U.S. FOOD AND DRUG ADMINISTRATION

BLOOD PRODUCTS ADVISORY COMMITTEE

July 20, 2009

Hilton Washington DC North/Gaithersburg
620 Perry Parkway
Gaithersburg, Maryland 20877

"This transcript has not been edited or corrected, but appears as received from the commercial transcribing service. Accordingly the Food and Drug Administration makes no representation as to its accuracy"

Proceedings by:
CASET Associates, Ltd.
Fairfax, Virginia 22030
703-266-8402

Table of Contents

P R O C E E D I N G S

Agenda Item: Opening Remarks

DR. FREAS: Mr. Chairman, members of the committee, invited guests, members of the audience:

My name is Bill Freas. I'm the acting executive secretary for this committee. At this time, I would like to go around the head table and introduce to the audience the members who are seated here at the head table.

I will start at the right-hand side of the room. The first chair is occupied by Dr. Henry Cryer, chief, trauma and critical care, Division of general Surgery, University of California, Los Angeles.

Next, Dr. Ann Zimrin, associate professor, Division of Hematology/Oncology, University of Maryland School of Medicine.

Next is Dr. Cynthia Lewis-Younger, managing medical director, Florida Poison Information Center, Tampa.

Next is Dr. Mark Ballow, chief, Division of Allergy and Immunology, Department of pediatrics, State University of New York at Buffalo.

Next, Dr. Katherine McComas, associate professor, Department of Communications, Cornell University.

In the next chair is our consumer representative, Richard Colvin, from the Committee of Ten Thousand, and clinical assistant in medicine, Division of Infectious

Diseases, Massachusetts General Hospital.

Next is Dr. Roshni Kulkarni, professor and director of pediatric and adolescent hematology/oncology, Michigan State University.

Around the corner of the table is Dr. Col. Francisco Rentas, director, Armed Services Blood Program Office.

Next is our chairman, Dr. Frederick Siegal, medical director of the Comprehensive HIV Center, Saint Vincent’s Catholic Medical Centers, New York.

Next, in front of the podium, Dr. Donald Trunkey, professor, Department of Surgery, Oregon Health and Science University.

Around the corner, Dr. Simone Glynn, branch chief, Transfusion Medicine and Therapeutics Branch, National Institutes of Health.

Next, Dr. Blaine Hollinger, director, Eugene B. Casey Hepatitis Research Center, Baylor College of Medicine.

Next, Dr. Thomas Fleming, professor, Department of Biostatistics, University of Washington, Seattle.

Next, Dr. John Perez, regents professor and director of the Natural Toxins Research Center, Texas A&M University.

Next, Dr. William Bower, Office of Blood, Organ,

and other Tissue Safety, Division of Healthcare Quality Promotion, Centers for Disease Control and Prevention.

Next there is an empty chair which will soon be occupied by Dr. Willarda Edwards, partner of Edwards and Stephens, Ltd., Baltimore, Maryland.

At the end of the table, I would like to introduce our non-voting industry representative, Dr. Celso Bianco, executive vice president, America's Blood Centers.

There are three committee members that could not attend today. They are Dr. Adrian Di Bisceglie, Dr. Maureen Finnegan, and Dr. Andrea Troxel.

I would like to welcome and thank all the members that are here.

I would now like to read into the public record the conflict-of-interest statement that is required to be part of this meeting record.

The Food and Drug Administration is convening the July 20 and 21 meeting of the Blood Products Advisory Committee under the authority of the Federal Advisory Committee Act of 1972. With the exception of the industry representative, all participants of the committee are special government employees (SGEs) or regular federal employees from other agencies and are subject to conflict-of-interest laws and regulations. The following information on the status of the advisory committee's

compliance with the federal ethics and conflict-of-interest laws, including but not limited to 18 U.S. Code Section 208 and 712 of the Federal Food, Drug, and Cosmetic Act, is being provided to the participants at this meeting and to the public.

FDA has determined that all members of the advisory committee are in compliance with federal ethics and conflict-of-interest laws. Under 18 U.S. Code Section 208, Congress has authorized FDA to grant waivers to special government employees and regular government employees who have financial conflicts when it is determined that the agency's need for a particular individual's service outweighs his or her potential financial conflict of interest.

Under Section 712 of the Food, Drug, and Cosmetic Act, Congress has authorized FDA to grant waivers to special government employees and regular government employees with potential financial conflicts when necessary to afford the committee their essential expertise.

Related to the discussions at this meeting, members and consultants of this committee have been screened for potential financial conflicts of interest of their own, as well as those imputed to them, including those of their spouses or minor children, and for purposes of 18 U.S. Code Section 208, their employers. These

interests include investments, consulting, expert witnesses' testimony, contracts, grants, CRADAs, teaching, speaking, writing, patents and royalties, and primary employment.

For topic I, the committee will review proposed strategies to demonstrate the effectiveness of new coral snake antivenoms. This is a particular matter of general applicability.

For topic II, the committee will discuss alternative clinical and surrogate endpoints for evaluating efficacy of alpha-1 proteinase inhibitor augmentation therapy in alpha-1 antitrypsin deficiency. This is a particular matter of general applicability.

In addition, the committee will hear updates and informational presentations on several topics. These updates and presentations are not for discussion by the committee, and therefore committee members were not screened for financial interests relating to these presentations and informational updates.

Based on the agenda and all financial interests reported by members and consultants, no conflict-of-interest waivers were issued under 18 U.S. Code 208(b)(3) and 712 of the Food, Drug, and Cosmetic Act.

Dr. Celso Bianco is serving as the industry representative, acting on behalf of all related industry,

and is employed by America's Blood Centers in Washington, D.C. Industry representatives are not special government employees and they do not vote.

With regard to FDA's guest speakers, the agency determined that information being provided is essential. The following information is being made public to allow the audience to objectively evaluate any presentation and/or comments:

For topic I, Dr. Alejandro Alagon has association with a firm that could be affected by the discussions.

For topic II, Drs. Mark Brantly, Kenneth Chapman, Asger Dirksen, and Robert Sandhaus have associations with firms that could be affected by the committee discussions.

In addition, there may be regulated industry and other outside organizations' speakers making presentations. These speakers have financial interests associated with their employers and with regulated firms. The FDA asks, in the interest of fairness, that they address any current or previous financial involvements with any firms whose products they wish to comment upon. These individuals were not screened by FDA for conflicts of interest.

This conflict-of-interest statement will be available for review at the registration table. We would like to remind members, consultants, and participants that if discussions involve any other products or firms not

already on the agenda for which the FDA participant has a personal or imputed financial interest, the participants need to exclude themselves from such discussions and their exclusion will be noted for the record.

FDA encourages all other participants to advise the committee of any financial relationships that you have with a sponsor, its products, and, if known, its competitors.

Thank you.

Before I turn the meeting over to our chair, I would like to ask everybody to check their cell phones and either put them in the silent mode or in a vibrate mode. We really would appreciate that.

Dr. Siegal, I turn the meeting over to you.

DR. SIEGAL: Thank you, Dr. Freas.

I'm just going to welcome everybody to the annual summer-vacation meeting of BPAC and comment that I think it's really exciting that, for once, the FDA is really going to address the problem of snake oil.

With that in mind, let's start right away. We are going to begin with topic I, strategies to demonstrate effectiveness of new coral snake antivenoms. The introduction will be given by Dr. Hon-Sum Ko of FDA.

Agenda Item: Topic I: Strategies to Demonstrate Effectiveness of New Coral Snake Antivenom - Introduction

DR. KO: Mr. Chairman, committee members, and distinguished guests:

I'm Hon-Sum Ko, from the Office of Blood Research and Review. Today I'm giving you an introduction to topic I, approaches to demonstrating effectiveness of new coral snake antivenoms.

In this introduction, I will first go over coral snakes and envenomation by coral snakes. Then I shall discuss a little bit about the currently licensed product from Wyeth and then go into the potential produce licensure pathways, including considerations for licensure of a new antivenom. Then I will go over the questions for this morning.

