panel this afternoon.

The session this afternoon is titled reporting manufacturing changes, and our first speaker will be giving the FDA perspective on changes to be effected/ changes to be effected in 30 days and pre-approval supplement categorizations, and that is Dr. Timothy Lee.

Reporting Manufacturing Changes

Changes to be Effected (CBE)/Changes to be Effected in 30 Days (CBE-30) and Pre-Approval

Supplements (PAS)

DR. LEE: Good afternoon. In the next 25 minutes I would like to just go through the topic that I would like to cover, basically some introduction to the reporting categories and also describe a little bit about CBER's approach to the categorization, brief description of the categories, and also I would like to provide a couple of examples to illustrate our approach to categorization and also illustration of the complications and complexities involved in reviewing.

Basically, actually I will reiterate some of the points that have already been made in this morning's session by the various speakers in comparability studies in the context of reporting manufacturing changes in the different categories.

I have included in a handout more details information on this topic for your reference. In the interest of time, I will skip over most of it actually and actually some of the material will be covered by the other speakers in this session in the afternoon.

In reporting manufacturing changes to an approved application, there are three reporting categories depending

on the potential for the change to have an adverse effect on the identity, strength, quality, purity and potency that relate to the safety or efficacy of the product. The manufacturing changes could include production processes, quality controls, equipment, facilities and responsible personnel.

Just to set the stage, I think you are probably quite familiar with this, that reporting categories are prior approval supplement, or PAS, changes being effected in 30 days and changes being effected where the 30-day waiting period is waived. Also, there are changes that are reportable in annual reports. The comparability protocol will be submitted, and I think Dr. John Finkbohner, Paul Gil and Frank Rauschen will cover that in more detail.

The approaches to categorization, basically I want to emphasize that the standard used to assure the safety and the efficacy of the product is the same among all the reporting categories. Also, to categorize the various manufacturing changes, we draw from our experience and our knowledge of the product and the process and the experience of those manufacturing changes as they are related to the safety and effectiveness of the product. We

also depend on risk analysis and assessment of those changes to the product.

The first category is the prior approval supplement. That covers major manufacturing changes that have substantial potential to have an adverse effect on the safety or effectiveness of the product. The products that are made using change can only be distributed upon approval of the supplement. The review schedule is four months and, as Dr. Chang mentioned a little earlier this morning, if a human clinical trial is required, then a ten-month review period will apply.

Here I have listed some of the examples that are for manufacturing changes that are reportable under PAS. This list is drawn from the guidance and is, by no means, exhaustive or inclusive but it does cover a wide spectrum of changes that will give us an idea of how those changes are being categorized. So, the process changes can include introduction of new or revised recovery procedures; new or revised purification procedures which might include a change in a column; change in chemistry or formulation or solution used in processing; and also change in the sequence of processing steps or addition, deletion, or

substitution of a process step. Those will also be reportable in a PAS. Also, if there is an additional viral inactivation step, those are also under the PAS category.

Basically, any change in the manufacturing processes or analytical methods that result in changes or specification limits or modification in potency, sensitivity, specificity or purity to establish a new analytical method, those are the changes that have to be reported under PAS. In terms of the stability protocol, if there is elimination of steps or alteration of the criteria, those are also PAS reportable.

Some more specific examples on PAS are scale-ups requiring larger purification equipment; change of manufacturing sites, with they be an addition of a new location or contracting of a manufacturing step to be performed at a separate facility; also, any changes in the location or major construction, those are also reportable under PAS.

The next category is change being effected in 30 days, or CBE-30 supplement, and those cover the changes that have moderate potential to have an adverse effect on the safety and effectiveness of the product.

A PAS could be, might be lowered to a CBE-30 if the proposed change has been validated using an approved comparability protocol. Subsequent speakers will cover that.

Under this category, the product made using the change may be distributed 30 days after receipt of the supplement. Within 30 days FDA will decide whether the submission is complete or the category is correct, and then the review will continue. Under PDUFA 2 the review schedule is six months.

So, during the review period, if FDA finds that the data fail to demonstrate continued safety and effectiveness of product we will try to resolve the problem with the applicant. If that fails, FDA might require the applicant to cease distribution of the product or to remove the product from distribution. So, it is very important to point out that in situations like this, it is the responsibility of the applicant to determine whether they should be prepared to release the product 30 days after submission of the supplement, recognizing that the release may be delayed because of the deficiencies in the supplement, or make other arrangements to better

accommodate such a possibility. The risk of distributing product also applies to other manufacturing changes that are reported under CB or annual reports.

When we have experience with certain particular changes, based on that experience we can decide whether for some changes the 30-day period can be waived, and those are reported under the CBE supplement category. Under this category, the product using the change may be distributed upon receipt of the supplement, and the review will continue, and the review period is six months. Again, the responsibility lies with the applicant whether they determine to release the product under such conditions.

I have also included in the handout a list of examples where manufacturing changes can be reported under CBE-30. Here, what I would like to do is to give one example to show that sometimes the changes might be complicated by the intrinsic mechanism of action of the reagent that is involved. In this particular CBE-30 supplement the applicant applies to extend the storage time of some buffers used to regenerate the chromatographic columns. One of the regeneration buffers that is used is a

rather concentrated urea solution which is used to regenerate a protein affinity column.

The sponsor would like to extend the period about two and a half times, and the buffer is slightly alkaline. From the literature and experience, alkaline urea will break down to a cyanate which will react with the amino groups in the protein, whether it be the lysine residue or the primary amine at the N-terminus, and carbamylation of proteins is known to change the property.

So, with this change, even though there might not be an immediate effect on the function of the column, we have to also consider that there might be long-term effects on the property of the protein ligand and, in turn, there may be potential adverse effects on the purity of the product. So, in this situation we had to ask the sponsor to address this issue in the context of validating the continual yield of the column over the period of time that they have reported. So, those are some of the issues that we have to look at when looking at some changes which on the surface may not be that significant.

The next category is the annual report. The annual report covers minor changes that have minimal

potential to have an adverse efficacy on the safety or effectiveness of the product. Product made using change might be distributed as soon as the change is implemented. These minor changes are reported within six day of the anniversary of the approval date of the product. The sponsor may also request in writing alternative reporting dates.

Under PDUFA 2, the FDA will determine within the first three months whether or not those manufacturing changes are appropriate, and then the review will continue until six months. If the FDA considers that changes need to be reviewed as a supplement, FDA will notify the sponsor to submit a supplement. If the FDA finds the product poses danger or issues are unresolved, we may require the sponsor to cease distribution of product or to remove the product from distribution. So, again I reiterate that it is the responsibility of the applicant to determine whether they should release the product or not.

In a list of examples for manufacturing changes reportable under the annual report, one of the examples deals with the modifications in analytical procedures with no change in the basic test methodology or existing release

specifications provided the change is supported by validation data. What I would like to do in the next minute or two is to give you an example that fits this particular category, but we have to look at the data more critically to evaluate its effect on the safety of the product.

This particular manufacturing change deals with a modification to an assay to determine the concentration of pre-kallikrein activator in the product. The changes, the modifications in the assay deal with changes in preparing the sample, preparation of the sample, the volume of the sample that is tested, and also the change in the range of the standard curve. There is no change in the basic methodology or the specification of the assay.

The sponsor internally classified this change as minor and as an annual reportable change. Using the modification, they have test released several lots of the product. Also, they have submitted release protocols to the FDA for official lot release. Based on the data that are generated by the modified assay, those lots are released. The sponsor also submitted a CBE-30. So, when we looked at the changes, one of the changes is in a buffer

and when we looked at the changes in the buffer we realized that the changes were not that minor. There are some changes in the pH and a stabilizer was removed in the modified assay, and then, most significantly, there is a significant change in the concentration of the salt that was used.

My colleague, Dr. John Finlayson, did some calculations and he concluded that the ionic strength is about three times higher in the modified assay. Also, based on his work in 1983, he pointed out that this particular enzyme is very sensitive to ion exchange.

So, we got a little concerned and the concern gets even higher when we look at the consequences of those changes. Using the licensed assay, the percentage of lots that meet the specifications is about 25 percent. When the company made the change in this assay all of the lots they submitted passed. So, because this is related to the concentration of the PKA in the assay, elevated levels of PKA result in hypotension and pose a significant safety concern. We contacted the sponsor and the sponsor stopped release of the lots which failed the licensed assay. On

further discussion with the sponsor, we came to realize that maybe the inconsistency is not with the assay itself.

Basically, I just wanted to use this example to show that when we do the characterization of the changes, in addition to looking just on the surface just at what the change is, we have to look at the data more critically and trust that the data is telling us something. When the data looks extraordinary, it probably means that we have to pay close attention to.

Basically, this is what I have to say on the categories. In summary, I just wanted to say that in our experience with looking at manufacturing changes we are constantly and continually learning from our experience with the product and the process. And, I just wanted to say again that when we get the data we need to look at the data very critically to evaluate the potential for change and adverse effects on identity, strength, quality, purity and potency as they may be related to the safety and effectiveness of the product. It would be helpful to consult the guidance documents often, and the FDA encourages the sponsors to communicate with us. I think whenever you have questions, it is good to communicate with

us so we can iron out those issues at a very early phase of the development.

Before I end, I would like to thank the people who have given me a lot of comments in terms of this particular presentation, folks at FDA, Dr. John Finlayson did some calculations and pointed to the significance of ion exchange in the PKA assay. Then, I would like to end by conveying something from one of the contemporary writers. Everyone quoted is not alive but at least Garrison Keillor is alive--

[Laughter]

--so it is a little bit different from everyone else: Be well; do good work and keep in touch. Thank you for your attention.

DR. HEALEY: Thank you, Dr. Lee. Moving on to our next speaker, it is Dr. Frank Rauschen, from Bayer, and he will give the industry perspective on changes.

Industry Perspective

DR. RAUSCHEN: Good afternoon. I would like to thank the organizational committee of this workshop for giving me the opportunity to provide an industry perspective and experiences for prior approval changes

being effected in 30 days and changes being effected supplements.

I will start my talk with some basic thoughts on post-approval changes and regulatory affairs overall in the handling and management of those changes. I will then go into the different supplement times and conclude with some industry experiences.

As Dr. Lee mentioned in the previous presentation, FDA has defined several general categories of post-approval changes that require a supplement if your information established in the approved application is changed, and this information can be found in 21 CFR 601.12. The changes include changes in product, production process, quality control equipment, facilities, responsible personnel and labeling.

Post-approval changes are a necessity and cannot be avoided. They are needed for continued improvement and industry's objective to ensure high quality procedure in accordance with current industry standards. I have provided a typical list of post-approval manufacturing changes which include improvements in product quality and safety, equipment, facility upgrades and modernization,

staying in compliance with CGMP requirements, and also improvements in logistics and supply. Those may include capacity increase to ensure adequate supply for the patient.

Regulatory affairs has a critical role in the management of change implementation, typically is consulted at an early stage. Regulatory affairs is a department that determines if a change may require a license supplement. If a submission is required, regulatory affairs will provide the submission strategy. It is also expected from regulatory affairs to predict accurately the submission requirements. This includes the content of the supplement, the reporting mechanism, the time line for the product distribution and the time line for approval. The overall objective is to achieve a short regulatory time line.

However, an accurate prediction of submission requirements is even more important. If the regulatory assessment is incorrect, the company may face delays in the change implementation and even the potential to impact product supply. So, as a company you need to have reliable information about when and how the change can be implemented. As a company, we need reliable information

that the planned changes can be implemented at the same time, under the same supplement type and, even more important, as a company you want to know and minimize the impact on product distribution and supply status.

FDA has issued several reference documents for industry specifically for post-approval changes that provide industry with guidance on post-approval changes, and lists specific examples for how to report those changes. That reference can be found in 601.12 and also in the guidance for industry documents that FDA has issued for specified biotech products, for biological products and also for whole blood, blood components and source plasma just recently, in 2001. FDA has also provided even more reference documents and a list is shown on this slide.

I now want to go into more detailed discussion of the different supplements, starting with the prior approval supplement. As Dr. Lee pointed out, major post-approval changes require the submission of a prior approval supplement prior to product distribution. As he also pointed out, the time for approval and broad distribution is, at the minimum, four months if no clinical data is required. However, deficiency or complete response letter

that you may receive from FDA may delay the implementation of the post-approval change and any product distribution time line significantly. Actually, the approval time line and the product distribution may become somewhat unpredictable. Comparability protocols may be an option to downgrade prior approval supplements to a CBE-30 and, thereby, reduce the time line for product distribution significantly.

This is a list of typical major changes that require a prior approval supplement. Among those changes you find changes for aseptic and filling and process area. You find new production scale. You also have, as Dr. Lee pointed out, comparability protocols.

As I mentioned before comparability protocols may be a strategic approach to downgrade prior approval supplements to a CBE-30 and, thereby, reduce the time for product distribution. They can overall minimize the regulatory impact on product supply. Some successful examples that use the comparability protocol to downgrade the prior approval supplement to a CBE-30 include changes to aseptic filling areas, new production scale and a new W5 distribution system.

Post-approval changes may be submitted under changes being effected in 30 days supplement. Different from a prior approval supplement, this type of supplement does not require FDA approval prior to product distribution. However, industry has to observe a 30-day waiting period. CBE-30s provide a significant time advantage over prior approval supplements in terms of the ability for a company to release and distribute product. The time line for approval of the supplement, however, may be up to six months or even longer.

However, there are some risks associated with changes being effected supplements in 30 days. Here is a list of risks that are associated with CBE-30s. For changes on a critical path the company should confirm the clock start and the expiration date of the CBE-30 with FDA. But this approach will confirm that the submission was filed under the correct submission type, on one hand and, on the other hand, it will also confirm that the company's interpretation and the FDA's interpretation in terms of the 30-day window are in agreement.

A potential risk for CBE-30 supplements, as Dr. Lee pointed out, is that supplements are not typically not

approved but product has been distributed after 30 days prior to the approval of the supplement.