This slide shows two of the coral snakes in the United States, the Eastern coral snake and the Texas coral snake. You can see the pattern of red bordered by yellow in these snakes, which gives rise to the rhyme, "Red on yellow, kill a fellow."

This is another coral snake in the United States, the Arizona coral snake, which has a similar kind of pattern, with red bordered by yellow, although this one is believed to be less poisonous than the others.

Here is a map showing the distribution of these coral snakes in the United States. We have the Eastern coral snake over a pretty large area over the Southeast.

But it’s believed that 50 percent of the envenomations by coral snake in this country are in the state of Florida, and then also the Texas area, with the Texas coral snake, and then the Arizona coral snake.

So you see that the distribution is over a very wide geographic terrain. Yet the number of envenomations is estimated to be only 75 to 100 a year.

For coral snake venom, the most troubling constituent is a curare-like alpha neurotoxin that competitively binds to nicotinic acetylcholine receptors, causing weakness and paralysis. There are contributing factors also, with phospholipase, hyaluronidase, and possibly other enzymes.

The next slide shows the envenomation by coral snake, which is believed to be not very efficient, as approximately 25 percent of the bites are believed to be dry. There are a number of factors:

- The coral snake relies on fixed retroverted hollow teeth to gnaw.
- The venom duct is not directly attached to the fang, but enters a small cavity in the gum.
- The fangs must penetrate the skin long enough to deposit the venom around the teeth and into the wound, so the coral snake holds onto the victim to gnaw and inject the venom.

What happens after that? The venom injected by the snake is believed to be typically 2 to 6 mg dry weight. With special techniques to milk the venom out of the snake, some large snakes may be able to give you up to 28 mg dry weight of venom. But the adult human lethal dose is only 4 to 5 mg dry venom. Local manifestations usually are not very remarkable, even though there may be risk of infection, although that still is rare. The most troubling thing is the development of bulbar paralysis, which has variable onset times. It may ultimately result in motor-function loss, need for respiratory support, and potentially death.

Here I try to show you two large series of envenomation by coral snake, in the days before the current licensed antivenom product was available. You can see from this table the mortality from these studies of the untreated patients was 9 and 24 percent, with an average of 18 percent. But if we take mean weighted for variance of these studies, then the mean is about 15 percent, with a lower bound of the 95 percent confidence interval of 4 percent. These have implications for doing a clinical trial, because a rigorous clinical trial based on a 4 percent historic rate of morality would involve a very, very large sample size, which is very difficult.

Now we come to the currently licensed product

from Wyeth, antivenin *M. fulvius*. This is the only licensed coral snake antivenom in this country. It is an intact equine IgG. It has been licensed since 1967. It is no longer being manufactured -- and that's the reason why we are all here today -- although there are still in-date products available. The potency and other quality attributes are maintained in the marketed lot. We evaluate the test results so that we can approve extension dates for the product. Currently it has been extended to October of this year.

This product has been recommended as safe and effective by the Panel on Review of Blood and Blood Derivatives, based on a literature review and testimony to the panel. Its potency is such that each vial will neutralize 250 mouse LD_{50} doses of the venom, which is about 2 mg venom. Adverse events basically are serum sickness and immediate hypersensitivity. There have been no reported deaths from coral snake envenomation in treated patients since its marketing in 1967.

The package insert for the Wyeth product gives the following advice in the use of the antivenom:

- Provide supportive care and give the antivenom to patients that have bite wounds from a coral snake as soon as possible.
- If a patient with no obvious bite wounds who is

observed develops signs and symptoms of envenomation, give the antivenom.

- The recommended dose is three to five vials, but one may give even 10 vials if the venom dose is presumed to be high.

Now I come to the potential product licensure pathways. By law, we need substantial evidence that a drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the proposed labeling. "Substantial evidence" means evidence consisting of adequate and well-controlled investigations, including clinical investigations by experts qualified by scientific training and experience to evaluate the effectiveness of the drug involved, on the basis of which it could fairly and responsibly be concluded by such experts that a drug will have the effect it purports or is represented to have under the conditions of use.

Here is a pathway called accelerated approval. It is intended for serious or life-threatening illnesses. It allows the use of surrogate endpoints that would be reasonably likely to predict clinical benefit or on a clinical endpoint other than survival or irreversible morbidity. You note from the title of the regulation that since it's intended for life-threatening illnesses, the

really important endpoints would be survival or irreversible morbidity. That is why a surrogate endpoint that is reasonably likely to predict clinical benefit or some other clinical endpoint that is not on survival or irreversible morbidity would require a postmarketing study to verify the benefit and describe the clinical benefit, since there may be uncertainty of the relation of the surrogate endpoint to the clinical benefit or of the endpoint studied with observed benefit to the ultimate outcome. These postmarketing studies must also be adequate and well controlled.

Again, what is to verify and describe clinical benefit where there is uncertainty? Basically, the effect of the surrogate has to be shown to correspond to a favorable effect of clinical benefit. Again, as the title of the regulation implies, for a product to treat serious and life-threatening diseases that warrant accelerated approval, then a meaningful clinical benefit is considered to be survival or prevention of major morbidity.

There is a lot of talk about a more recent pathway, the animal efficacy rule -- the approval when human efficacy studies are not ethical or feasible, based on evidence of effectiveness from studies in animals. This animal efficacy rule is really to be implicated only if no other pathway is available. It requires the use of at

least two animal models that have been validated. You need good-laboratory-practice studies for effectiveness to support the efficacy. Again, similar to accelerated approval, using this pathway still will require postmarketing studies, such as field studies, to verify and describe the product's clinical benefit and to assess its safety when used as indicated, when these studies become feasible and ethical. For approval, there may be need to be potential restrictions to ensure safe use, and patient labeling.

Here we come to the challenges for doing clinical trials on a new antivenom. I'm sure the other speakers will be addressing these as well, so I will go very briefly.

First, study population:

- As I showed earlier, the patients are dispersed over a wide geographical terrain, and they are not pre-identifiable.
- Also you have only about 75 to 100 envenomations a year, so the target population is small.
- Some patients are drunk and confused.
- There is possible confusion with other snakes.
- Again, I mentioned earlier the potential dry bites -- a substantial proportion in the people bitten by a coral snake.

- Then there is the uncertainty of progression of venom toxicity.

What is an appropriate control to do a clinical trial on an antivenom? My colleague, Dr. Jessica Kim, will be discussing this later in the morning, so I'll just leave it, for the time being.

Possible solutions to these challenging issues:

- Dry bites and confusion with other snakes: The envenomation may be confirmed if there is a validated plasma venom assay.
- For uncertainty of venom toxicity progression, we may consider evaluations other than mortality, such as the need for intubation of ventilatory support, progression of any clinical manifestation, even development from no symptoms, and the changes in plasma venom levels, if there is a validated assay.

Here we try to propose a way to get licensure of a new antivenom:

- First, to establish a dose of antivenom for a clinical trial. One can determine the relative potency of a new product by comparing with the currently licensed product in the inhibition of lethal challenge in an established animal model and then match the new product dose by proportional adjustment to the Wyeth antivenin labeled dose using this relative potency information.

- Knowing the dose, one can conduct a clinical study with the new product in a small trial, including evaluation of clinical outcomes, such as progression, intubation, ventilatory support, mortality, and so on, together with plasma venom measurements with a validated assay.

For an investigational product to address this problem, there is the possibility of having an expanded-access protocol to the needed antivenom. Such a protocol may provide valuable safety data, although generally it's not considered adequate and well controlled to support effectiveness because of the lack of measures to minimize bias.

For such a product, because it's dealing with a serious and life-threatening condition and may be addressing an unmet medical need, then it is probably eligible for regulatory assistance, including fast-track and priority review.

Finally, I come to the questions for you:

First, is a clinical trial to address efficacy of a new coral snake antivenom feasible and practical within a suitable timeframe, using the licensed product as an active control or using a modeled historical control of no antivenom treatment?

Second question: If your answer to the first

question is no, then would you agree that the following data may be reasonably likely to predict clinical efficacy of coral snake antivenom?

- First, to determine a dose, using the relatively potency that I described earlier, one can use venom neutralization of the new product against the licensed product to inhibit a lethal challenge in an established animal model; then basing the dose of the new product on a proportional adjustment of the dose of the licensed product with the relative potency information.
- With the dose determined, do clinical studies on a small number of envenomated patients treated with the new product, such as 10 or more, using an estimate of improved clinical outcome compared to the historical controls with no treatment, using, for example, a point estimate of 15 percent mortality and, at the same time, a consistent decrease in venom levels after treatment.

Of course, we would also need pharmacokinetic and safety data -- and these can be from normal volunteers -- before marketing, and then a postmarket study would be required to confirm the product's safety and clinical efficacy. The postmarket study would take place as a continuation of the pre-licensure clinical trial that has used the, for example, 15 percent mortality. But as you accumulate more patients, then you may be able to meet a

lower mortality as the point estimate.

That is what I have right now. Are there any questions from the committee?

DR. SIEGEL: Thank you, Dr. Ko.

Questions for Dr. Ko at this point?

DR. GLYNN: Could you clarify the difference between 2(b) and 1(b)?

DR. KO: You are talking about the pathways?

DR. GLYNN: No. I'm talking about the questions.

Maybe I can rephrase my question. You are saying that if you answer no to 1(b), then one of the possibilities is to do 2(b). I'm just asking, what is the difference between 1(b) and 2(b)?

DR. KO: For 1(b), essentially we are referring, really, to a historical control trial that is fully powered to be addressing a point estimate, for example, using a lower bound of 95 percent confidence interval, such as I showed earlier, with 4 percent or something like that. Then for 2(b), we are going to try to use, for example, a point estimate of 15 percent, such as I showed earlier in the table, weighted mean for variance from historic data, instead of a really rigorous way of using, say, 4 percent of the lower bound of 95 percent confidence interval, which may be the ultimate clinical outcome we would expect with a postmarketing trial.

That is why I said the postmarketing trial can be a continuation of the premarketing trial. You can look at it that way.

DR. SIEGEL: Dr. Epstein?

DR. EPSTEIN: The difference here is statistical rigor. In 1(b), what we are saying is that you either have a prospective controlled trial against licensed product. I'm going to talk about the issues there. In the other case, you model the control historically, but, as Dr. Ko explained, you then have a confidence range down to a 4 percent mortality. That would require a very large trial, to show that at 95 percent confidence, mortality with the new product is no greater than 4 percent.

In 2(b), what we are saying is that we will accept a small clinical trial -- as few as, perhaps, 10 patients -- as long as the point estimate of mortality is no worse than the point estimate from the combined historic studies. So you are giving up statistical rigor, but we still think that, accompanied with the other modeling, that may be a way forward. But that's the very thing we want to discuss.

The simple answer is, the difference between 1(b) and 2(b) is abandoning statistical rigor.

DR. KO: There will be more discussion later by Dr. Kim, our statistician, on this issue.

DR. SIEGEL: Maybe what we ought to do, instead of taking a lot of questions so early on, is to get through the program and then come back. There are lots of questions, and I'm sure that Tom Fleming has lots of answers.

Let's proceed with Dr. Craig Kitchens, from the University of Florida. He is going to talk about "Coral Snake Envenomation: Pathogenesis, Clinical Effects, Morbidity and Mortality and Distribution."

Agenda Item: Coral Snake Envenomation: Pathogenesis, Clinical Effects, Morbidity and Mortality and Distribution

DR. KITCHENS: I want to thank the chairman and the committee and BPAC for allowing me to come up here.

I'm going to try to give you some enlightenment to the clinical syndrome that we see with coral snake envenomation. Through the tale I'll weave, there will be ample opportunity to understand why confusion rules the day and will certainly make our task of proving that anything works and how to use it harder to do than, say, with pit viper envenomation, where you can usually tell that someone has been bitten very easily.

I'm from the University of Florida, and the University of Florida is interwoven with the city on this event here. A very old study, a paper written here in the

American Naturalist, was whether or not these snakes -- and I'm going to use that generically -- described as a colorful snake, long and thin, with no shoulders -- that's about all we had to define it in those days. We didn't email pictures back and forth 100 years ago. There was debate about whether this was a poisonous reptile or not. There was a man here, Mr. Schindler, at the United States National Museum, who was bitten on the index finger of the left hand so firmly that the finger had to be pulled off. I think they meant the snake was pulled off, not the finger. This bite was sent from a live specimen from Gainesville, Florida to the museum, because I guess someone said, "Send me some of these snakes."

Now, it turned out, if you read the fine print of this thing, that the snake was put in the snail mail and, I think, arrived to Washington about two weeks later. So the snake was somewhat irritable when it came.

Four hours later, Mr. Schindler was drowsy or unconscious, symptoms which lasted for a couple of days. A Dr. Taylor from Washington treated him with sodium bicarbonate and bismuth subnitrate, which I don't think we are going to study, a teaspoon every five minutes for six doses. He didn't get any better. So they reverted then to ammonia and French brandy -- at least one of those products is harming the other, it seems like -- a teaspoon every

five minutes for five or six or seven doses. The patient felt better by the third day. I'm happy for that.

The authors conducted a review of records they could find from the surgeon general's office in this town. They also wrote to several Texas physicians. We'll hear from those. One Texas physician noted that the lack of swelling and discoloration, which are characteristics or Texan and other pit vipers -- likened these symptoms more to those of sea snakes, a very good observation. From Gainesville, Florida, Professor Baird (phonetic) noticed that there might just be two kinds of snakes that we are calling coral snakes, one poisonous and one not -- an astute observation.

True, who is the author of this paper, concluded that coral snake bites were poisonous, but rare due to -- his words -- the lack of abundance of these serpents, which is partially true. They are very secretive little animals. They like to come out most in the morning about 4:00 or 5:00 or 6:00 a.m., when most of us are either just coming down from our drinking sprees or just fixing to get up and go to work. They have a sluggish disposition. Stories of mothers cleaning out their children's pockets and finding a live coral snake that has been in the pocket for several days without being bitten are not only credible, but true. They are a very small size in their mouth, as opposed to

the large gaping maw of pit viper. His conclusion was that the general notion that they are harmless is erroneous.

These are they. This is a pair of them. I want you to notice something, because we're going to come back to this. There are three colors: black, yellow, red. Black, yellow, red. That's on all these and all their mimics. But it's just the order.

There is the mimic. This is a scarlet king snake, which I think suffers negatively in my home state from mimicry. When I was a young man, this was a very, very common snake. It's friendly. It doesn't come out at 3:00 or 4:00 in the morning. It plays all day. It gets stomped on all day. So it has gone from a very common snake to almost an endangered-species snake. It's an ideal pet. It's a fun little snake to have. But it should not be confused with that guy.

These are more scarlet king snakes. You can see red, yellow, and black. There is not any yellow on this guy.

There is a very lovely article, which we reference here, in *Science* from about 20 years ago on mimicry -- how all these snakes, which are not kin to each other -- it's not like evolution or anything -- snakes somehow -- their evolutionary pattern figures out that if you look dangerous, some people may regard you as that.

The question is, how does anybody learn that? Does a mother dog have to teach her pups not to play with that snake because it might bite you? How does she learn that without dying? They ask a whole bunch of interesting questions of how this came to be. Our goal today is not to talk about that, but that is a very, very interesting reference, if anybody wants to read about it.

The dangerous pair -- that's poisonous, not poisonous; poisonous, not poisonous; poisonous, not poisonous, in Central America.