This again is a list of moderate changes that qualify to be submitted under a CBE-30. It includes changes in non-sterile processing areas, and also introduction of unlicensed product into a licensed production area.

Comparability protocols may also be used to downgrade the changes being effected in a 30-day supplement to a CBE supplement and, thereby, even waive the 30-day waiting period. For the most part, comparability protocols are not often used to downgrade CBE-30s for one-time changes. The time advantage of 30 days is not significant enough compared to the effort of going to a prior approval supplement for the comparability protocol first.

However, this approach may be used in situations when a single change occurs multiple times. For example, a contract manufacturer who wants to introduce on a regular basis an unlicensed product into a licensed facility, instead of submitting each time the changes being effected in a 30-day supplement and having to wait 30 days prior to the start of product distribution, an approved

comparability protocol could actually allow to downgrade a the CB-30 to a CB supplement and the contract manufacturer would be in a position to start release and distribution immediately. Those are two successful examples that use the comparability protocol to downgrade the CB-30 to a CB supplement.

For some moderate changes FDA has waived the 30-day waiting period which leads to the CBE supplement, or sometimes called CB immediate supplement. Overall, there are only very few examples in the FDA guidance documents for post-approval manufacturing changes that are qualified to be submitted as a CBE supplement.

Risks associated with the CBE supplement are identical to the ones outlined for the changes being effected in a 30-day supplement. Again, a potential risk is the product is typically distributed before the change is approved.

Again, this is a list of post-approval changes which were submitted under CBE supplement. As I indicated before, only very few moderate post-approval changes qualify to be submitted as a CBE supplement.

Coming to the last part of my presentation, I provide some industry experiences for supplements and compare protocols. As I mentioned in the beginning, accurate prediction of regulatory requirements is critical to minimize the impact of post-approval changes on product supply. However, experience has shown that an accurate prediction may sometimes be difficult. The reasons for that may be the level of detail provided in FDA guidance documents which vary, and has led to different interpretations of reporting requirements between industry and FDA. As a result, supplements could be down or upgraded after the submission was filed. In particular, if the submission is upgraded the company may be impacted in its ability to release product.

Experiences with comparability protocols indicate that CBs have become a well-accepted strategy to downgrade major facility and equipment changes. CBs are also used to obtain confirmation from FDA that the chosen approach and also the supplement content is acceptable to FDA. The decrease in reporting requirements is typically one tier. However, you should keep in mind that substantial revisions to an approved comparability protocol or any reference SOP

or method in the reference in this comparability protocol requires a separate prior approval supplement.

In some instances, and we have heard this in several speeches this morning, comparability protocols are used to downgrade the requirement for a clinical trial, and a typical example would be if a product change is combined with a facility change. So, you might be able to downgrade the need for clinical studies with a comparability protocol, but for the facility portion a prior approval supplement and pre-approval inspection may still apply.

To facilitate approval of comparability protocols you should consider the following format and content details. You should provide specific changes or a list of changes. You should always include relevant SOPs and methods. You should provide a commitment to report all deviations and investigations, and you also should provide a commitment that minor revisions to method and SOPs are included and explained and justified with the submission of the supplement for the manufacturing change. You should specify detailed acceptance criteria, and for process changes you will be required to provide criteria and a

decision tree that trigger additional preclinical and clinical studies.

Coming to the end of my presentation, I will just summarize what I just talked about. Post-approval changes, as I mentioned before, are a necessity and regulatory affairs play a criteria role for the implementation of post-approval changes. Comparability protocols have become instrumental to minimize the impact of critical changes on product release ability.

At the end, I will leave you with two recommendations for discussions, I believe that the FDA guidance, and in particular the update of existing guidances, industry's ability is strengthened to accurately predict reporting mechanisms for post-approval changes. I also believe that guidances on the content and form of comparability protocols will help industry and facilitate the approval of comparability protocols and meet FDA's expectations. Thank you very much for your attention.

DR. FINKBOHNER: At this time we have about ten minutes, I guess, reserved for Q&A. Feel free to stand up, wave your hand or use the lavaliere, otherwise, if you have questions that you have jotted down on a 3 X 5 card,

somebody will come through and pick those up and we can handle them up here. Any questions?

Q & A

DR. VELLUCCI: I am Laura Vellucci, from Ortho. Could you please clarify if you have a committee review letter and you make your response to FDA, does the clock start again, and how long does FDA have to respond to a committee review letter? Is my question clear?

DR. FINKBOHNER: Yes.

DR. LEE: If I understand your question correctly, let's say you have a PAS, when we send you a CR letter the response time is four months. Then, if it is a CBE it is six months.

DR. VELLUCCI: Sometimes you submit a prior approval supplement and four months to the day you can sometimes then get your committee review letter. So, four months has lapsed, or almost four months has lapsed and you have, you know, 15 questions that they would like more answers to and you provide that in a timely fashion. Does the clock start again? How long do you have to anticipate because sometimes that just varies--

DR. LEE: I will defer my answer to my more experienced colleagues here.

DR. FINKBOHNER: That is a very good question. When we receive a complete response to the letter that is sent outlining the deficiencies, it is six months from the received date at CBER.

DR. VELLUCCI: So, in four months you will do the initial review but then it is another six months before you could actually get approval?

DR. FINKBOHNER: As a maximum time. Just to clarify, for a prior approval supplement we have a maximum of four months to respond with an initial response of some sort officially. Hopefully, depending on the type of issues, those can be discussed through telephone conferences and potentially be resolved in a less formal way. If the issues require additional validation studies or would require additional time for documentation to be gathered, that would extend past our action time and we will often send a letter. Those, again, are maximum times. Then, once we receive a complete response to the deficiencies outlined in the letter, we would have six months maximum to respond. I don't know what your

experience has been in that case, but I can tell you that today I signed off on an STN assignment approval letter. So, that was in house less than 23 weeks. So, again, these are maximum times.

DR. VELLUCCI: Okay, and I have one other question, if I may. When you read the changes to be effected document, it is very clear, the different categories. But sometimes you have an older license and you may not have everything in that initial license that you would have if you were doing a BLA today. So, there could be changes to--I don't know--maybe how you qualify on your incoming raw material, minor changes. I guess really my question is how far back do you have to go for every change, even if it is not included in your license, especially some of the older ones, change to a test method or change to a component used in a test method to test an in-process material? Do you understand my question? Where can we kind of draw the line with having to report a change?

DR. FINKBOHNER: Actually, what I would recommend is something that Tim Lee has mentioned, I would always encourage you to contact the agency when you have these

kind of situations. For instance, regarding the raw material, it would depend an awful lot on a case by case situation for that material. If there is a question that potentially implicates the source material as being a BSE-involved country or other kind of issues that are raised to a higher level of concern, we may want to have some additional discussions with you. If it is something that is a USP monograph material, depending on the kind of change but if it is in conformance with compendial requirements, that wouldn't usually be a major issue for us. So, again, it is going to be case by case depending on the type of issues that play into it. That is why, again, I would encourage what Tim Lee had said, contact us whenever there is a question.

DR. VELLUCCI: Thank you.

DR. ZEID: Bob Zeid, TLI Development. I was wondering if you could share some thoughts on what kind of supporting stability packages you have seen with a range of annual reports all the way through prior approval supplements. Does it range from just commitment to follow long-term the lots that have been done, or do you see three months accelerated?

DR. LEE: Usually we require both, both real-time stability and also accelerated stability, and also usually if you are comparing before and after manufacturing changes we would like to see a tend analysis for the product that you used before and after the change.

DR. ZEID: And with regards to the accelerated stability, is there a minimum that you are looking for, a minimum time duration like 30 days, 60 days?

DR. LEE: I think it depends on the product, how stable your product is and what the condition of the accelerated study is.

DR. ZEID: One other quick question is the impact of a comparability protocol on lot release characterization testing, or lot release testing, have you ever had a situation where somebody submits a comparability protocol and through expanded characterization testing, or new information, somehow that might impact now their lot release testing protocol?

DR. FINKBOHNER: What you are asking is the potential for having changes in the release testing requirements due to the outcomes of the data collected as part of a comparability protocol assessment?

DR. ZEID: Let me ask it a little differently. You have a battery of release testing for the approved application and then sometimes too with the consent decree, NOIR, etc. or other extenuating circumstances you may have a lot release test protocol of lot by lot or skip lot testing and this includes an expanded version which may or may not be the same as all the characterization testing that was in the application. I guess my point is now you come along with a comparability protocol and you have sort of the state-of-the-art or the new information that you have learned, the modified process, could this now impact what you would actually put into a lot release protocol? Of, how are the two linked? Because lot release protocol or lot release testing is sort of a fuzzy area for a lot of people.

DR. FINKBOHNER: Actually, I am going to give you a really wishy-washy answer on this. I think it really depends an awful lot on what the specifics are. You mentioned consent decree and NOIR, if there is an ongoing compliance action it is going to depend an awful lot on what the specifics are that are built into the consent decree, clauses--I forget the correct term for that-- and

we would be in contact and consultation with our compliance groups to look at how those issues would be melded together, but I don't think we can really give an answer to that.

MR. CHERNEY: Barry Cherney, I am an attorney from the FDA. If we were doing a review of a comparability protocol and a characterization test and we felt that the test was particularly useful as a lot release test, we could ask you to implement that, and we have done that on rare occasions.

DR. KALTOVICH: Florence Kaltovich, SAIC. I have had the opportunity to see a lot several times where companies have sent in their interpretation of what they would think should be in an annual report, or it has also been FDA's decision to change that. Do you have any data to show how many times FDA--or clarification on issues in the guidance document perhaps where these decisions were not made correctly by the manufacturers for comparability protocols?

DR. FINKBOHNER: Thanks, Flo. For those who may not know, Flo and I had offices next door to each other about seven years ago before she left the agency. Anyway,

if the question was do we have a data set and a thorough analysis of this type of downgrade, I am not aware of any thorough analysis of this. I have seen anecdotally items that have been, quote/unquote, under-reported by one tier and as much as multiple tiers. I think a worst case one was a new filling machine in an aseptic area that wasn't submitted as an annual report, which was a little extreme of a downgrade. So, we have seen a little bit of everything. I would think, as a general rule, it is more the exception than the rule that there is a change in reporting category when you look at the overall number of items that are reported in the total annual reports that come in.

DR. KALTOVICH: I also have an example where something that was sent in an annual report was not reviewed by the agency for well more than three months, and your slide said three months. In fact, this was almost a year that the annual report had been submitted and, in this particular case, the FDA reviewer decided that it should have been at least a CBE-30 and required a supplement at that time from the manufacturer. How often do you see

something like that? Or, are your review times truly three months as you stated in your slide?

DR. LEE: Well, I don't have any experience with that so I am wondering whether my other FDA colleagues will fill me in.

DR. FINKBOHNER: It is extremely rare.

DR. CHANG: Again, I don't have a database for that, but one thing I want to emphasize, which Dr. Lee emphasized in his presentation, is that when the sponsor makes a decision on a particular manufacturing change or particular type of their reporting category, it is your responsibility to assure that that particular reporting category does not jeopardize the safety and efficacy of the product. In terms of the actual study to support the change, it is the agency's expectation that you should do a validation to the same extent and evaluation of that particular change so that, hopefully, you will not make many decisions that we will not accept. But, again, I don't have data exactly to say how many times the finding is not acceptable where that happens.

DR. POLLAK: Lewis Pollack, NABI Biopharmaceuticals. I would like it if FDA could comment

on manufacturing a new investigational product in a licensed facility and submitting that as a CBE-30 because I seem to have gotten some different responses from FDA to that question.

DR. LEE: Go, ahead, John.

DR. FINKBOHNER: Actually, a number of those supplements come into our review group. If a facility is licensed to produce a single product, the first time introduction of the product into that licensed area so that it becomes a second product in that licensed area, be it investigational or a another licensed product, it would be a prior approval supplement because, at that point, what we are doing as the first time assessment of moving to a multi-product facility is checking changeover procedures, line clearance, assessing cleaning validation needs, etc., and if there are any special segregation aspects that need to be in place. For instance, we wouldn't be very thrilled if somebody wanted to bring gene therapy vector production in with a licensed recombinant. So, there are some specific segregation issues of types of products.

DR. POLLACK: After you have more than one product in a facility, to bring another investigational product that would always be a CB-30?

DR. FINKBOHNER: In general, yes. The only times that I have ever seen that we have considered upgrading that to a higher level--we did have a situation where, luckily, they contacted us before they implemented the change. They did want to bring a gene therapy product into a plasma fractionation area, and we weren't real thrilled about that idea. Again, when in doubt, always feel free to contact us.

DR. RITTER: Nadine Ritter, consultant. I would like to ask a question that goes back a little bit to Bob Zeid's question about the relationship between analytical methodology that you choose for characterization versus lot release testing, especially for products which have been licensed a long time, like plasma-derived products have been, for which there may or may not be very good compendium methods available. For example, in the recent past I was involved in contract testing and we would frequently get calls from clients who would like to have me perform compendial tests of a particular plasma-derived

product and I couldn't even get the equipment and people who could run it retired a long time ago. Yet, the new methodology has not yet been upgraded in the USP. I know there is movement in the USP to do that. What is your opinion, or what comments could you make about companies which are stuck with doing old compendial type of methods when there are new methods? And, we know what the rules are for going from a compendial method to a new method where the management defaults back to using the old, less sensitive methodology for the existing products rather than upgrading to the new one?

DR. FINKBOHNER: It seems to me the people to answer that question are probably in the audience, actually, in terms of particular-specific assays.

DR. RITTER: This is a general question, just in terms of your perspectives. I mean, we have been saying this for a while, it is your expectation that a new methodology will be used whenever it is appropriate. I think Bob Zeid's question brought that up, which is you do a comparability study, you have the new methodology, you know the answers but now you have to go back and apply it to products which are already licensed and maybe you don't

want to know what is in that band, or you don't want to know what is under that peak and you would rather use the old, less sensitive method because at least it doesn't tell you anything you don't want to know.