This is our range of snakes. We have already seen those. I did something which came out -- at least one of my sets of statistics came out right. If I take all these states, except for Florida, and add up all their population, I get a big number. If I add Florida to it, with all of its 18 million people, we get 36 million people, of which 18 million are in Florida, so half are in Florida. Notice -- because other maps that are done on this will show, like, for pit vipers and moccasins that there is a carved-out area in the bay area, in the Miami area, in the Gold Coast area, because those snakes don't like to be around populations and people and construction. Coral snakes could care less. They are out in the front yard all the time. They are happy living down deep in the grass. So they are not excluded, like most of the pit

vipers are, from high-density populations.

The North American pit vipers, just in case you don't know, are very complex, having 30 and 40 different enzymes inside of them. They cause almost immediate pain, swelling, and discoloration -- that triad -- within minutes in the vast, vast, vast, vast majority of bites. Most damage is local and not systemic.

The problem with an ER doctor-type thing is, if he or she sees 20 snake bites in their entire career in the Southeast and the Southwest part of the country and maybe 1 percent is from a coral snake, then you can see that there is a diagnostic bias toward, "I've seen snake bites, and they all swell and get discolored." Here's someone who doesn't have swelling and discoloration; ergo, it's not a poisonous snake. We have seen some grave errors that way.

For the coral snake bite, there are no local signs or symptoms. I have not seen any redness and swelling, except when you ask the patient later on, "What did you do when you realized the snake had bitten you and you were in transit?"

"I was sitting there just rubbing the heck out of it, rubbing it and rubbing it and rubbing it, squeezing it, squeezing it, squeezing it."

I think that's what causes the vast majority of any redness or a little bit of edema.

It’s alarming how these folks can come in and be -- they have been drinking or smoking a little pot or something like that. Euphoria is not necessarily an alien or dangerous symptom to their way of thinking. But then they get diplopia and dysarthria, airway troubles, aspiration, and can to from almost being perfectly intact neurologically, except for some inebriation, to needing to be intubated.

The sign I look for the most and I teach my residents is dysarthria. If you are totally drunk, you have slurred speech, but when you talk like this, it sounds like someone with ALS, because they have palatine paralysis.

This is just a foot from a very mild pit viper bite. You can see pain, swelling, discoloration. You can’t see the pain, but, trust me, it’s there. It feels hot, fiery.

This is a very, very, very, very serious bite. I’ll read you her case in a moment. Notice that it’s 8:20 in the morning, so she gets bitten in the morning, because she is out painting something. This is a bite that was very, very serious. She was bitten right here. What do you see? Nothing. Not a thing. If you strain your eyes and close them, you might see a tiny bit of redness here. But, trust me, that’s from that rubbing of that.

In our series that we wrote up, 20 years or so ago, of 39 bites, we got all these variety of symptoms. None of them are particularly specific or sensitive enough to really hang your hat on. We want to avoid the neurologic signs and symptoms, because reversing some of these neurologic events is very difficult with out treatment now, because of, probably, the high affinity of the toxin for the nerve end plate. You almost have to have fang marks to get inside the ER when you come in with a snake bite.

We show here that it takes about two hours before you get to our emergency room. Our emergency room is usually a step up. They go to a smaller emergency room in our state and then come to us.

This is who we treated. We treated a lot of people. I'll come to a concluding slide on that momentarily.

Let's just go through five cases. We describe 35 cases, but we wrote out in a little longer hand several of them.

This is a house painter, who was going to be the brother of the sister who was painting the house who had the very bad bite on her left hand. He's painting a house. They are doing it in the hot summer. You do that stuff at 6:00 and 7:00 in the morning before it gets hot. So they

are out there painting the outside of the house. She reached up under some azalea bushes with her paint brush and came out with this guy stuck on here. They were both on opposite sides of this small house. So she starts yelling and hollering, “Help me! Help me!”

He hears this on the other side. She decides to go running to him and he decides to go running to her. Unfortunately, they did run in opposite directions. So he’s running around the house and she’s running around the house with a snake stuck on her. Then when he finds her and pulls the coral snake off her hand, it literally turns around and bumps him in the hand. So there is no engagement. There is nothing that goes on. He hits this calloused finger, bounces off, and then gets stomped to death.

We observed him for 24 hours, because we were looking at his sister. First of all, if the risk of the snake bite -- because it wasn’t a bite -- was zero, then you can’t afford any risk from the treatment. Trust me, it’s considerable.

Here’s a different one. This is the next elevation up. A 15-year-old boy reached under a log at a state camp and, withdrawing his arm, found that a snake was attached to his left ring finger. He recalls shaking his arm for four or five seconds before the snake let go. When

he arrived at the emergency room 90 minutes later after the accident, the vital signs were normal. Fang marks from which a small drop of blood could be expressed were found on the fingertip. We are going to use this as a sentinel sign. In other words, you don't see a hole, like you do with a pit viper, which injects a classic hypodermic needle in you. You can't see anything. But if you press on the finger, a small drop of blood comes out, so you can deduce that a hole is there. The hole is there, large enough to allow the blood to come out -- theoretically, large enough to allow venom to get in. So we use that as a key point in making our risk analysis.

He complained of only minimal tingling of his finger. It was interpreted as a local sign. There were no systemic signs. Skin testing, which we did in those days, with horse serum was negative. Five vials were given intravenously during the next two hours without any untoward reaction. Nothing happened. We watched him for a day.

So here's a guy that had a positive identified coral snake, a hole in his finger, no symptoms. He gave the history of the snake hanging on in order to chew for an estimated five or six seconds.

Third case: I remember this guy. A 36-year-old man had been working on a horse farm. There is a lot of

horse farming in Ocala, right south of Florida, where they raise national champions. He had been in that industry all his life. He was admitted to having drinking [sic] beer with some colleagues. While at the farm, the group observed a snake. They all agreed it was a large coral snake. A wager was placed -- usually with the ante being whatever beer is left inside of the cooler -- of which person would handle the reptile. This patient stepped up to the plate, picked up the snake, and was bitten on the right index finger. The snake held on with a chewing motion, as was described.

Now, the average person will say, “Why did you let this thing do this, continue chewing?” You don’t ever get a good answer.

When the snake was pulled off by a friend after several seconds, the feeling of separating layers of Velcro was experienced. That’s another similar thing, that there’s this feeling of Velcro being pulled as this series of small teeth, which almost all amphibians and reptiles have, is pulled out of the skin. That means that there was engagement at least through some layers of dermis.

The patient came to the emergency room five hours later. He was grossly and obviously inebriated. His blood alcohol level was 52 micromoles, which is 238. You can’t drive in any state I know with that. The only physical

finding was minimal swelling of the finger. He had been rubbing it. Fang marks were present from which small drops of blood could be expressed. He had no systemic symptoms. A skin test was negative. However, after receiving the first drops of antivenom, he had the only honest-to-God anaphylactic reaction I have never seen to snake venom. I have seen lots of reactions, but this was the full-court press. So he got three drops of antivenom, out of a diluted, probably, liter or so. So we had to follow him, kind of natural history. He developed shock, responding to fluids, epinephrine, dopamine.

Despite the severity of the bite, it was decided that no additional antivenom would be administered. He looked pretty good, after he was resuscitated. Within the next several hours, he developed diplopia, dysarthria, generalized fasciculations. He was intubated electively, endotracheal. The reason we do this -- which I'm worried about this being one of the criteria, we go forward -- is that we don't want to wait until someone can't breathe to intubate them, but we do it because, since they can't swallow because of the pharyngeal muscle paralysis, we do this prophylactically, to prevent people from having aspiration pneumonia.

He never did require mechanical ventilation. He developed total body weakness, being unable to move any

muscles, with the exception of his diaphragm and his hands, with which he communicated. Left was yes and right was no, I guess.

His serum creatinine came to 18,000. Six days later, his fasisculations abated and his strength returned. He was intubated for about six days.

This is the painter, the 24-year-old woman who was painting the house and had the chewing for about 12 or some-odd seconds as the race ensued around the house. She remembers very well the extraction, Velcro-type feeling. She had intense paresthesias up to her elbow. Testing with horse serum was negative. Because of the length of the snake, prolonged exposure, and early symptoms, she was regarded as probably having a major envenomation. Accordingly, 10 vials of antivenom were -- this is the only time I have never given 10, I think -- over the next three hours. However, the paresthesias progressed to her upper arm. Five more vials were administered. I don’t know whether it was necessary or not. Nevertheless, intense paresthesias progressed to her sternal area, and five more vials, which all were given within 12 hours of the accident. She never had muscular weakness or paralysis.

So I think the stuff may work. I don’t know. I don’t know how much she needed.

This was another scary. A 19-year-old woman was

in a suburb of Ocala when she saw a snake. Because she thought it was a scarlet king snake, she picked it up. She was bitten in the webbed space between the third and the fourth fingers of her left hand. She said the snake held on for 20 seconds while it made chewing motions.

Again, I never got a good response to, “Why did you let a snake gnaw on you for 20 seconds?”

I did get a good answer, if it comes up in the question-and-answer period later on, for one that’s written that we saw after this.

She went to the ER -- that’s the bottom line here -- and the doctor in the ER said, “You’ve been bitten. I don’t see any swelling. There’s nothing wrong with you,” because there were no local symptoms. She went home and just quit breathing. Taken to a hospital. Then she was intubated. We gave her a lot of antivenom. But she was already paralyzed, so we couldn’t get anything better. So she was paralyzed and on a respirator for five days. It took about six weeks for her to get normal muscle strength.

So we have seen all these -- now, the one thing in common with all these, these are all documented coral snake bites. We’re not talking about mimics here or, “Mama, a pretty snake bit me,” from an eight-year-old. That’s what’s scary. A lot of snakes are pretty. You don’t know what to do.

We'll digress here briefly. When we wrote this thing in the 1980s, we got some letters to the editor. I'll read those, who can't read it. This is, "Since erroneous identification of the Eastern coral snake was the most common situation resulting in snake bite in Kitchens and Van Mierop's article, an old Boy Scout dictum would have been helpful to the victims. The photographs of the Eastern coral snake and the scarlet king snake were quite helpful, but the dictum referring to the contiguousness of the colored bands on each side is a pearl: Red and black, friend of Jack; red and yellow, kill a fellow."

Well, both snakes have red and yellow, and red and black. So this guy got it wrong. In deference to my Harvard colleagues -- now, I didn't point that to the editor in my response.

Then the same thing. Someone from Nairobi, where they don't have a whole bunch of coral snakes said the same thing. I replied, "We too are aware of the dictum mentioned by Drs. Murray and Worsten. Nearly everyone in the area where coral snakes are endemic is aware of the dictum or its modifications thereof. However, it has not proved to prevent victims from being bitten by a snake. We offer two explanations for this failure. Often patients are" -- for "often," read "always -- except for the painter, because she's a lady -- "often patients are

inebriated and don’t seem to remember the rhyme correctly.” These were both wrong. Now, if you said “red on black, touching black” -- but just “red and yellow” -- they are all that way.

Number two, it becomes altered, something like “red on yellow, friend of a fellow,” something like that. In other words, it gets wrong. Our experience is that the dictum, or at least the recitation of it at the time of confrontation with the snake, is imperfect. Accordingly, we advise people to simply avoid any snake that looks like these, which would actually solve our problem. If we ran out of antivenom, then can’t we just pass a law that you can’t get bitten by a coral snake, and then we can all go home.

But, unfortunately, the practice continues.

So why is this hard to do? Why are we giving this a rough time?

- Snake identification is difficult because of the mimics.
- Inebriation -- not just alcohol.
- Delayed ‘fessing up. This happens a lot in teenagers. The guy saying, “My mama told me, don’t play with snakes, and here I went out and did it.” So they’ll wait around for hours and hours. You see this with pit vipers, too. I said, “How did this leg get this big?”

- Children -- I have absolutely given up on showing pictures to children. Anything that's long and skinny with no shoulders, they'll say, "That was a snake, mom." You can show them a mamba. You can show them a krait. You can show them anything you want -- "yes, yes, it's that one." So I have given up on that.
- No local signs of envenomation, compared to the pit viper.
- No surrogate tests are readily available. We can deduce, with very, very high CKs, that people have been bitten by canebrake rattlesnakes. We can deduce by hypofibrinogenemia that people have been bitten and envenomated by certain rattlesnakes.
- We probably have a fairly high mortality if we are wrong. That's the problem with the little eight-year-old kid that says, "Mommy, this pretty snake bit me." You don't know if he has been bitten or not. You sit there and, in your mind, you see this kid on a respirator about eight hours from now. It's easy to pull that trigger while he is doing well, because we can't reverse this very well.

There was a U.S. death reported just a few days ago, where three migrant workers were playing around by the camp and again imbibing. This victim did not go to the hospital, and within five hours, he was witnessed to have a seizure and stop breathing. His buddy went to the hospital

and got therapy. Autopsy revealed no other cause of death. The patient's serum was found to be positive for coral snake venom on ELISA testing. This was the first death since the Wyeth product became available in 1967.

Our triad for exposure to coral snake venom is:

- We like to see a confirmed coral snake. Some people don't bring the snake in. Some people bring it in alive. Some people bring in little pieces.
- Punctate wounds consistent with a bite site with expressible blood. You don't see anything, you press on that area, and you see a little -- like when you want to donate blood and they check first your hemoglobin. They stick you and say, "Aha, you missed," and then they squeeze, and here it comes up.
- Then the history of hanging on several seconds, preferably with a history of the Velcro separation phenomenon.

We like to have two of those. If we don't have it, we often will wait and see.

Some fuzzy math. We have problems with reporting bias, referral biases, because people in a small hospital -- they know it's not a coral snake -- are not likely to send in a person, our personal experience, and some of our best guesses.

So the number of coral snake bites per year has

already been alluded to be right around 100. I would estimate that the number of coral snake bites due to true coral snakes would be about 50 percent of that. We still think in our area that about half the snakes that are judged to be coral snakes are not coral snakes. The percentage of those true coral snakes -- true exposure to be about 25 percent. The lethality of exposed patients treated with Wyeth so far has been zero. The lethality of patients exposed, hospitalized without antivenom treatment -- I submit my two guys, the one with anaphylactic reaction and the one with the delayed apnea -- is zero. But the lethality of exposed patients who are never treated, by the natural history, we have heard maybe 15 or 20 percent. I'm guessing a little higher, because I have tried to cull out the non-coral snakes and the non-coral snake envenomations, leaving roughly 25 percent, which then 25 percent negative. Four times enrichment of something that had a lethality of 15 percent, you could see, would get you right about the same number.

That's where I am. I'll stop at that point. We'll do questions now or later or whatever the chair wishes.

DR. SIEGEL: Thank you very much, Dr. Kitchens. That was a very nice discussions.

Next we will hear from Dr. Steve Borron, South

Texas Poison Center, “Challenges for Clinical Trials.”

Agenda Item: Challenges for Clinical Trials

DR. BORRON: Good morning, Mr. Chairman, committee.

This gentleman on the left of the slide is the reason that clinical trials will probably be possible.

There are major challenges to performing a clinical trial of coral snake antivenom. I spoke to a smaller group in February, and my conclusion at that time was that it would be almost impossible to do a comparative trial of a new antivenom versus the Wyeth antivenom because of the large number of patients that would be required. But I have come to a little bit of a different conclusion, based on some of the background material we were supplied and some more reading.

I will go through these major challenges, some of which have already been mentioned -- the low incidence of bites, the geography and the time logistics, differences in toxicity among various species, improvements in supportive care, study power and sample size necessity, cost/benefit ratio, existing importers of foreign coral snake antivenom, and multiple foreign manufacturers. Then I’ll talk about some possible solutions.

We have already heard that there are fewer than 100 coral snake bites per year. Using the American

Association of Poison Control Centers data, I estimated that there are about 82 bites that are reported to poison centers. Now, they don't get all the bites reported to them, because some physicians feel they can treat them without assistance and don't call. But that includes all of the snakes across the southern United States that you saw on the map there. For the Eastern, Texas, and Arizona coral snakes, there really are not a lot of bites.