DR. CHANG: Well, we have CGMP in place, right? We need good manufacturing practice with current standards. Yes, we do have some situations where we experience that some old testing method is not the-state-of-art testing and the company wants to make a change. It is science-based regulation. So, for what you said, we would like to have some justification in the application. We have a lot of scientists in the agency who can also, you know, make some scientific judgment there. So, we would look at specific cases and see where some study can be performed. In this case communication with the agency is very useful for these kinds of issues. As you said, yes, we encourage using more scientifically sound technology in production and quality control assessment.

DR. HEALEY: Is there one more question?

DR. GOLDMAN: Mark Goldman, Chiron Corporation. I actually have a couple of short ones I think. The first one, as equipment becomes outmoded or unavailable you no

longer have a comparable system. The annual report guidance says specifically you can change an exact piece of equipment with no change in process parameters, but what if that piece of equipment really doesn't exist anymore and you are doing substantially the same thing? Let's make it easy, the piece of equipment is a centrifuge but it is not the same kind of configuration, for example. There is no example under the CB-30 listing that addresses that, that you need to jump up to a BAS; it doesn't seem to be warranted in all cases. I was wondering if you could respond to that.

DR. FINKBOHNER: Actually, unfortunately, I don't have my guidance with me but I thought there was a reference to equipment of similar operating principle being under the CBE-30 category because we have handled supplements that way many times. I don't have the exact quote and citation in front of me, but there is that flexibility there. It doesn't have to be a prior approval supplement.

DR. GOLDMAN: The other question was we have had the experience in the past where we have had CBE-30s reclassified as PASs. In that case, we get a letter back

saying this is a PAS and we are considering it as such. But, more recently, we have had the experience where at the end of four months we were given a letter that said you must resubmit this as a BAS, starting the whole process over again. Do you have any kind of standard approach to this?

DR. FINKBOHNER: I am not familiar with that situation myself.

MR. CHERNEY: I am not familiar with that, but at the end of four months you would be getting a complete review letter and, in responding to that, you would then start with the six-month time clock. That is what is happening there I think. Is that not right? It is not a PAS; it is a response to our complete review letter, but that response is under a six-month time clock. Unless it was a refuse to file, but it wasn't a refuse to file, right?

DR. GOLDMAN: It was a refuse to file.

MR. CHERNEY: Oh, it was a refuse to file.

DR. FINKBOHNER: Actually, this sounds like it may be something we could have a little more detailed discussion to find out more particulars. I don't know if

we could give you a general response without more details. Maybe we could discuss it further off-line or maybe after the session you could meet with myself and we could identify the appropriate people in the agency to continue discussion on that point. I would be interested in discussing that myself.

DR. GOLDING: Just one general comment, from the perspective of the reviewers, the CBE and the CBE-30 mechanism, I mean, I think it puts the FDA at risk and the company at risk in certain respects, and that is that if there is something potentially harmful that could arise from the manufacturer going ahead and distributing the product we have less time to review it. In fact, with a CBE it could go out the next day. So, there is a chance for potential harm. I would just underscore what was said earlier, that if there is a question about the submission and the status, the FDA should be consulted, or else it should be submitted as a PAS. There is clearly a risk both to the company and to the FDA if something falls between the cracks and it is potentially harmful. I have seen at least one example in the last year where a product was distributed on the basis of a CBE and it turned out to be

an extremely potentially harmful situation. So, I would advise caution about how those different mechanisms are used.

DR. HEALEY: Our next speaker is John Finkbohner.

Comparability Protocols

FDA Perspective

DR. FINKBOHNER: Thank you. Thanks for the good questions. First, I am going to have very much an overview presentation but before we go into this, I would like to ask just a general question. Can everybody raise their hand for just a second? Let's just see if your arms are working. Great, a few rotator cuff injuries. I just want to get a feeling for the folks who are in the audience here. How many folks are in regulatory affairs? A good number. So, a lot of this is very applicable to your job in terms of dealing with the regulatory strategy and so forth. I was wondering how many folk here would characterize themselves as being in a quality assurance role. Great. Production manufacturing folks? A few. Let's see, how many FDA-ers are here? Okay. Good, we have a nice mixed crowd here.

I do want to add one other clarification to Frank Rauschen's presentation. His slide 14 had mentioned that CBE-30 would be the category for the introduction of an unlicensed material into a licensed area. Thank you for asking that question because we have a chance to clarify that. That is only in a situation where you have a previously licensed multiple product production area.

Anyway, I am really going to be focusing on two different aspects of FDA's perspective on the utility of comparability protocols from the manufacturing supplement sense, that being mainly the utility and then some of the problems that we have seen so, hopefully, we can help you to avoid some of those pitfalls when it comes time for submitting your next comparability protocol.

A couple of quick notes back to the regulation. I know you have all heard about this way too much today so we will keep it straightforward. It allows for a potential downgrade of the reporting category, as has been noted many times, usually one reporting category tier. It is important to remember that we cannot downgrade a prior approval supplement that might require a pre-approval inspection. Okay? So, that is a caveat that needs to be

kept in mind so don't even consider a comparability protocol if you are building a new filling facility, please. I know I have said that a few times before.

Anyway, when are situations when a comparability protocol can be useful in terms of the establishment description section and some of the CMC data related to facilities? It really falls into two major categories, those being a one-time change where the impact on market distribution could have a negative impact on public health, and then also in situations where there may be multiple identical changes to be phased in over a period of time.

In addition to these two general categories where comparability protocols are useful for a manufacturing facility change, we have also seen them have quite a bit of power and utility in complex changes where there is a multi-phase regulatory strategy with multiple submissions coming in over a period of time to implement a number of changes. An example of this might be if you have a licensed aseptic filling area which is going to undergo a series of renovations, say, taking a non-critical zone of the aseptic filling suite, putting up a temporary construction wall, expanding that area into additional

space to make a second filling line; a second phase where the construction area is dropped and reconfigured for requalification of use of the licensed area with the final prior approval supplement for the new filling line. The final prior approval supplement for the new filling line cannot be downgraded because that requires an inspection.

However, some of the issues related to changes in flows, especially for a lyophilized product aseptic filling suite, there have been cases where comparability protocols have been very useful in a series of four to five different supplements to cover a phase-in of a major project like this. So, again, they are a very powerful tool.

What are some of the problems that we have seen? Again, this is, hopefully, to help you and us because when we get a nice, clean comparability protocol it makes it easier for us to review as well. So, we are very happy to be able to get some of these ideas out to you.

Some of the issues that we have seen with comparability protocols submissions--again, this is the original comparability protocol submitted as a prior approval supplement, and before I get into these bullet points I would like to emphasize that it is the

comparability protocol. So, if you think of the whole idea of a validation protocol or qualification protocol, it should have that same kind of parallel thought process, a detailed description of the types of methods that are going to be utilized to assess the impact of the change; the acceptance criteria that will be utilized; and any test methodologies that will be utilized.

So, some of the problems that we have seen in assessing comparability protocols, and we have had to ask people to rewrite them and revise them, have included not including relevant copies of SOPs or, I should say, copies of relevant SOPs. We don't necessarily want to see the SOP for changing out the mop bucket in a janitor's closet in an unclassified corridor. No, relevant SOPs might be something involving change-over procedures if it is going to be a critical assessment of a multiple product facility that could have potential impact on aseptic processing. So, you have to use judgment there as to what would be a relevant SOP and what wouldn't. A relevant SOP would be how you are going to handle your samples that are being collected as part of this assessment under the protocol.

Another issue that we have seen, unfortunately, is not prospectively defining the acceptance criteria that will be applied to all of the methodologies used to assess the change.

Not including a thorough description of the methods to be used--I was involved in the review of one that said we will assess bioburden. That is nice. Is it microbial limits? Is it some sort of unofficial test? What method is going to be used and what are the limits going to be?

Not including all the relevant aspects of the assessing of the change, for instance, in this earlier example of a series of supplements for an aseptic filling change, we have to make sure that all of the relevant aspects are being assessed, such as requalifying classified areas; conducting the immediate challenges to verify the capability of aseptic refilling the product, etc.

Also, we have had some problems in the past with versions of SOPs being changed after a comparability protocol has been approved. If it is a minor change that doesn't impact the overall assessment or procedure, then it is not usually going to be a major issue and I think Frank

had mentioned the utility of incorporating this into a comparability protocol, that if there were going to be minor revisions to SOPs that that be built into the protocol as to how that would be handled in the subsequent submission. But if the protocol calls for using test methodology A to assess the change and approve it under that sort of venue of testing, and then the SOP is revised to put in place a completely different methodology utilizing a different method that may cause the comparability protocol to be applicable for the specific change that it has been approved for.

Again, as I mentioned earlier, this is technically a protocol and it can become problematic if the submission is structured in such a way that your operators can't perform the assessments and collect the samples needed to actually have your subsequent submission ready to submit. It sounds like straightforward and common sense, but we did have a situation where the subsequent submission that came in was rather mangled because the operators really couldn't follow what had been written to use as a protocol, at least that is the explanation that was given

to us when we called and said can you please explain what this is.

It is also important to remember that a subsequent study really needs to follow the protocol as it is approved, and if there are going to be minor changes to cross-reference SOPs and methodologies, there may be some utility in building some limited degree of flexibility into the original protocol. I would tend to discourage that because as soon as you move into that approach you can be opening up a can of worms and take an awful lot longer to get nailed down.

Some of the potential problems that we have seen with the subsequent submissions that have come in are not meeting the prospectively defined acceptance criteria, and if the study fails to meet the protocol criteria then effectively you have failed to demonstrate comparability of the protocol so, potentially there would be downgrade in that reporting category.

As Tim and several folk have said, I strongly urge anyone who has any questions about a change in a reporting category to contact us and discuss the situation in more detail prior to getting into a situation where

there may be a problem due to a reporting category misunderstanding. So, always feel free to call us. Thank you.

DR. HEALEY: Our next speaker will be Paul Gil, of Bayer Corporation.

Industry Perspective

DR. GIL: Good afternoon. My name is Paul Gil. I am with Bayer Corporation. I would like to expand on the topic of comparability protocols a bit and, hopefully, not be redundant because we have heard quite a bit on it in this session. I want to focus specifically on manufacturing changes and specifically facility changes and improvements to facilities.

Just to outline the talk, I am going to talk very, very briefly about the FDA Modernization Act and the origin of the initiative to streamline the regulatory approval process. Changes to approved application I think we went over quite extensively. I would like to further focus on facility changes and upgrades, and look at some specific examples of facility-related change categories, talk a little bit about protocols, complication protocol strategy and design and, finally, tell you a little bit

about a case study that we have for a comparability protocol, specifically for changes in HVAC, including an aseptic filling area which I think, as Dr. Gross alluded to earlier, is a reduction from a PAS to a CBE-30 which gives you the most bang for the buck, as he put it.

Just very briefly, the FDA Modernization Act included various initiatives and programs and included the streamlining approval process for drug and biological manufacturing changes, and that has evolved into guidance documents and the comparability programs that we have in place now.

As we heard, 21 CFR 601.12 prescribes the requirements of reporting changes for licensed biological products to FDA, and those type of changes include changes to the product, the labeling, the production process, QC, or quality controls, equipment, facilities and responsible personnel.

Just very briefly again, the annual report includes those changes that have the minimal potential to have an adverse effect on the safety and efficacy of the product. Examples of annual report type changes for facilities would be water or steam supply systems using

equipment of similar materials of construction, design and operating parameters and not affecting specifications; room upgrades such as improved finishes on walls and floors; and replacement of equipment with that of identical design. I believe that was the wording in the guidance document.

Also, in the annual report changes would include in that category upgrades in air quality, material or personnel flow; equipment relocation with no change in air classification; modifications to pretreatment stages of a WFI system and installation of non-process related equipment or room upgrades, and those would include warehouse refrigerators or freezers.

Again quickly, the CBE-30 category are those changes that have a moderate potential to have an adverse effect on the safety and efficacy of the product. Example of facility-related CBE-30 changes would be adding new walls to increase environmental control; downgrading of a room or area classification, of course, not including the aseptic area which we all know, as soon as you get into a class A filling area you are talking about a prior approval supplement. Addition of duplicate process trains with no process change; addition or reduction in number of pieces

of equipment; new product in a previously approved multi-product area, as we talked about earlier; and new or modified purified water systems.

The PAS would be those changes that have a substantial potential to have an adverse effect on the safety and efficacy of a product. With respect to facility changes, those would include major construction for new or modified areas. Those would include aseptic processing areas, HVAC systems that would feed these aseptic processing areas, WF5 systems, any type of change that would have the potential to stress or challenge the system; environmentally controlled manufacturing areas such as new buildings, production areas or rooms; anything that might have a potential to affect the cross-contamination precautions that are in place; and critical utility support systems, anything that has a potential to affect the air quality, the water quality or the steam quality.

The burden of proof, of course, is with the manufacturer, with the sponsor. The burden of proof is that you must demonstrate through appropriate validation and/or other studies the lack of adverse effect on the

safety, identity, strength, quality and purity of the product.

Comparability protocols come into play because those are the mechanisms by which we can really streamline the regulatory approval of manufacturing process changes. As described in CFR, they establish the tests to be done and acceptable limits and specifications that are to be achieved to demonstrate the lack of adverse effects, thus the burden of proof, for specified types of manufacturing changes on the safety and effectiveness of a product, getting a little bit more about the definition of specified types of manufacturing changes.

I think as we touched upon, the comparability protocol is submitted as a prior approval supplement. It describes the plan for implementing the change. It includes, as Dr. Finkbohner pointed out, the validation studies and stipulates the acceptance criteria, the acceptance limits, and, hopefully, once approved the comparability protocol may reduce the reporting category by at least one tier and, as we talked about, that would be from a PAS to a CBE-30 or CBE-30 to CBE or annual report.

The comparability protocol allows the change to be facilitated with thorough planning, and I think that is key here, and abbreviated implementation time frame, the time from implementing the change to distributing product is going to be shortened by this process. It increases the flexibility to bring important and improved biological products to market more efficiently and expeditiously. It entails consistent requirements well understood before the change is implemented. This ties into the planning and also incorporates discussions with FDA. We have a mixture of older facilities at our site and new facilities, and the older facilities are particularly challenged with constantly upgrading those facilities to improve and stay current in GMP design. Several years ago, when we embarked on a master plan for upgrade of those facilities we had a conference with FDA and basically outlined our plan and discussed various phases of implementation of these changes and, at the same time, we talked about reporting categories so we at least had a baseline of agreement for how these should be reported.

For a comparability protocol the manufacturer should carefully assess the product change and evaluate the

product resulting from the changes to the preexisting product. Results of testing must demonstrate that the manufacturing process change does not affect the safety, purity or potency effectiveness. You have to take into consideration that the same change may have variable effects on different products. Also, implementation of a change in multiple facilities under the same license, we have several comparability protocols in place now which cover two of our facilities at Bayer. Also, you must take into consideration that comparability protocols are optional and not mandatory, and the benefit there, of course, is that they may allow for more expedient distribution of product.

Appropriate uses of CPs, appropriate for discrete situations, specific changes with manufacturing experience. Now, I think one of the classic examples is if you would like to upgrade four virtually identical filling lines, you plan to implement changes on all four so you are actually repeating, you could actually expedite those changes and the distribution of product following those changes through a CP because the same change and the same validation will be necessary for each of those filling lines. So,

typically in that case it would be more of a repetitive protocol. So, it is specific for a type of change and the applicant. Of course, validation is key. The equipment must be qualified and the process must be validated. It results in a comparable product which has defined characteristics and meets established measurable criteria.

Inappropriate uses of a comparability protocol, as Dr. Finkbohner pointed out, are those submissions which require a pre-approval inspection, new filling facility, new product manufacturing facility. Also inappropriate would be having outstanding compliance issues or not defining the change; it is either too broad or too complex. Inappropriate use would be having the predefined acceptance criteria not available or omitting any type of discussion and explanation OOS results. Also, if the change adversely affects the product, or the change results in a newly characterized product.

I will talk about a case study for HVAC. When we strategized on how to develop a comparability protocol, we knew that we wanted to make multiple changes in the aseptic filling area. That would all fall into the pre-approval supplement category. We wanted to maximize the impact and

the utility of the comparability protocol. So, we took the approach of using one comparability protocol for several changes, making sure that when we described the changes and reported on the changes that they were clear and distinct. We wanted to maximize the impact so we picked the comparability protocol topic on aseptic filling areas for HVAC changes. Again, that is the biggest bang for the buck, a PAS down to a CBE-30. We focused the comparability protocol changes on the most critically controlled environment, which is the aseptic filling area.

When we first met to devise the comparability protocol, we reviewed in detail the validation, engineering, regulatory, quality, the planned changes that were involved to install a new system or modify existing systems that we had in older facilities. These are systems that supply the aseptic processing areas. Some of the things we wanted to do were addition or replacement of air handling units; modifications to fill line equipment which may change the air quality or the air flow patterns; reconfiguration of some of the air supplies or returns; and make some modifications to the fill-line curtains.

So, when we got together to discuss the activities that would clearly demonstrate that these changes would not have an adverse effect on the process of the product, we came up with certain criteria. So the changes were all found to have similar criteria. All of those changes were required changes, validation in the form of IQ and OQ qualifications; air flow visualizations to be sure that we didn't disrupt or change any air flow patterns--all of these changes would involve extensive environmental monitoring to demonstrate control of the area, and all would involve proof of the pudding, which we call the media fills. So, any one of those changes all had the same criteria, and that was the basis of us bundling those together in a single comparability protocol.

The comparability protocol details the change scope qualification activities, the acceptance criteria and the supporting data required to demonstrate the lack of adverse effect following the change.

In developing the CP, again, our strategy was to describe all the anticipated system changes. We incorporated several of these planned system changes into one comparability protocol. We identified all the SOPs

that were associated and validation test functions. We talked about extensive environmental monitoring and defined our qualification and monitoring scheme. We specified the acceptance criteria, and we also said that we would need to describe any actions if the results were unacceptable. We tried to foresee anything that could go wrong and, as you know, if you fail to meet the acceptance criteria that could cause rejection of the change according to the CBE, resulting in knock-down grading by a tier to a CBE-30, rather. Again, we talked about the planned implementation of these changes to the aseptic filling area and the schedule for doing so.

The supporting data must substantiate the development of the strategy and provide adequate validation documentation that appropriate chemical, physical and microbial controls are in place, and the type of supporting data is dependent on the type of change and must be evaluated on a case by case basis.

We try to get out of the mode of submitting all the IQ/OQ validation data, which I don't believe FDA really wants to see for these changes. What is key is that we had to identify certain key aspects and things that we had to

be sure were met for IQ/OQ particular examples on these changes. So, we developed a list of IQ things that we would check for, and anything that had to do with the qualification of this aseptic filling area following the change for IQ and OQ.

We reasoned that to demonstrate fully these changes, the PQ data would have to be submitted in its entirety. PQ data, that is, environmental monitoring, total particles, all those data for a particular period of extensive monitoring needed to be presented to show the agency and demonstrate the level of control in that area. So, we did not take any short cuts or attempt to summarize any of the PQ data. That particular stack of monitoring data was included, as was the proof of the pudding, the media fills. All the media fill data was submitted, the number of units and any type of deviation that we had was explained.

Now, there is a fine line between not being able to find an attributable cause and having that lead to a rejection of the CP, or, if you think about the entire volume of sample as an attributable cause, that is certainly not going to be a reason, we would think, for

rejecting the CP. So, you really have to look at the deviation itself. Obviously, trends will be really seriously looked at. But if there are no trends and just the occasional sporadic deviation that can be explained, I think that is a different story.

Some of the top misconceptions that we have found on comparability protocols through our training program have been that once you have a comparability protocol approved with the agency for a specific type of change, you always have to submit that particular change under the CP. What we found is the notion of scope creep, which is we have intentions of doing a specified change and through engineering evaluations or other means we find that we want to expand that particular change to be above and beyond what is described in the CP, yes, we can do that but it convolutes, as in some cases it may, the issue. Sure, we can submit it, but it is not going to go into the comparability protocol because, remember, the comparability protocol is a protocol; it is a series of steps that one must follow in order to get that reduced reporting category.

Another misconception is that an approved CP alleviates the need to submit data; only a summary is required. As I mentioned, that is possible in some cases but when it comes to PQ data, media fill data for HVAC changes, I don't think any short cuts are warranted, and I think any short cuts in those areas would be particularly risky for CP change approval.

Also, another misconception is that comparability protocols are quickly approved and modified--I will just submit the change and get that through; no problem because some acceptance criteria may not be accurate or be able to be met. Well, the course of action is the same course of action as to get the CP approved in the first place and it involves submitting it as a pre-approval supplement. I might add, you must have extensive rationale, sound, scientific rationale for making any changes to the CP.

Lastly, one of the misconceptions about CPs is that it is acceptable worldwide for worldwide registrations. It simply is not. A change to an HVAC system for an aseptic filling area which could be reduced for U.S., under a comparability protocol, from a PAS to a CBE-30, in Canada for the same submission there is no

mechanism for that, for example. They would not allow distribution of the product following the change until they have had a much longer review period. So, there is no mechanism to shorten that in certain regulatory environments.

Finally, I wanted to talk about the elements of the CP submission. As I said, the comparability protocol we designed in an executable format is a protocol that one follows and it is very detailed. It prescribes what type of qualification testing and analytical testing is required. So, as we compile these elements of the submission, we always want to make sure that we have the complete description of the changes and a system overview. We want to present to the reviewer clearly what it is that we are changing, what it was before and after the change. We always include detailed drawings of the system changes. We try to highlight the areas on those drawings of the aspects of the system that have been changed. We include the executed complication protocol. We have the approved CP. When we make the change and follow-up with the testing we actually execute that protocol and include a copy with our submission for approval.

Validation summary, again as I said, we summarize and list the key aspects of our IQ/OQ and we always include all the PQ data, not just a summary of that. Again, I think it is very key that all deviations are fully investigated, explained and documented. Again, having unexplained results could jeopardize the use of the CP to reduce the reporting requirements.

Again as Dr. Finkbohner pointed out, copies of SOPs and labels, you know, he made it clear that it is the key SOPs that need to be approved and not all SOPs to change the mop bucket, as he put it.

In conclusion, I think the key here is planning the comparability protocol, the details of the change, the qualification and the acceptance criteria and the burden of proof. I think if you can list and detail the change, and you know what you are going to do and you guard against scope creep during these changes and plan it properly, for both industry and FDA the use of comparability protocols are a win-win strategy. It allows reduction in the reporting category, and it also will actually extract and present the key elements of the submission that the FDA wants to see. Thank you.

DR. HEALEY: Thank you, Paul. We have some time now for Q&A before we go to a break. So, once again, please stand up if you have questions and remember to use the microphone, and state your name and affiliation. Also, I see some questions coming in on 3 X 5 cards.

Q & A

DR. FINKBOHNER: This is directed to FDA. I will do my best to read it. If you have an approved comparability protocol and have a single deviation, parenthetically OOS, would CBER accept this with data justifying it as an isolated incident with a root cause determined?

I will make a first comment on that. Again, I know it sounds horrible but it really depends on a case by case situation. It is horrible, isn't it? Let me give an example. Let's say that as part of an originally approved comparability protocol there is a specification that all of your pharmaceutical grade water samples should be taken to the QC lab and tested within six hours without freezing, and a QC technician tests them at the eight-hour mark. That is technically a failure to follow protocol. Is there a significant impact on the water testing results for the

way that one set of samples was handled? If you have data to support that there is no significant impact, it is probably a reasonable deviation and data to support the appropriateness of that deviation.

If it is a deviation in terms of predefined acceptance criteria for an autoclave where the validation protocol is written in such a way that you have a thermal mapping of the chamber with 18 thermocouples, which all must give you data of at least 121.5 degrees Celsius and two of your leads--how many people here have ever validated an autoclave? So, you know exactly what I mean. If you crimp one of those thermal couples through the access port you are not going to get data. If you have written the protocol in such a way that you don't allow for that kind of a failure, then you have technically failed your own protocol. Is it significant? Maybe not if it is not a critical thermal mapping location in the chamber based upon other studies you may have.

Anyway, to go back to the earlier statement, it depends. Potentially it is okay, but the deviation should be thoroughly investigated and if there is a corrective and preventative action plan that needs to be put in place or

supportive data to justify that the data collected is still reasonable to support the change, then it may be all right to move forward with it.

DR. GOLDING: Something that Dr. Gil said reminded me of a situation which maybe we can correct. He mentioned that it is possible in certain situations to do bundling of comparability protocols. What happens at our end sometimes, and I have seen several examples of this, is that we get submissions which are based on the product and the license which have the same information. It may be comparability protocol; it may be a change in a bottle or a stopper. Multiple reviewers get this and the reviewers may not know that the other reviewer has an identical set of data but for a different product. So, my plea to the industry is when you are bundling up issues like this, in your cover letter you should state that at the time you are submitting this particular submission you are also submitting a similar submission with similar data for Factor VIII or IGIV, albumin or whatever so that the reviewers immediately read that, hopefully, and get together and expedite their review and come up with a consistent comment.

DR. GIL: Just to clarify what I mean to say, one particular comparability protocol could be used to describe several different changes. For example, we wouldn't necessarily bundle multiple changes into one study because then, should the results of the study prove to be a failure, for example, we would not know what change made that cause. So, what we do is we basically have a comparability protocol where you would select what the change was, for example, curtain changes on aseptic filling or a new HVAC system, but I don't think we would actually bundle them together in one submission. We would use the same comparability protocol to validate and to demonstrate control following those changes.

DR. HEALEY: We have a couple of write-ins up here.

DR. FINKBOHNER: Yes, we have another question on a card, could you comment on industry's perception that if they call the FDA to get an opinion on reporting categories, i.e., CBE-30 versus PAS, the answer is almost always the more conservative option.

[Laughter]

I am trying to think of the right comment for that. We get a number of these type of calls and often the follow-up question will be will you please fax some additional information to describe this change in detail so we can give you a more reasonable and thoroughly considered response? A phone call with a three-sentence description that we are expanding our aseptic filling facility and reformulating two of our licensed products and we want to see if we can submit a comparability protocol for change one and maybe for change two and three as a bundle thing. Now, we have to have a little bit more detail to give you a rational response and a thoroughly considered one. So, in general when we have a lack of information we will fall back to the more conservative advice.

DR. LEE: I have a question here but I am not quite sure whether I can read it. I think probably I need help from my colleagues to figure out the answer here. Do you have a single product facility if you wish to reformulate it and make clinical supplies in your facility, is it still a PAS? Did I read the question correctly?

DR. FINKBOHNER: Yes.

DR. LEE: I think I am going to defer the answer to my more experienced colleagues.

DR. FINKBOHNER: Can I restate the question just to confirm we have it right? If I understand it correctly, you have a licensed product where you would like to introduce a second formulation for additional clinical trials into a licensed area, and would that be a PAS? So, basically a secondary investigational formulation for an already licensed product. What do you all think? I think it would really be something that would have to be discussed between the product office and our group and some other groups in the Center, but in general I think Dr. Golding raises a good point, from the equipment and facilities viewpoint, since it is not a currently license multi-product area that would be the equivalent of having a new operation performed in the area and we would want to discuss further with you the type of segregation of operations if there are product dedicated parts or if there would be shared process contact parts. I think all of us are nodding that that sound like a major potential impact.

[Discussion away from the microphone]

DR. HEALEY: I think the additional question was except for the segregation.