That said, most of the bites that do occur -- probably 50 percent of them occur in Florida and another 30, 40 percent in Texas. We average about 20 to 25 bites per year in south Texas.

You have seen this map already. Over on your left are some data from the American Association of Poison Control Centers. You can see that over the last three or four years, where the reports are out, there are fewer than 100 case mentions of coral snakes, very few major outcomes, and no deaths. I suppose that this recent death that was reported by Norris will end up in the 2010 report.

Looking over at the right, you see some data from Texas. Over the last nine complete years, we have had 217 coral snake bites, 10 of which were severe -- about 5 percent. Only about a third received antivenom. I think that's a point that will be important in considering a clinical trial, because there are clinicians, including

toxicologists, in Texas who don't believe that *Micrurus tener* requires treatment with antivenom.

This is kind of a summary of what we see in Texas. You can see that rattlesnake bites and copperhead bites largely outnumber coral snake bites. While things are spread out quite a bit, 22 percent of our bites come from Harris County, which is where Houston is located, and 50 percent of the bites are in Harris (that's Houston), Bexar (that's San Antonio), Travis (that's Austin, Texas), and Oasis (Corpus Christi). So about half the bites are in fairly concentrated areas.

This is the distribution on the map. The farthest west is out here in Ector county. There has been one earlier in the Sanderson area. But this is a recent 10-year map of where bites have occurred.

There are problems of geography and time. Many of these bites occur in rural areas. As Dr. Kitchens has already mentioned, the participants in these activities who are getting bit are often drunk. The bites are not particularly painful. They don't show any swelling or discoloration. So the person who is bitten often does not seek medical attention until hours later, when they start having paresthesias or other symptoms that suggest that things are going wrong.

This is an issue in any kind of clinical trial.

It is our supposition that the antivenoms do not work as well once clinical symptoms have installed. So there will have to be some consideration of how many hours out after a bite would be an acceptable timeframe to go ahead and administer the antivenom.

In terms of logistics, with only 100 bites or so -- 80 to 100 bites -- in order to have sufficient power to detect a benefit, almost every bite that occurs would need to be included in a study, to be able to complete the study with sufficient power in a year's or two years' time. This implies having rapid availability of the antivenom to literally hundreds of hospitals spread out across the Southeast and Southwest. Many of these hospitals will have never seen a bite. Many of them do no clinical research. So setting up a clinical trial in some of these hospitals and having adequate coral snake antivenom supply to have it available in a short period of time is definitely problematic.

It further implies the need for a very, very large number of investigators who would be available 24 hours a day and who may never see a bite. To those of us who publish, not having your name on the paper after participating is not a great incentive for participating in a clinical trial.

If we assume that there is an investigator in

every hospital, then the number of IRBs that would have to be involved -- that's institutional review boards, for those who don't do clinical trials -- would be potentially enormous. Agreeing on a unified protocol between multiple IRBs is always difficult, and between hundreds of IRBs, it might be even worse.

The next thing that we come to is species variable toxicity. All coral snakes are not alike. The Eastern coral snake, as you have already heard, was responsible for pre coral snake antivenom mortality of greater than 10 percent, perhaps 15 or 20 percent.

You have already heard about the one death that has been reported in the last 40 years, following the availability of Wyeth.

There have been no deaths in Texas since 1883, when True published his paper. Two of the deaths in True's paper came from Texas, but no substantiated deaths due to coral snake envenomation have occurred in Texas since. So we are dealing with a very low lethality in the Texas coral snake and a significant lethality in the Eastern coral snake. The Arizona coral snake, to my knowledge, has never been responsible for a death and is viewed as being even less toxic than the Texas coral snake. I can tell you that Texans are generally proud of having the most toxic and mean animals, so if we can trade with Florida, we are

willing to do that.

The differences in toxicity clearly have implications for a clinical trial. As I mentioned already, some clinicians in Texas don't believe that coral snake antivenom is required for Texas and Arizona coral snake bites, unless the patient develops definitive neurological findings. So even if a drug and a trial were made available, some clinicians might choose not to participate because they don't believe it's necessary. The low lethality of Texas coral snakes supports that.

The differences in toxicity between the Eastern and Texas coral snakes could also cloud the recognition of any beneficial effect. If you don't expect any major toxicity or death in the first place, it's going to be very difficult to see a reduction.

But on the other hand, if you were to do a clinical trial and eliminate the Texas and Arizona coral snakes, you eliminate maybe a third of the bites, which is going to increase the amount of time that is required to complete any kind of clinical trial.

I will not pretend to be a statistician. I will leave the power calculations to those who follow me. But just looking roughly, the death rates in the pre-Wyeth era actually varied quite a bit. Wilson, in 1908, reported eight bites that he could catalogue in the United States,

and six of those died. Neill, in 1975, reported on 21, of which 19 percent died; Parrish, in 1967, on 11 and 9 percent died; Findlay Russell, from Arizona, in 1980, 21 percent death rate. Dr. Kitchens, who is here, didn't see any deaths in his series, but he had a 10 to 15 percent rate of serious symptoms. Again, this is in the post-Wyeth era.

Just doing some rough statistics, if we took a pre-Wyeth mortality of 15 percent and we assumed that the Wyeth product reduced the mortality to 1.5 percent and we assumed that any new coral snake antivenom is as effective as Wyeth, if we do a quick power calculation, a single-arm study would require only about 26 subjects. That's to see this huge difference. In other words, we would expect no mortality with the new product. If you tried to compare the Wyeth product to a new product and see a 25 percent reduction in mortality, you would require thousands of patients. It's clinically impossible.

In terms of cost/benefit -- and this is certainly something that the pharmaceutical firms are going to look at -- if our baseline serious toxicity is 15 percent and antivenom treatment lowers it to 1.5 percent, we are saving two patients a year. You have to ask, what is the net worth of this? To those two patients, it's priceless. But if there are 82 cases, and even if the manufacturer can

manage to pass off a price of $5,000 per treatment, their income on that treatment is $410,000 a year. If we assume that only about a third of those are actually going to receive the antivenom, then they would make the enormous sum of $137,000 a year, gross.

What's it going to cost to do the studies that are necessary for approval? What are the liability risks of new vaccines for the manufacturer? I would ask, what is the incentive to seek approval?

These are all things that I think will be considered by the manufacturers.

Finally, we have groups in the United States, particularly in Florida, that operate under what is called the zoo IND, which allows them essentially to import foreign antivenom for compassionate use for zookeepers, et cetera. Particularly in Florida, they make these available not only to zookeepers and animal handlers, but to people in the community, their hospitals, et cetera. What incentive will they have to participate in a clinical trial? Would they give up their current stocks of competitive products if a group decided to study one particular coral snake antivenom in the United States and they have another one available?

Theoretically, more than one manufacturer might wish to submit an IND. If this happens, this increases the

sample size necessary for a single study in order to be able to tell a difference between the two or implicates competing additional studies, which may confuse the picture. A study with sufficient power to distinguish the efficacy of two CSAV products is doubtfully feasible. However, we might see differences in safety, and that might be a useful exercise.

Some possible solutions: I would propose that the best way to attack this large geographic area and small number of patients, if you are to do a clinical trial in the U.S., is to use poison centers as study centers and their medical directors as principal investigators.

Parenthetically, I now live in El Paso, so I have no vested interest in this. We have lots of rattlesnakes, but no coral snakes.

The advantages of poison centers are:

- They are already under HRSA oversight and they report their data to the CDC, so they are used to dealing with the government.
- They currently receive the majority of reports of bites from hospitals.
- They have educators who could go out to the communities where these bites may occur and encourage the hospitals to call whenever they have a bite.
- Poison center medical directors are all board-

certified medical toxicologists. They understand the toxicity and the treatment of coral snake bites, and probably have more experience than any other group of physicians in treating coral snake bites.