DR. FINKBOHNER: Again, I think it would depend partly on what the controls are in place to demonstrate that the process streams would not potentially crosscontaminate. Again, it depends on the specifics but I think most of us are nodding our heads that it doesn't sound like a real red flag for us in terms of removing something from an already licensed formulation. Don't forget, call us.

DR. GEIGERT: John Geigert, consultant. I would like to follow-up on your comment, John. We have talked about calling the FDA and getting maybe mixed advice occasionally. We get a more conservative answer than a company chooses to want. There is also the other case where the company may not give the correct information to a reviewer and they get a lower tiered response. I have actually run into that and had to come in and tell them they had to go back and give the right answer. You have to ask the right question to get the right answer.

Part of it is communication. We have the guidance. It was written in terms of an illustration of

how to handle a number of these changes, written in 1997. It obviously was very quickly put out. It was thought through but it went out very quickly and, of course, industry didn't have much chance to comment, at least that is the impression I get.

Is it possible for that to be updated? Clearly, the FDA has a wealth of information now dealing with all these changes and, very clearly, if that could get somehow communicated to help industry understand exactly what the current thought is at the agency on these different levels, it might eliminate some of the confusion either in terms of conservatism on one end or incorrect questions going into the FDA that lead to misleading comments. You have a number of procedures. You have SOPs which you guys have on your web site. You have all sorts of mechanisms so it may not be the most long-term mechanism but some way of getting this out because there are a lot of people who are struggling in the industry. They don't know where to put their change, and if they don't ask the right question or they don't ask the right person they can get misled, unfortunately, and create major delays for them.

DR. FINKBOHNER: Thank you for your suggestion.

DR. JONECKIS: Those guidances, as John said, have been out for sometime and, actually, we are taking a loot at those at present to decide what type of revisions are needed. Unfortunately, under good guidance practices which we all now follow, putting things and SOPs on the web sites for the intention of guidances is sort of verboten. So, we will have to go through the formal guidances procedure but we are actually looking at those in combination with some other actually sort of associated regulation changes that have been sitting since actually the last PDUFA.

DR. GIL: I have a question. It says that with regards to having the comparability protocol as an executable protocol, for example, an IQ/OQ, is it as detailed as a normal IQ protocol or does it only consist of a summary covering the most important steps?

What we did basically is to reduce the volume that is submitted for IQ/OQ, we took each of the test functions in the IQ/OQ and listed what the test function consisted of, what the expected outcome or specification needed to be, and then what we did, we had actually almost like a check-off, like an initial where we included the

reference in the validation protocol where that test function could be found, where those results could be found. So, basically we created a list of the test functions, a summary, if you will, and for each test function and acceptance criteria we verified and listed the reference where that particular test function result could be found. We did that for both the IQ and the OQ.

PARTICIPANT: [Not at microphone; inaudible]

DR. GIL: Not for IQ/OQ. We did submit the completed form, if you will, listing the test function and showed that we were diligent in verifying each of those critical items.

DR. HEALEY: Any additional questions before the break? If not, why don't we take our break and we will see you back here at 3:45.

[Brief recess]

DR. GOLDING: I would like to get the final session started, so if everyone would gravitate to their seat. I have the honor to introduce the next session. As you can see from your programs, we are going to be going over some case studies, hopefully, to shed some light on the comparability question. The first speaker is Angela

Blackshere. She is the director for global regulatory affairs Baxter BioScience and Biopharmaceuticals.

Case Study Presentation from Industry

DR. BLACKSHERE: Good afternoon. I am pleased to share Baxter BioScience's experience on comparability to help in the understanding between FDA and industry on this topic.

The case study I will present looks at the use of comparability studies to reduce clinical trial requirements. It also involves exchange of intermediates, which will be further discussed in tomorrow's session.

Before I go into the manufacturing change, a little history is necessary. Baxter acquired Immunolg in 1997 and, as a result of that acquisition, had access to another IGIV product, immune globulin intravenous human. So, Baxter is licensed to manufacture two IGIV products and for the purposes of this presentation I will refer to one as pathway 1 IGIV and pathway 2 IGIV. Both products are manufactured from Fraction II intermediates, which is a common intermediate for manufacture of IGIV products. However, the Fraction II intermediates are manufactured by slightly different modified co-manufacturing processes.

The change that we proposed to the FDA was to include the option of utilizing Fraction II intermediate from the pathway 2 process as starting material for the pathway 1 IGIV.

I have a schematic here to help show the differences between the two processes. Option A is what we were currently licensed for, starting with plasma, going to Fraction I plus II plus III in the pathway 1 Fraction II and continuing with pathway 1 downstream process that includes S/D treatment and purification and, ultimately, the pathway 1 IGIV product.

The option B that we wanted to include in our license was to start with plasma or coagulation factor deficient plasma, which is simply the plasma that is subjected to different treatments so that factors can be removed to go on to further manufacture of the products. So, you continue with Fraction II plus III as opposed to I plus II plus III, and get a pathway 2 fraction intermediate and continue with the pathway 1 IGIV downstream process, which again includes S/D treatment and purification.

The purpose of the change from Baxter's perspective was to increase manufacturing efficiency, as

well as flexibility, increase yield. As a result of combining the two processes, we increased yield which ultimately results in increased product availability.

Our regulatory strategy was to look at the FDA guidance document on comparability. In addition, we had plenty of experience from a manufacturing perspective and clinical perspective over the years. So, we wanted to submit a prior approval BLA supplement with no clinical data, based on this guidance document which you have seen several times today. The guidance document states that manufacturers of biological products may make manufacturing changes without conducting additional clinical efficacy studies if comparability testing demonstrates to the FDA that the product, after the manufacturing change, is safe, pure, potent and effective. Coming from a regulatory standpoint, this is something that the team likes to hear from us normally.

Also, this guidance document states that determinations of product comparability may be based on chemical, physical and biological assays and, in some cases, other non-clinical data. And, if a sponsor can demonstrate comparability, additional clinical safety

and/or efficacy trials with the new product would generally not be needed.

Our goal was to demonstrate that the pathway 1 IGIV manufactured from the pathway 1 Fraction II, which is what we were already licensed for, is comparable to pathway 1 IGIV manufactured from the new starting material, pathway 2, Fraction II, by the performance of analytical and biological testing, and also to demonstrate that this pathway 1 IGIV, the new version, meets the same prescribed standards of safety, identity, purity and potency.

To begin our analytical and biological testing program we wanted to start our program by demonstrating the comparability of Fraction II intermediates manufactured by the two different pathways by performing the following characterization tests: Molecular size distribution was evaluated which gives us an assessment of integrity and purity and safety. Purity was evaluated by electrophoresis. Amidolytic activity, which is an indication of safety. Anticomplementary activity gives us an indication of both safety and efficacy. Fibrinogen content, information on purity; non-IgG protein content on purity, and IgG subclass distribution on efficacy.

We continued with our program by looking at comparison of the final container products. We looked at the same characterization program but we also looked at it showing that the pathway 1 IGIV product manufactured with the pathway 2 Fraction II intermediate would meet those established final container specifications for the licensed product. So, in addition to these tests that were performed at the Fraction II stage, as a part of the release criteria we also looked at antibody titers and pre-kallikrein activator activity, antibody titers giving us an indication of safety and efficacy and pre-kallikrein activator activity of safety and efficacy as well.

The results of our testing showed that the Fraction II intermediates were comparable, except there was detectable amidolytic activity in the pathway 2 Fraction II intermediate. This amidolytic activity is an indication of potential proteolytic activity. We continued with our characterization program and the results showed that at the final container the products were comparable and this amidolytic activity was no longer detected. So, our downstream process was robust enough to remove the differences that were observed at the Fraction II stage.

The FDA response to this data was that they did not require an efficacy clinical study; they did not require the performance of pharmacology, toxicology animal studies or in vitro studies. The FDA had some safety concerns because of the amidolytic activity that was observed, so they requested the performance of a safety clinical study. They also requested process validation data on a large scale, since all of the data that we generated previously was on a smaller scale, to demonstrate removal of this activity by pathway 1 IGIV process.

The safety clinical design that we proposed was to compare the incidence of infusions with the investigational product, associated with related adverse events, to the incidence of adverse events observed for infusions of the licensed pathway 1 IGIV product that we had previously studied in a Phase IV safety clinical trial. So, we had data that we could compare to look at the adverse event profile of this new version of the product.

The proposal was to evaluate 40 subjects with primary immune deficiency, and for the statistical analysis we needed to show that the adverse event rate was no larger than 40 percent. We submitted an IND to manufacture the

pathway 1 IGIV from the pathway 2 Fraction II intermediate. The safety clinical design was accepted and IND not placed on clinical hold. Then we received an FDA letter allowing us to proceed with the study.

Although I am not at liberty to comment on the progress of our clinical trial studies, I think this is a good example because it shows the comparability program that was utilized; the FDA decision-making and the applicability of the guidance document to reduce clinical trial requirements for manufacturing changes. Our beginning goal was no clinical trials, however, we did not have to perform an efficacy trial and had to perform a limited, small safety study. Thank you.

DR. GOLDING: We are going to continue with the next case study. This will be presented by Ghiorghis Ghenbot. Ghiorghis is a senior manager at Aventis Behring.

Case Study Presentation from Industry

DR. GHENBOT: Good afternoon. We have heard quite a bit about the principles of CP and cases where this comparability protocol can be applied. So, in this particular case I am going to focus our attention to a case study where qualification of an alternate source of murine

ascites derived monoclonal antibody reagent for the manufacture of Factor VIII was used.

The outline of the presentation follows. We want to define exactly what the objective was and then go over the justification of this particular activity, and then look at the aspects of the comparability protocol and consider what will and what will not change in this activity, as well as highlight the essential components of the protocol, and then give you an activities overview. In addition, we will look at the acceptance criteria for the monoclonal antibodies as a reagent, not as a therapeutic agent. Also, the acceptance criteria for the immunoaffinity column.

The objective is clearly stated here. We had one particular objective in this case. We wanted to qualify an alternate supply for the monoclonal antibody reagent used in the immunoaffinity preparation of Factor VIII. But, at the same time, we wanted to demonstrate by a comparability protocol that the safety, identity, purity and potency of the monoclonal antibody reagent, as well as the product itself was not affected by the change.

The justification of this CP lies in 21 CFR, Part 601.12(e) which says that an applicant may submit one or more protocols describing the specific tests and validation studies and acceptable limits to be achieved to demonstrate lack of adverse effect for the specified types of manufacturing changes on the identity, strength, quality, purity or potency of the product. It seems a little bit redundant to say it over and over again, but that is the bottom line because these are regulatory guidelines.

The key aspects of this comparability protocol as summarized here. The product affected, of course, was Factor VIII which is a licensed product. The process step was a discrete step, and this is the immunoaffinity purification step. The reagent in question, of course, is the antibody which is being used as a reagent again, not as a therapeutic agent. The supplier in question, of course, is that you are changing one supplier and going to another supplier. The basis of qualification is, as I just mentioned, the FDA approved comparability protocol.

In this exercise two things need to be clarified. There are things that will not change during the entire process, and there are things that will change. What will

change, of course, is the supplier. In other words, maybe the first supplier was so rich because of the money that we gave them that they decided to change their business or to go to other areas, and we had to look for a different supplier.

What will not change, however, is very, very important in this comparability protocol. We have the testing specifications for procedures used for the ascites production, which are going to be conducted by the alternate supplier and will not change. The testing specifications or procedures used for the monoclonal antibody purification, conducted by the Aventis Behring manufacturing facility will not change. The coupling of the monoclonal antibody to the solid support, the chemistry, the linkage chemistry, will not change. The currently licensed Monoclate-P or Factor VIII for the manufacturing process of this product will not be affected. And, the product specifications, of course, should not change.

Furthermore, the essential components of this comparability protocol as follows: A comparison of the mouse strains and care, hybridoma cell line, the clone

itself, the statement of work and mater batch records between these two suppliers should be identical. Acceptance criteria for comparison of cell culture and ascites product processing parameters should be identical between the two sources. The acceptance criteria for functional comparability tests of the ascitic fluid, the purified monoclonal antibody, and the resulting monoclonal antibody immunoaffinity resin as a purification reagent, and conformance to current licensed specifications, of course, should not change.

Now, if you look at the summary of activities, you can basically break them down into the old supplier and the new supplier and basically, as I said, in a nutshell we have a statement of work, master batch records, specifications, master cell bank, culture medium and procedures, as well as the source material. Where it says plus/plus means they should be identical to the extent possible between the two sources.

What are the acceptance criteria for the monoclonal antibody obtained from the suppliers? You could categorize the quality parameters with respect to safety, purity, quantity and potency, and you could look at the

specification testing as well as the description for the specification testing. Most of this is actually in the guidelines for the monoclonal antibody reagent and I don't need to go into too much detail.

We had additional characterization for the purified monoclonal antibody because it is a new supplier, of course, and the testing methods include some of the standard analytical procedures, as well as some of the newer approaches. You can look at IEF as well as SDS-PAGE under different conditions; HPLC and well as CZE, and Western Blot Analysis for the subclasses, as well as binding affinities by titration colorimetry.

Furthermore, if you look at the immunoaffinity column itself, you could qualify the column in terms of manufacturing a small-scale affinity column, and you want to look at the final production of Factor VIII.

In conclusion, this is a case of a successful comparability protocol, and we are very happy about it and hope that our future comparability endeavors will be as successful as this one here. In this case we have looked at an alternate supplier of a monoclonal antibody reagent and it has been successfully qualified. Furthermore, the

safety, identity, purity and potency of this monoclonal antibody reagent, as well as the Factor VIII product itself has remained intact.

I want to add that this comparability protocol, this exercise is a concerted team effort, and I wish to thank all the parties involved in this particular case, including tech ops, R&D, QC, QA, manufacturing and regulatory. I am only their spokesperson here. They would be very happy to come here and talk about this further. I appreciate your attention. If you have any questions, I would be very glad to answer them.

DR. LYNCH: We have a ten-minute Q&A period following this section. So, I think I would like to move on to the last speaker, Dr. Basil Golding. He is the Deputy Director for the Division of Hematology. Dr. Golding is also the head of laboratory of plasma derivatives, which is the part of hematology that has jurisdiction over the albumin products, IGIVs, IGs, alpha-1 PIs, hemoglobin--he is a very busy guy, and he will give the FDA perspective case studies on manufacturing changes as well.

Case Study Presentation from FDA

DR. GOLDING: Thank you, Tom. I am going to present some case studies that will give you a few examples. You can see from the title that the examples include plasminogen and pre-kallikrein activator.

Just some general comments related to biologics comparability, as we have said multiple times during this meeting, plasma pools are very complex mixes of many, many different molecules with potential for protein effects. We know the identity of many of these molecules, but there are probably many molecules that we don't even know that they exist in the plasma.

The other point I would like to make is that minor production--minor as is in quotation marks on purpose--minor production or purification process changes may change the product and the level of impurities in the product, and small changes to large molecules may be important but very difficult to detect even by sophisticated physicochemical methods.

Impurities may be active. They may affect the activity of the product. They may relate to immunogenicity of the product, and they may affect the stability of the product. As I already mentioned, probably not all the

impurities are known, both in the source plasma and in the product. In addition, products may change during storage. Actually, in one of the examples you will see this exemplified.

Products may be complex and difficult to characterize, for example immune globulins. As you know, immune globulins may be millions of different molecules and it is impossible to characterize what kind of immune globulins you have in a polyclonal immune globulin preparation.

So, plasma fractionation, some of the variations and variables of the actual process, starting from the beginning in the donor selection, we have source plasma versus recovered plasma. These may differ in a number of points. The biomarkers may be different. Usually the viral markers are higher in source plasma. The number of donors involved in collection of a plasma pool would be probably different. You need more recovered plasma to achieve the same volume. In addition, viral safety issues, efficacy, dimer content, demographics may all be different in these different IGIV products. So, these are different variables that go into the donors.

In terms of manufacturing efforts, the Cohn-Oncley, which is still the commonest used manufacturing process, is a multi-step process which has several parameters at each step, including pH, temperature, ionic strength, alcohol concentration and protein concentration. Chromatographic steps are also often used either alone or in addition to the Cohn-Oncley process and this can induce variables, and the viral clearance method used may have an effect on the product. The viral clearance could be solvent detergent; it could be heating; it could be nanofiltration or other methods.

Other variables which could influence the product are the excipients. In recent years we became aware that sucrose in some of the products was associated with an increased incidence of renal failure, and this resulted in us sending out a "dear doctor" letter, and asking manufacturers to include in their labeling a warning statement regarding sucrose and renal failure.

The product itself may form aggregates, and aggregates can be associated with activation of complements, and there has been some suggestion that they

may relate to making the product immunogenic and inducing autoimmune disease.

During manufacture, class and subclass removal may occur. For example, certain chromatographic steps are effective in removing IgA and this is important because IgA can induce anaphylaxis in recipients that have elective IgA deficiency. But we have also seen that the same manufacturing step actually can, and often does, remove IgG4. There isn't any good evidence for IgG4 playing a protective role against infections. So, this is still an open question, whether this is a detrimental effect.

But we have also noticed that some manufacturing involves the loss of IgG3. Just to point out, IgG3 is a subclass that is very sensitive to proteolysis, and I would propose that it is important when new manufacturers come on board that they should look at subclasses, and in particular look at IgG3 because degradation of IgG3 could be a sensitive marker of something in the process that is degrading proteins.

Contaminants that we have become aware of over the years, and which may have an impact on safety and efficacy are the pre-kallikrein activator, and I am going

to go into this in more detail. Alpha-2 macroglobulin fraction does contain inhibitors of some of the vasoactive substances. This is actually something that you may want to have in the product. And, plasminogen is a proenzyme which, in a liquid product and under the right pH conditions, could degrade the immune globulin and would impact on stability. This is a theoretical concern, but we learn more and more about bacterial walls and, in addition to LPS, we now know that bacterial DNA and cell wall components are very active in inducing responses, proinflammatory responses from the innate immune system. Currently, we test for pyrogens and we test for LPS, but I don't think we have tests in place that could detect small amounts of these components which may be associated with adverse events to immune globulins.

Getting back to the question of how minor changes could have a major impact, the first example I am going to deal with relates to plasminogen. In the precipitation of Fraction II plus IIIw in the Cohn-Oncley manufacturing scheme, depending on the pH of this step, you may get a different outcome in terms of plasminogen contaminating the product. During this step we get Fraction III which is

precipitated and discarded, and we get supernatant III which is used to go on to make Fraction II and IGIV. What has been found is that at a pH of 5.4, at a higher pH, the product is often much more stable than if this step is done at a low pH.

The reason for that is that when you do it at the low pH some of the plasminogen ends up in the supernatant, whereas, if you do it at the higher pH no plasminogen ends up in the supernatant. All the plasminogen goes into the precipitate, into the Fraction III which is discarded. So, the outcome is that when you do it at the lower pH the product is less stable than when you do it at the higher pH.

I have mentioned before that source plasma is different from recovered plasma, and I am just outlining some of the differences. Source plasma is frozen early so there is less possibility of activation of enzymes, and probably less potential for vasoactive compounds to be activated. In terms of recovered plasma, it is frozen at variable times or not at all, and there is a potential for more activation of enzymes and more vasoactive compounds being found in the final product.

Hypotension was notice many years ago, in the '70s, and the paper was published by Barbara Alving and John Finlayson was also on this paper. So, this was worked out at CBER, and they noticed an association of pre-kallikrein activator in plasma protein fraction. This was a precursor product to albumin and is used for volume expansion.

The example I am going to cite is an example of a manufacturing change which was actually designed to improve the safety profile but resulted in more frequent side effects. So, this was done some time ago, and the manufacturing change was not reported to the FDA at that time. But the manufacturing change rationale was as follows: A paper was published by Bland and co-workers showing that PPF, protein plasma fraction, given at time of cardiopulmonary bypass could cause hypotension. They suggested in the article that this may be due to bradykinin-like molecules that are normally cleared by the lungs.

One manufacturer changes its process by increasing exposure upstream in the manufacturing to surface activation by stirring for a longer period of time

in filter-aid to purposely generate bradykinin, and then presumed that at the ultrafiltration step the bradykinin, which is a small peptide, would be removed by ultrafiltration. So, they went ahead and instituted this manufacturing change.

Just to remind you of the pathway, Factor XIIa is a pre-kallikrein activator. When Factor XII is exposed to negatively charged surfaces, you get generation of Factor XIIa, which is a fragment, and this Factor XIIa can act on pre-kallikrein to give you kallikrein, and the kallikrein can acto on kininogen to give you bradykinin. Bradykinin has very potent vasodilatory effect on small blood vessels and can cause a very sharp drop in blood pressure. Normally bradykinin will go through the lungs and will be degraded by kinases in the lungs. When you are performing cardiopulmonary bypass you actually circumvent the lungs and the bradykinin can then go into general arterial circulation and has great potential to cause hypotension.

So, this is taken from the paper by Barbara Alving where she looked at different lots from the manufacturer that were associated with adverse events and lots that were produced in the same time period from other

manufacturers, most of which had very low incidence of hypotension but did have some cases of hypotension. These two columns are looking at the levels of PKA. The first column, on the left, are the implicated lots, in other words, lots that were associated with hypotension. In the right column are lots that were not associated with hypotension. So, these are the lots from the manufacturer that introduced the increased stirring and, as a result, were found to have increased PKA levels in the product and were associated with an increased incidence of hypotension. This is before cardiopulmonary bypass, during that procedure and during general surgery.

The conclusion from that study was that there was a marked increase in reporting of hypotension following the manufacturing change, and the hypotensive episodes were related to the rate of infusion and the high levels of PKA in the product.

Now I am going to go on to a much more recent episode where pre-kallikrein activator was associated with hypotension, this time in five percent albumin product rather than in PPF.

Again, the manufacturing change was not reported to the FDA but was very similar in that the manufacturer, at one of the upstream steps, decided to increase the stirring time prior to fractionation. This was associated with an increase in case reporting of hypotensive episodes, and we responded by sending an investigator from the FDA to the company.

I scanned this in from the 483 reports so the quality of these graphs is not so great. But on the left the axis is PKA expressed as the percentage of the FDA control. The recommended level of PKA in PPF is usually higher than albumin; it is around 35 and this is because for PPF there are very clear instructions for how to infuse it, and the infusion rate has to be really slow to avoid adverse events to PPF. But for albumin the recommended level is 10, expressed in IU/ml but it is actually a percentage of the FDA control.

The X axis here shows minutes of stirring. So, what the company was asked to do, and they presented it in graphic form, is to relate the time of stirring to the levels of PKA in the final product. What you can see is that there is a cluster of high PKA levels at a particular

range of stirring in minutes, but somewhere between 300 and 600 minutes. Then it seems to fall off again. But most of the lots were less than 600 minutes. So, there seemed to be a trend of increasing PKA levels as the stirring time was increasing. Here is a very marked increase in PKA level.

This is looking at a time course, taking two lots that, at release, were well within the range that we would like to see, below 10, but with time during storage the retention samples were tested for PKA, and what we noticed was something very interesting. During storage the PKA levels increase in both these lots. This one particular lot peaked at 12 months and this other lot peaked later, but there was an increase and then a decrease of PKA activity.

So, just to remind you that the actual assay is a functional assay and we don't conclusive evidence but what we think is happening there is that, because of the increased stirring, you are generating PKA but in addition to the PKA there are inhibitory molecules at the time of release and that is preventing the PKA activity from being seen here, but the inhibitory molecules degrade with time

and the PKA levels increase. Then eventually the PKA itself degrades and you have a fall.

So, what was actually happening with these lots is that even some of them were within the release specifications, they were increasing the PKA levels with time and that is when you started to see the hypotensive episodes.

This again shows that there was actually a period in time when the lots that were exposed to the increased stirring had increased PKA levels. Again, you are looking at PKA levels on this axis, and these are days after release of those lots. So, these lots we are looking at here were all within spec at time of release, but with time after release there was a window of time when they exceeded the specification for PKA, and this is the time when they were associated with hypotensive episodes.

In conclusion, the two separate incidents of increased stirring times resulted in increased PKA levels and increased incidence of hypotensive episodes. In the first case the manufacturing change did not reveal any changes in release testing. There was no PKA testing at the time. In the second case PKA levels after release

increased but were within specification for the lots that were released, obviously, and for those lots within specification after release PKA levels often exceeded specifications during storage, and these were the lots that were associated with the hypotensive episodes.

Conclusions are that the production of immune globulin is a multi-step process. So-called minor variations can have far-reaching effects on safety and efficacy. Each IGIV product should be regarded as unique, and the term "generic" does not apply to biologics. Thank you.

DR. LYNCH: I guess according to the schedule we have a ten-minute Q&A on the instant presentations you have heard. For my money, I think the case studies really bring into focus some of the issues that we have been wrestling with in a more general sense throughout the day. So, I would encourage anyone with a question for any of the three speakers here to come forward and be heard.

Q & A

DR. GEIGERT: John Geigert. For Dr. Ghenbot, you said you did a comparability protocol. Was that from a PAS

down to a CBE-30, or CBE-30 down to an AR? What was the actual downgrading? It wasn't mentioned in your slides.

DR. GHENBOT: CBE-30

DR. GEIGERT: From a PAS down to a CBE-30.

DR. GHENBOT: I think so.

DR. GEIGERT: Also for ascites, there is no mention of any kind of virus testing comparability between suppliers. Surely, you did that.

DR. GHENBOT: Yes. In other words, I couldn't go through tests of all those specifications that we have for this particular reagent. All the required tests were done appropriately, and we did not see any negative issues.

DR. GEIGERT: And for Dr. Blackshere, if you had not seen a difference in a safety assay, suppose it had been a purity assay, do you think the FDA would have asked you to do a clinical safety study?

DR. BLACKSHERE: I think the FDA may want to comment on that but I think another route we may have taken was to spend more time on process validation and demonstrated it at the large-scale prior to the submission.

DR. LYNCH: A question, Angela, you did not pursue a comparability protocol for the changes you

described, but presumably you had some opportunity to discuss the program with the agency before you implemented it.

DR. BLACKSHERE: Well, we had a first meeting with the FDA just to discuss the strategy, and after we had more data and actually saw the differences in that particular assay, we had a pre-IND meeting and we went from there.

DR. SEAVER: This is for Dr. Ghenbot. When you talked about your acceptance criteria for these tests, I am trying to distinguish what you can present to us versus what you really did, and I think one thing that surprises me is that everything was comparable to reference or peak, or something like that. I am used to seeing that for, like, Phase I data but for products I am really used to seeing the peak at X, Y, Z time and something else like that, and much more quantitative type acceptance criteria. Could you comment on that?

DR. GHENBOT: In terms of details, I can't really say that the limits were between X, Y, Z and this particular product fell X-plus or Y-minus. So, I am not at liberty to say exactly what the specs are. However, let me

say that what we have seen was absolutely acceptable and, more importantly, I think if you look at the final product, if you compare that product which was generated using the previous monoclonal antibody reagent supplied by the previous supplier compared to the new one, it was virtually indistinguishable.

DR. VIDOR: Arlene Vidor, Baxter. I have a question for Dr. Golding. This is more of a regulatory policy question. Obviously, there is a lot of discussion with sponsors when you are faced with the possibility of having to do a clinical study to support a manufacturing change, and it may tip the scale for you in favor or not doing the manufacturing change even if it is an improvement or an enhancement because you may have logistical issues to contend with, or resource issues and there may just not be enough clinical subjects to go around. I was wondering if the agency has considered the possibility, from the standpoint of practicality, of having a modular review process where one could submit the CMC section of the application while running out the clinical trial so that there could be some concurrent review going on, thereby shortening the time frame for review? I guess it would be

more similar to the device review paradigm, but I wonder what CBER's thinking is on that or what your thinking is on that.