You could team the poison centers with medical helicopter programs, making the flight personnel sub-investigators. You could rapidly transport coral snake antivenom to any hospital within 150 to 200 miles. You could potentially transport the patient to the hospital aligned with the poison center. If that were to occur, helicopters would probably require some payment for no transports, and there could be logistical issues with following up patients who received the antivenom and then either refused transport or, for whatever reason, were not taken back to the hospital of the principal investigator.

If you used designated poison centers as study sites, the IRB submissions could be reduced. Instead of having all the hospitals where patients would be treated put through an IRB approval, you can instead of IRB approval at the study sites and then have letters of agreement set up with the hospitals to allow investigators to recruit study participants in those hospitals.

That’s my impression. I’m not sure of all the legalities associated with that. It’s just an idea.

Unlike in the scenario that I described in the

beginning, in this case you could store the coral snake antivenom in a hospital pharmacy where the helicopter lies, which means you wouldn’t have to put it in hundreds of hospitals all over the region.

If, as was mentioned in the first talk, venom levels were determined to be a useful surrogate measure of efficacy, the flight teams could assure collection of pre- and post-treatment venom levels as a secondary outcome measure.

If you did consider a helicopter and poison center-based model, you probably still would want to have some storage of coral snake antivenom and a few additional investigators at the largest hospitals out away from the poison center, so that in the event of bad weather, you would have available coral snake antivenom that could be transported by police or other ground vehicles to the hospital that needed it, along with a handful of investigators who could go in and enroll patients.

This just gives you an idea of the times involved. This is San Antonio, Texas. Austin is here. Houston is over here. Our helicopters can be at the Gulf Coast in about 45 minutes. So it’s very easy to cover the entire south Texas area with one or two helicopter programs. Houston has one. We have one. There is one in Austin. There is one in Corpus. So it would be very easy

to have helicopters cover the area. The same is true for Florida. There are helicopter programs and poison centers dotted all throughout the state. There are three poison centers in Florida and six in Texas.

In conclusion, a human efficacy trial may be feasible if we compare a new coral snake antivenom with the historical pre-Wyeth mortality rate rather than with the efficacy of Wyeth. Again, comparing head to head, Wyeth with the new antivenom, in humans would require years and years and years.

Logistical issues preclude having a local investigator, local storage of CSAV, and IRB approval in every hospital that is susceptible to receive a bite. Using a hub-and-spoke poison center/medical helicopter-based study operation seems to make sense to me.

I thank you.

DR. SIEGEL: Thank you, Dr. Borron.

We'll next hear from Alejandro Alagon, M.D., Ph.D., of the Autonomous National University of Mexico, "Animal Models and Surrogate Markers for Coral Snake Envenomation."

Agenda Item: Animal Models and Surrogate Markers for Coral Snake Envenomation

DR. ALAGON: Good morning.

I'm going to speak about animal models and

surrogate markers for coral snake envenomation. I work at the Institute of Biotechnology at the National University of Mexico. I have a strong collaboration with the main antivenom producer in Mexico and, I would say, also on a worldwide basis.

I want to start by saying that in Mexico we have a big experience in using antivenoms, because we have a big need of a scorpion antivenom. In my country every year, more than a quarter of a million patients receive antivenom because of scorpion stings.

These are my introductory remarks.

Treatment of envenomation has depended almost entirely on the individual clinician experience in assessing the severity of envenomation. The efficacy of treatment is related to the neutralization potency of the antivenom used, the route by which it is administered, the dose, and its pharmacokinetics. The introduction of enzyme immunoassays has permitted a more scientific approach, allowing the estimation of circulating venom and antivenom concentrations at any time after the bite in patients or experimental animal samples, mostly blood.

I would like to continue with this very nice slide that was published by the group of the late Dr. Cassianbon (phonetic), in which they showed that the intravenous injection of antivenom induces an immediate,

complete, and lasting neutralization of venom components, as well as a rapid redistribution from the peripheral compartment to the vascular one. What they did was, they used rabbits as the animal model and they injected venom from this North African scorpion. This is the dose. But at the same time, they also radiolabeled a minute amount of venom so that they could follow the venom, following the radioactivity, as well as they could follow the venom using ELISAs.

This is the kinetics of the venom. The venom was injected at time 0. You can see that, more or less, the two curves run in parallel, and the maximum absorption levels in the blood were around 60 minutes. In this experiment, what they did was, they injected an antivenom, an F(ab) 2 antivenom, at time 60. This is really very interesting. The free venom levels immediately went down to zero and remained like that for the time they checked the rabbits. But what is very interesting is that you see that the labeled venom actually rose quite a bit -- this is a log scale -- after the antivenom administration, meaning that not only the venom in the blood was neutralized, but also that the presence of the antivenom in the blood was inducing a redistribution of venom from the peripheral compartments or from the site of the injection -- from those two sites.

Also from the same group, there is another experiment I want to show to you. It is the effect of bio-doses(?) of antivenom on blood venom levels. In this case, they were using viper venom, a European viper venom. Again, the model was a rabbit, and the venom was also given subcutaneously. Total venom means venom as measured by radioactivity and free venom means venom measured by the ELISA procedure.

When they gave small doses of antivenom, there was a small decrease in the venom, but then it starts to bounce up. When they increased the dose of antivenom, the venom went down to zero, but then again it started to build up. When they gave a much higher dose, the free venom drops and remains like that for that amount of time. Again, the same effect on the total venom concentration in the blood as in the case of the previous slide is also seen with this viper venom.

This is a very simple scheme that can help us to understand what the previous experiments tell us. Here we have the venom depot. You can think of it as being the bite or the injection site. Then you have the blood compartment and the tissue. This is the situation before antivenom administration and after antivenom administration. From the venom depot, in which you have a very high venom concentration, some of the it diffuses out

and is continuously diffusing out from the depot site to the blood. From the blood, it goes to the peripheral compartments and then it binds to the target molecule. Of course, depending upon the venom, there are venom components, for example, in the case of rattlesnakes in which some venom components -- their targets are located in the blood. But, for example, in the case of coral snake, the target is outside the blood. It is the neuromuscular junction, as you have heard several times this morning.

I just want to point to something here that sometimes we don't picture very clearly. Most of the venom remains in the site of the injection or the site of the bite for hours and even days. We have measured that on a large-animal model that I will talk a little bit about in a while. But I can say that, for example, if you give a subcutaneous injection or intradermal injection of venom and then, after several hours, you go and remove the skin and make a small square and extract as much venom as you can and then measure the venom in that lesion, what you find is that most of the venom is still in the site of the injection.

So if you give antivenom by the indicated route, which is the intravenous route, then you have a big antivenom concentration in the blood, and it binds very fast the venom components. In less than four minutes, all

the venom antigens disappear from the blood. Then, just by mass effect, what you may think is that the venom in the peripheral compartments actually starts diffusing out from the tissues and binds to the antivenom. Of course, there will be also antivenom in the peripheral compartment and eventually some antivenom will also reach the injection site. But this is just a very simple slide. This will promote the dissociation of the venom from the target, and also you will get venom coming out at a faster rate from the injection site.

There are a number of different ELISAs for measuring different venom antigens. This is just an example. What you have is a sandwich ELISA, in which you use a captured antibody -- which, for example, in this case, can be immunopurified antibodies -- then you incubate the sample, and the venom antigen is bound to the solid phase, and then you have a detection antibody, which in this case is rabbit hyperimmune serum, and then a secondary antibody, which is an antibody against this one labeled with an enzyme. What we usually get are ELISAs that can easily detect as low as 1 ng/mL of venom. Usually, with that sensitivity, you can follow all types of envenomations, except the blood envenomations, in which you have much lower venom levels. But, for example, these types of ELISAs have been very useful for measuring venom

antigens in scorpion patients, as well as rattlesnake patients.