DR. GOLDING: Well, I don't think I want to make a policy decision that would affect many reviewers and many people right here on the spot, but I think that that type of review, revolving review or fast track review, has been invoked under certain circumstances. For example, if there is an acute shortage or there are justifiable reasons to do that, I think the agency could be persuaded that that is something that we would allow. But I think there would have to be something in addition to just standard procedure.

DR. LYNCH: The PKA examples that you gave, of course, are classics. Looking at the more recent of those events, one associated with five percent albumin, with the benefit of hindsight, a couple of things really strike you. One is the number of non-conforming. It was a very high percentage and one wonders why that didn't set an alarm bell off. The second is, of course, the increase in PKA over time as the product was stored, making one wonder why this didn't come to light during routine stability studies.

And, are these factors that should be taken into account to perhaps avoid a repeat of the same incident with a different protein?

DR. GOLDING: Well, I can't answer for the company, obviously, but just in their defense, if I can put myself in that situation, in their defense, many of the lots did actually pass in the sense that they were less than 10 IU/ml, and the problem was that they went up with time. While they were doing the increased stirring, they still had many lots at the same time that had the reduced stirring and it seemed to be a convenience. That is my guess. We never got that as a direct answer. Depending on the shift and the time of day, they might stir it for longer or less.

So, mixed in with all the slightly elevated lots there were a lot of lots that were still well within specs. The usual lot that was stirred with the lower stirring time had almost undetectable PKA. There was an increase for most of the lots, around 5 and 8, when there was increased stirring, even at release. But they released those lots, and those were the lots that went up with time and fell. So, I agree that, you know, looking at that data, obviously

with hindsight and maybe with real time, people in QA should have said, well, our usual spec is 10 but our usual finding is less than 1 and now we are seeing multiple lots that are 5, 6 or 7 and that should have triggered some kind of response and they should have started to look at it in more detail. But what actually triggered them looking at the data in more detail was adverse event reporting, which could have been avoided I think if it was looked at more carefully.

DR. GHENBOT: I have one follow-up question to that. We talked quite a bit about product heterogeneity for plasma-derived products. We do know for a fact that for recombinant products, transgenic products, there is a certain degree of heterogeneity and that is either because of the expression system, depending on what system you choose, or because of lactation time. I your experience, could you break the two apart and say you have to follow this particular approach in order for you to specify and say that your product should not be any more heterogeneous than this level?

DR. GOLDING: Let's try talking first about recombinant products. Again, I think, just based on basic

principles, that if you are making a recombinant product and it has a certain heterogeneity by some physicochemical method, and possibly it could be due to glycosylation or other differences, but once you have documented a certain heterogeneity and you go ahead and you do your clinical trials, I think there should be a specification that this is the heterogeneity that we are going to allow for release and there shouldn't be major departures from that heterogeneity in subsequent lots. If there, that could be a problem.

Regarding the recombinant product, I think it is easier to come up with a rational answer to your question. When you are talking about something at the other extreme, like IGIV, having heterogeneity in terms of having multiple antibody molecules there is actually good. So, you know, you want all the subclasses. You want as many specificities as possible. And, I would argue the other way, that in terms of immune globulin, if you start to lose specificity or lose subclasses which you see in some products with IgG3 levels or undetectable or very low, or we have seen products where they are having difficulty in meeting the release specifications in terms of specificity

against some of the required testing--measles, or polio, or diphtheria or hepatitis B or surface antigen. In that case, the loss of the heterogeneity is a problem.

But in dealing with other products, like albumin or Factor VII, again I think that you define from a physicochemical point of view your product up front. You do your clinical trial; you show that the product is safe and effective and then you don't want to have departures from that in subsequent lots.

DR. GHENBOT: Thank you.

DR. BAKER: This is Don Baker, from Baxter. I was having trouble formulating this question and I hope I have it right now. There is a saying in law, something to the effect that extreme cases make bad law. I was contemplating the examples that Dr. Golding was presenting, and one could interpret that to say that every minor production change ought to receive the most extreme review. In reality, I think for most of those changes, those minor changes, we are going to have 20/20 hindsight. If they are detrimental, we are going to find that out in the field. We probably actually won't discover that even with an extreme review. And, the cost of subjecting every minor

change to an extreme review will inevitably substantially delay progress in permitting changes in production processes and bringing these improvements forward. So, I would like to get the committee's perspective on just what is the take-home message from the last set of examples.

DR. GOLDING: Maybe I can start. The message that I was trying to deliver, and I think many speakers have alluded to this during the course of this first day of the meeting, was that what we are dealing here is a very complex situation; that the starting material is very complex and the manufacturing in some instances is very complex. The characterization of the final product, just using physicochemical means, is probably not complete. What we have learned from these cases that I have shown and from other examples is that very low concentrations of certain impurities that are very active can have major clinical effects.

So, I think the message is that we have to be very careful about looking at any changes and making decisions about whether analytical methods and preclinical data are sufficient, and whether we need to do safety trials or even efficacy trials. Now, I am not saying that

one should go from one extreme to the other, and I think your question is a very good question, but I think the problem that we have, all of us have, FDA and industry, is that our knowledge is incomplete. So, it is very hard to make good decisions when you don't have all the information. That is why making some of these decisions is very difficult.

I would argue that if there is sufficient doubt, one should err on the side of safety and doing the clinical studies. But one has to balance this with availability of product and allowing the industry to survive economically. So, I think we have to make some kind of balance here and it is a question of judgment, but I think that it is very difficult to make these judgments without having all the information and without having enough data to decide whether this should be approved based on no clinical studies or not.

DR. FINLAYSON: As a contemporary of Francis Bacon, perhaps I can be forgiven for taking a simplistic approach. But the answer, I think, to the question of what is the take-home message in this is he who will not learn from history is condemned to repeat it. I mean, how many

times do we have to repeat the experiment of learning that a pH change at this stage of fractionation will give you plasminogen in your product, which may be activated to plasmin, which may chew up your product, which may have safety and efficacy consequences? How many times do we have to repeat the experiment of learning that if you stir in the presence of a negatively charged surface you can have activation of the contact activation system and this can generate vasoactive substances? How many times do we have to repeat the experiment of saying that if you think stability means putting one lot a year on a stability program, you are going to get bitten by it?

Now, when you say stability, people go into this knee-jerk response and say, oh, but that means you have to measure 27 different things and we don't have enough people in QA to do that. Maybe you don't have to measure everything on 16 lots; maybe you have to measure one or two things on virtually every lot. So, to me, the message is very simple--look at what happened to yourselves; look at what happened to your competitors and maybe, heaven forbid, listen to the FDA and think about it.

DR. LYNCH: Has John now inhibited the rest of his colleagues?

[Laughter]

DR. CHANG: It is very hard to add anything now.

DR. LYNCH: Everybody is afraid to speak!

DR. CHANG: Actually, this particular case made me include one sentence in one of my slides. Dr. Baker, remember, we had this discussion at the last workshop on how exactly we should assess impurity for plasma-derived products. One of the measures that actually you proposed is whether or not we can have a certain level of impurity, such as lower than the natural concentration in the plasma, and we looked at some of the cases for some of the impurities. Then, I included in my slides that I recommended that toxic impurities should be identified.

So, how are we going to identify the toxic impurities? One way is learning from history. As Dr. Finlayson pointed out, perhaps the industry should have a database somewhere to capture all the lessons learned and to share that information between the companies. One workshop like this is a good way to learn between different companies and also the agency. So, I think we should

identify the toxic impurities and then control them to a level that will not have a safety concern.

DR. LYNCH: I actually would not think of disagreeing with John--it is a bad career move--

[Laughter]

--but I think he may be arguing too much here. In fairness to Don Baker, I don't think he was suggesting that we raise the specifications for PKA in albumin. We certainly have a catalog of bad actors that we have learned about through painful long history. We know enough, I hope, to avoid those problems. But, as Dr. Golding points out, there are myriad other proteins for which our understanding is far less well developed. So, using a material like these materials, be it plasma or a derivative thereof, has an unavoidable risk associated with it.

As a reviewer, when you are faced with making a decision what you would like to do is have that risk reduced to zero, but to satisfy that requirement means that you never actually make a decision at all. That is the only way to avoid risk entirely. So, the objective, I would suggest, whether it is in this context of making changes to an approved product or developing a new product,

is the same, is to balance your risk reduction with the anticipated benefit that the product offers.

That is what is hard. That is the hard optimum to strike. If one does that right, there will be the inevitable case where you find out subtle problems with the product only when much greater clinical experience is derived from use in the market than you could possibly accrue during a controlled clinical trial. Otherwise, your controlled clinical trials or your comparability studies would be endless. A zero risk standard paralyzes the industry. I think the FDA has sometimes been questioned unfairly when these long-standing problems come to light, you know, why weren't these prevented up front? Well, because they are doing good risk assessment. A proper balance has that risk associated with it with subtle problems that cannot come to light, except with extensive clinical experience and I think we need to bear that in mind as well.

DR. SEAVER: Sally Seaver, Seaver Associates. Since I am a consultant I don't get kicked by anyone, but I have a quick question, is the PKA assay for this product a

stability indicating assay or an assay that is done during stability studies?

DR. GOLDING: To my knowledge, the FDA has recommended it. It hasn't become part of the regulations. So, we recommend all manufacturers of PPF, albumin and now immune globulin as well, to measure PKA at time of release and during stability. But there is no regulation that mandates measuring PKA.

DR. SEAVER: The reason I bring this up is because I think one lesson that I have learned on comparability studies, and I tend to deal more with specified biologics, is really the usefulness, and this is addressing your question, Don, of the accelerated stability studies. I can remember in particular an example that I think was talked about by Amgen where they had a relatively pure protein. They had several 0.2 micron filtration steps before their final real 0l2 micron sterile filtration step, and they wanted to remove one of these and the product was essentially homogeneous on gels or chromatography, and they did it. What happened was that this 0.2 micron filtration step, in retrospect, was also removing trace proteases

which you couldn't see on any of these analytical methods but which showed up in spades on the accelerated stability.

That is not the only incident. That is a particularly graphic incident that I can recall, but for a long time in biologics we all ignored accelerated stability data because it couldn't be used for real-time data and, therefore, we all decided it wasn't useful. It turns out that accelerated stability studies are incredibly useful and I think, if I would say anything to you, it is that I have seen numerous things where people did changes that they didn't think were very much, and when they did accelerated stability they were picking up all sorts of changes.

DR. CHANG: Now probably people will understand why we insist on requesting accelerated stability data for comparability studies.

DR. LYNCH: Just to follow-up on the stability issue, in your experience, Andrew, in real-time stability on a relatively short program do you accept partial data, with a commitment to supplement that to fulfill, like, a two-year real-time protocol? Or, would you in all cases wait until the two years run?

DR. CHANG: Yes, we accept basically six-month data with a commitment to carry out the real-time stability study. As Dr. Tim Lee pointed out in his presentation, we want to look at trend analysis and that is used as one tool to look at whether or not the product is comparable under stress conditions.

PARTICIPANT: Dr. Golding dropped my favorite bombshell at the very end about the comparability and where the CBER division is going with the generics, because if a company within themselves can compare two products what is the prevention of an abbreviated BLA?

DR. GOLDING: I am not sure that I understand your question.

DR. LYNCH: It is a generic biologic question.

DR. GOLDING: Right. Let me back up a little bit. If you have two separate manufacturers making a product, because these are such complicated products from beginning of the manufacturing to the end, we would not regard them as generic. That is clear. What you are asking is if the same company, using the same manufacturing process, made some kind of change, is that the scenario?

PARTICIPANT: No, I asked the first question because if, after all this, everyone does all these tests and they cannot distinguish between company A's version of the product and company B's version of the product, and company B wants to go to market without any clinical trials because they have done all their homework--

DR. GOLDING: Then it is no go.

PARTICIPANT: It is no go?

DR. GOLDING: It is no go. Just the simple answer to that is it is no go because this is such a complex process that there is no way that we can call these products the same just based on physicochemical characterization because there are just too many variables, too many unknowns for us to be able to look at two different products and say they are the same or different.

DR. JONECKIS: In addition, I would just like to make the point that at the time your product is approved it has been through full clinical studies and testing and, in a sense, what you are doing with comparability is making a leverage off of that. In terms of the chemical sense, you are basing it solely on the chemical characteristics in the determination of a bioequivalence test to make the leap of

faith that you will have the same safety and efficacy, just to follow-up on what Dr. Golding said. That is the other concern to think about.

DR. LYNCH: I think that is a fair statement of CBER's position. What is interesting, of course, is that many biologics are regulated by the Center for Drugs and, of course, those biologics, which are just as complicated as the CBER biologics, may not have a choice. The generic drug provisions apply to those as well. So, it will be interesting to see how this dichotomy unfolds over the next few years.

DR. JONECKIS: I don't want to turn this into a discussion but a couple of things to say is that several speakers have discussed this issue in public forums, and that is not necessarily the policy aspects but the scientific aspects. Dr. Cherney, who is the Deputy Director for the Division of Therapeutic Proteins, has discussed this in various forums and Dot Scott has had some discussions with some of our regulatory colleagues. I have had several discussions with various groups, including the Generic Pharmaceutical Manufacturers Association. So, we are willing to discuss this, but in follow-up to what Tom

said about CBER, Dr. Chu, who is the Office Director for the Office of New Drug Chemistry, is actually proposing that they would not follow the standard generic approach used for chemical entities but more what is known as the 505(b)(2) approach which allows for additional clinical testing of perhaps some demonstration of efficacy and/or safety, including immunogenicity studies.

DR. LYNCH: Yes, that is my understanding as well. It will unfold over the years. Angela, I had a question. Especially during John Finkbohner's presentation the issue of bringing additional products into a dedicated facility came up. Actually, your presentation provoked a question in my mind about bringing an unapproved source material or manufacturing intermediate into a licensed area. Was that issue raised in your program, and how exactly did you address it? Did you have to take any extraordinary measures to change over between your validation runs and your routine manufacturing, or anything like that?

DR. BLACKSHERE: I think we relied on our segregation procedures, including validation, in order to

make sure there was no impact on the existing licensed product.

DR. LYNCH: Is that a general issue? I would assume that if one had an unqualified source material that that might create a problem with the rest of the facility. How generally, if there is a general answer, is that addressed? I mean, one has to do validation in order to qualify the new source material, but one can't do that if one can't bring it into the facility.

DR. GOLDING: You know, I think this has multiple levels to it. It depends, again, on what is the source material and where is it coming from. Just to give you some extreme examples, where we have had manufacturers that are using U.S. licensed plasma and non-U.S. licensed plasma, obviously that raises a lot of issues. The next level is what is actually happening? Besides the material being plasma, is there something that is done to that plasma before it is added to the equipment?

There was one situation where the plasma was treated in a way with a toxic material. It was under an IND. It was added to tanks and there was no cleaning validation in place to show it removed that toxic material.

So, obviously, there are many issues that could occur on the way, but the essential point is that if you are introducing something that is non-licensed into a licensed facility, we need to have validation along the way that the source material, the cleaning of the tanks, or that it is done in separate equipment is all in place and validated before this is done, and done up front, in order to allow this to be done and not impact the licensed material in a negative way. Angela referred to those items.

DR. SCHWARZ: Hans Schwarz of Baxter. If I may clarify, the facility is licensed for both intermediates, to answer your question.

DR. LYNCH: So, that doesn't create a major problem because they are both approved, but approved for different processes.

DR. SCHWARZ: The downstream process is the same.

DR. LYNCH: But to bring, let's say, an intermediate manufactured by a third party into your facility would create the issue that I guess I am trying to tease out here. That is an issue that at least deserves discussion before one does it to make sure that the appropriate safeguards have been put into place so that

ongoing manufacturing of the license product should not be impacted.

DR. SCHWARZ: Correct.

DR. GIL: Just one further comment on that, what we have done with the source material that has been unlicensed, and using it in our manufacturing area for qualification testing is that we have made sure that that source material has been thoroughly tested and had no quality difference from the licensed source plasma. We have also notified the agency and told them what our plans would be to look at cleaning validation, etc., and subsequently submitted data to show that, in fact, there was no contamination, if you will, of the licensed facility equipment. So, we did some extensive monitoring and validation of that as well.

DR. LYNCH: Any other questions? Are there any questions in a more general sense? Are there any questions on any of the issues that were raised today, any of the presentations? I guess now would be a good time to do that. We will have another sort of round table forum tomorrow. There are some interesting presentations, particularly on the intermediates and characterization

thereof that raise some of the most difficult issues, I think, but now may be an opportunity to review today's presentations. Are there any specific comments now?

PARTICIPANT: I don't really have anything profound to say, other than it certainly did raise the question in my mind, since it is very useful to have this type of a forum, about the concept of a database. Dr. Finlayson's comments really struck me very strongly, and I feel that somehow or other we don't always get a firm grasp on what is going on and what the collective experience is of the various companies. We may hear about it through the grapevine, through different networking, but there is really no forum to exchange this information or make it available. I think everybody may have the good intentions of wanting to have as broader base of knowledge as possible, but it is just not always available and maybe there is a possibility for industry and the FDA to work together to somehow sanitize the database so that we don't divulge proprietary information and, yet, make all of the relevant information and findings available to the public so that we can benefit from our collective knowledge and experience. But I don't know how to do that.

DR. LYNCH: Mark, do you know whether there is a mechanism that exists maybe to waive certain aspects?

DR. WEINSTEIN: I don't know that there is a mechanism that exists. That is, of course, what I started out saying this morning, that there are, in fact, issues that we know are paradigm cases, case studies as we have discussed today, and that we should think about that, about ways of doing it. You know, I think that that is one outcome of this meeting that it is important to maybe form a small committee through PPTA and FDA representatives to think collectively about how we might be able to establish some sort of information clearinghouse that would allow sanitized, as you say, information to be available to the industry and be able to take advantage of the FDA database. I think this would help advance the field quite a bit.

DR. LYNCH: You have actually been pretty successful with monitoring product distribution data, in certain respects, by the use of a third party. That initially was a great problem of confidentiality, proprietary commercial confidential information. So, maybe that is a model.

DR. WEINSTEIN: I am not exactly sure what the nature of that third party would be, but it is worth thinking about.

DR. LYNCH: Is Florence still her? Florence is volunteering; she will do it for you!

DR. WEINSTEIN: With that is a contract sort of thing, perhaps PPTA would be involved in that. Georgetown Economic Services I guess is the group that was chosen for the collection of the data.

DR. LYNCH: I don't think they know squat about PKA though.

DR. WEINSTEIN: No, they don't but it is the sort of thing that might go out for some sort of competitive bid.

PARTICIPANT: I have a question. As I think was mentioned, there are a number of reviewers looking at different submissions and focused on different products, within the agency, and particularly within CBER, how do you monitor and track findings from reviewer to reviewer so that the second reviewer would make a consistent judgment? Is there an internal FDA database existing now?

DR. WEINSTEIN: We are in the process, in fact, of setting up such a group at least within the Division of Hematology, where we would actively attempt to do that very thing, set up a database, to have an educational program to make certain that all reviewers have the same information before them. Part of this program, at least as we envision it--this hasn't happened yet but it is in process and actually we have somebody who is gathering data even now and setting up an electronic system to be able to disperse it within our group--but part of this should be an outreach program where, in fact, we are going to attempt to do that. Setting up an educational program is at least one of my highest priorities.

DR. GOLDING: At the moment our database is sitting in the third row and his name is John Finlayson.

[Laughter]

DR. LYNCH: I have kind of a background question to set up another question later on from your presentation earlier on immunogenicity. It was a very intriguing presentation scientifically but today it doesn't help very much wrestling with that issue. Raising a concern about immunogenicity almost guaranties a clinical trial because

there are no good animal models that will allay that concern. There are animal models that will kill a proposed change if they indicate that the immunogenic profile has been somehow changed, but it won't get you over the hump of doing an immunogenicity trial.

So, I guess I would like to ask this question in two parts. One is are these concerns focused don specific products? Obviously, a history of inhibitory antibodies is most pronounced in the history of coagulation factors, not so much we albumin and immune globulins. So, is the concern uniform? Secondly, when setting up to address an immunogenicity concern with a clinical trial, does one rely on the kind of standard comparability trial that, in this case, may involve a great number of patients and a great deal of time, a great deal of exposure time? Or, can one look to other study designs? The ISTH, for example, has proposed certain criteria for concluding that there is not an unacceptable risk of immunogenicity for these coagulation factors.

DR. GOLDING: I think, first of all, that any product--you mentioned some of them that are obviously related, and I will ask John to correct me if I am wrong,

but I think there was, at least in Europe, one treatment of immune globulins with a chemical that induced a change. I am not sure if it caused decreased binding or actually make those molecules immunogenic. Be that as it may, I think potentially the theory is that any chemical change in a molecule that is then injected into a human could be immunogenic. Aggregated immune globulins were typically used to get antibody responses in animals.

I think there is a potential for any molecule that we give, any protein that we give to a human to become immunogenic under certain conditions. Now, the testing in animals, obviously, is not perfect. Doing clinical tests in humans is also not perfect, and I think it comes back to what you said about clinical trials anyhow. For a lot of the immune responses, say, for example, IgE responses, those are relatively rare and if you had that it could be anaphylaxis and death. So, that is a very profound adverse event but you may not see that if your trial is 40 or 100 patients. It may take Phase IV studies to release to the market and then find later on that r procedure is immunogenic.

So, I think we have a problem and I don't think there is a good solution. What I was trying to suggest in my talk is that there are newer approaches that could be directed towards this problem and that we should start thinking about these newer approaches. It might take years before we get to a point where we can say this in vitro testing system can allow us to predict immunogenicity, but I think it would be a start in the right direction.

Besides the in vitro testing of human cells, the other approach that may be worthwhile is using co-called humanized mice that have human antibodies but I think, from that point of view, it is going to take a while before we have enough mouse strains out there that represent the human population.

DR. GHENBOT: Can I add just one comment to what you said? I think early this year there was a workshop on immunogenicity during the WCB conference, back in January or February. There was an entire workshop which was devoted to immunogenicity. If I recall the facts correctly, there was no one particular model that was proposed during that presentation, which tells me that this an industry-wide problem. I just want to put into

perspective that this is not only industry, but industry that covers all biotech biological compounds.

DR. CHANG: I just have a follow-up to that before I try to address the second question. That is right. Up to today, we still haven't found a good animal model or in vitro study to address immunogenicity. I think that Dr. Golding pointed out that we still need effort in this area to develop some new method and maybe five or ten years later we can find a model. But we don't have a model available today for that today.

Tom, to come to your second question, and other colleagues can comment on that, for the Factor VIII which apparently is one of the molecules that has more immunogenicity concern, we have not yet asked a company to do a side by side clinical trial to look at the immunogenicity issues. A general approach that we have is to rely on historical data as the control. The database generated from the Canadian study that looked at the inhibitor formation in previously treated patients is about four percent. That information can be used as some kind of parameter to design a clinical trial to see whether or not a new product after a manufacturing change can have

immunogenicity higher than that or not. So, we have not yet asked for a side by side. But that may not necessarily be a universal approach for other products. If you have a study model where you can use less patients, if we evaluate it and find that it is a scientifically sound approach, then we will let you do that.

DR. LYNCH: The previously treated patients are a great tool. But there still is an issue of what are appropriate criteria, how many exposure days; how long do you have to challenge the patient to achieve a certain level of assurance for no more than a five percent chance of greater than five percent increased risk. At some point you start slicing the baloney pretty thin. The numbers can get high but one has to wonder whether the value is actually much value added.

I think your European colleagues have sort of strayed a bit from the rigorous statistical design in adapting their criteria for acceptable immunogenicity study and I am wondering if FDA--I know in the past the ICH standards were kind of influential, let's say, and I am wondering if there is a change there to a more or less rigorous standard.

DR. WEINSTEIN: We certainly look at the data that we get from our European colleagues and see what its basis is, but there does seem to be a somewhat different philosophy there regarding the necessity for statistical rigor. In some cases there simply isn't a feeling that one needs to have a very strong statistical basis for making decisions where, at least up to now, we have felt that we did want more of a statistical estimate that had some reasonable rationale behind it. We have found that, at least in some studies that have been carried out by our European colleagues, there really was a rather arbitrary selection of numbers depending on the numbers of patients in a particular country where a study was being done.

DR. SCHWARZ: Well, it is late but let me speculate about the future. I think the future will make things more complicated. We can assume that the susceptibility of a patient with hemophilia A to develop inhibitor probably represents a polygenetic disorder. Very rapidly we will be able to identify, on the genetic background, what those additional risks factors are. So, the question will come up will a manufacturer then prospectively select the patients with hemophilia A which

have a low risk, an intermediate risk or a high risk to develop an inhibitor, and then what are the appropriate numbers of patients and how the regulators will look at that, if you then want to evaluate a new product, in terms of the risk of developing inhibitors.

DR. LYNCH: We are really coming to the close of our allotted time. There is one question that was submitted in writing, and it is a follow-up on the accelerated stability discussion. Since accelerated stability is being requested, is it necessary--this is to you, FDA guys--to perform all tests as if a real-time study was being conducted? I suppose this may be a response to John's comment that it may not be necessary or even desirable to do--I don't know, visual appearance necessarily if what you are looking for is PKA activation. Do you have any guidance on that, Andrew or Chris? Simplifying the question, if you are doing an accelerated study, should you have all the--

DR. JONECKIS: Right. Generally, for an accelerated study for the purposes of supporting a comparability protocol, yes, in my experience, people generally will just replicate the existing tests using the

standard sterility protocol, as well as perhaps doing some additional ones if they are looking for particular things. Is it necessary to do that? Again, you know, you are making assumptions if you are only going to look at certain things in your accelerated test. Over the longer period of time, I guess the second associated thing is do you have to do all of the tests all the time? And, I don't think that is the case. There are, I think, certain provisions where you can do, as Dr. Finlayson pointed out, only some of the tests during periods of time. In fact, you can amend the sterility protocol and at least for some other product classes we have allowed that to make it more relevant types of protocols. So, I think there is already an existing mechanism that is out there to do that. So, I think you could do that under the existing stability protocol.

DR. GOLDING: Could I just add one thought to that? You know, when we look at stability data from a company, and we have seen supplements asking for those, they provide you with five years of stability data, and they have been looking, for example, at the immune globulins and they have been checking the monomers and the dimers every three months, and then they find that for over

five years they don't get any useful information by doing it every three months; they could do it once a year or once every six months or once every nine months. What I am trying to say is that it is a good idea to do as much testing as possible, not as little as possible, but once you have shown that certain parameters that you are looking at are extremely stable, you can start cutting back on the testing and only do the testing that is going to reveal changes in your product.

DR. CHANG: Well, one thing I want to add is that you should know your product better and what kind of degradation will occur. I think that is very useful information to keep your stability sample longer than you normally do to learn, and don't throw those stability samples away but learn what degradation occurred and what the mechanism it could be for that degradation or aggregation, and what kind of risk there could be. So, that information can be used in a stability study to support comparability. So, you can build that into study to look at those factors. But what Dr. Joneckis said is what we currently do.

DR. LYNCH: If any of you have any doubt about the import of what Andrew just said, I am here to testify that Dr. Finlayson has samples of albumin that are older than I am.

[Laughter]

With that, it was a wonderful first day. I want to thank all the participants, the speakers, organizers, moderators, my mother, and we will see you back here bright and early tomorrow morning, 8:30. Thank you.

[Whereupon, at 5:30 p.m., the proceedings were recessed, to resume at 8:30 a.m., Friday, May 31, 2002.]

- - -