This is a paper that we recently published with Dr. Leslie Boyer. It shows that the clinical resolution of a scorpion envenomation in children treated with antivenom correlates with the disappearance of venom antigens in plasma. Over here you have different clinical evaluations -- for example, percentage of patients with abnormal eye movements -- the open squares were treated with placebo and these were treated with antivenom -- and then patients with thrashing limbs and then patients with a abnormal respiratory distress, and the overall toxicity as evaluated by the investigators. Here is the amount of the sedative that was provided to the children. It shows a huge difference in the dose of midazolam that had to be used in the non-antivenom-treated children. These are the venom levels.

In this case, we took three samples -- baseline, one hour, and four hours. As you can see, after one hour, you don’t detect the scorpion venom anymore. The same thing at four hours. In the midazolam or the placebo-treated children, you still can find significant amounts of venom in the blood.

These are published studies in rattlesnake envenomation. Recurrent hypofibrinogenemia and

thrombocytopenia are associated with recurrent venom antigen detection in plasma.

Let’s start with this first. This is a patient treated with an F(ab)2 antivenom. F(ab)2 antivenoms have a higher persistence in the blood, as you can see here. In this patient, the venom level started here, and after the administration of the antivenom, it went down to zero and it stayed like that for many, many days.

In this case, this is an example of a patient treated with F(ab)1 antivenom. F(ab)1 antivenoms are removed very fast from the body, as you can see here. These are the antivenom levels. They drop very fast. Then the effect on the venom levels was the same as in the other case. It went down. But after several days, it started to build up again.

What is very interesting is that when the rattlesnake venom went up, then the platelets and the fibrinogen went down again, whereas in the other case, once they bounced up to normal values, they stayed like that.

Here, the recurrent antigenemia is associated with the recurrent thrombocytopenia and hypofibrinogenemia.

This is another study done by the group of Warell and Pigston (phonetic) in Nigeria, in which they followed patients bitten by *Echis ocellatus*, which accounts for more than half of the snake bites in Africa, in which they also

found the recurrent venom antigenemia associated with the fast disappearance of ovine Fab antivenom(?). So you have the venom levels going down and up, down and up. The triangles indicate here that they have to administer extra vials of this antivenom because they were having problems with the clotting of the blood.

I will focus a little bit more on animal models. You can think of small animals and large-animal models. All venom researchers or venom producers rely mostly on the mouse model to determine the potency for venom, as well as the potency of an antivenom. But usually these studies are carried out by pre-incubating the venom with the antivenom. Of course, these are very useful assays. But the problem is the extrapolation from what you find in such small animals to larger animals, such as human beings. This is where a large-animal model must come to rescue us, in trying to understand this better.

If you use a sheep, for example, you can take samples and you can measure venom/antivenom very easily. For example, the rabbit, which is not as small as a mouse, also allows you to make lots of blood sampling. But the problem with the rabbit is that they are very resistant to coral snake venom. Also a large animal can provide the means to do pathophysiological studies and do more pharmacokinetics and pharmacodynamic studies that are

closer to what happens in the human patient.

These will be the objectives of our large-animal model:

- To evaluate the efficacy of antivenoms as measured by the disappearance of venom antigens in blood.
- To determine dosage of antivenom as assessed by the continuous absence of venom antigens. That means no recurrence in antigenemia.

I now show you our results. This is the injection of a dose of 5 mg *Micrurus fulvius* venom subcutaneously. You see the venom going up and then the antivenom -- in this case, we injected three vials of Coralmyn at one hour -- it went down. But then, after 10 or 15 minutes, the venom antigens started to go up again, whereas when we injected 10 vials of the same antivenom, we didn’t find the recurrence of venom in the blood. So we believe that a big-animal model can not only help us to study the efficacy, which means the elimination of venom antigens in the blood, but also will help us to actually determine the dose that should be used to completely neutralize the amount of a bite. You can see here, we use 5 mg. It has been proposed that 5 mg can kill a man. We know that 5 mg can kill this animal after 24 hours with a single -- very similar to what occurs in humans.

These are my final remarks. Plasma venom levels

have not been used routinely as endpoints in human studies of envenomation. However, several studies suggest that venom levels may be valuable for monitoring the suppression of venom antigenemia and for developing correlations with ongoing venom toxicity.

In addition, the presence of detectable venom at baseline, albeit retrospectively, serves as a confirmation of the validity of clinical diagnosis of the envenomation, and the disappearance following antivenom treatment supports the expectation that antivenom works by binding venom and removing it from bioavailability. Accordingly, we favor the notion, as pointed to by Professor David Pigston move than a decade ago, that the disappearance of venom antigenemia following antivenom administration is a valid endpoint to objectively measure the efficacy of antivenom.

Thank you very much.

DR. SIEGEL: Thank you, Dr. Alagon.

Next we will hear a summary of the January 2009 NIH conference on coral snake antivenoms, given by Dr. Steven Seifert, from the University of New Mexico School of Medicine and the New Mexico Poison and Drug Information Center.

Agenda Item: Summary of January 2009 NIH Conference on Coral Snake Antivenoms

DR. SEIFERT: Thank you.

In January of 2008, the envenomation special interest group of the American Academy of Clinical Toxicology sent a letter to the FDA highlighting the impending loss of coral snake antivenom and its potential consequences. That resulted, ultimately, in a meeting, which was sponsored by the NIH Office of Rare Diseases and coordinated and organized by FDA/CBER. This meeting included about 50 individuals -- clinicians, basic scientists, venom researchers, antivenom developers, and representatives of the NIH and FDA -- to discuss potential solutions. Much of what you are hearing today was presented and discussed at that meeting.

These slides will include the essential summary points from that meeting, the things that were presented and generally agreed upon without controversy, and a post-meeting questionnaire that I sent out. The respondent number was 17 from that, involving 12 clinicians, 2 basic scientists, 2 representatives from industry, and 1 member from the FDA. That questionnaire, however, was sent out and responded to prior to the release of the issue summary that you see in your packet now. I have added some additional slides at the end of my presentation to what you have in your packet to respond specifically to that.

The summary statements will be followed by an

"SS" to indicate that those were general points of agreement and presented at the meeting. The questionnaire responses will be followed by either a "C," when there was overwhelming majority consensus, "SM," for supermajority, "M," for majority, and "NC," for no consensus.

Obviously, you have heard that the coral snake antivenom is no longer being produced, and we are running out very quickly. We anticipate having no new supply of antivenom capable of replacing used stock sometime this year or next, and then over time, over the next few years, existing supplies will be used up and/or will expire.

It was published that the estimated mortality pre-antivenom was around 10 percent. You have heard varying numbers, between 10 and 20 percent. It's important to realize that there are a number of variabilities that change what we might expect to see today. You have heard about the one death that we have had.

One thing that we can say is that in the absence of antivenom, an increase in morbidity and mortality over what we are currently experiencing certainly can be anticipated.

This is from a paper that was just published in *Clinical Toxicology* -- I should have put the reference at the bottom; I apologize -- that Leslie Boyer and I analyzed, five years worth of U.S. poison center data, from

2001 to 2005. You can see that about half the cases that are reported to poison centers, in fact, occur in Florida. That number is 34. Texas sees about a third or so of the total number. That number is 18. Only one other state has more than one a year, and that's Georgia. Then, of course, the western coral snakes are not really applicable to this. We don't see clinically significant envenomations in Arizona and New Mexico.

From our data, we find that antivenom is being used about half the time when we have coral snake envenomations reported to poison centers. The majority of those actually are in Florida, because Texas poison centers have generally moved to a model of waiting and seeing because of the decreased morbidity and mortality that they see with their snakes.

Historic control data is available in a number of settings. You have the poison center data that I just showed you. You also have a variety of published series in case reports. Those are pretty limited. There are a number of difficulties that you should keep in mind when looking at historical data as your comparative point:

The dry-bite rate is probably much higher than reported. People don't submit reports and say, "Hey, I had six bites, and they were all negative." In addition, when they do report envenomations, they tend to be the more

serious cases. These are the things that are interesting or otherwise worthy of publication. So there is a publication bias. When we are looking backwards and trying to estimate what the morbidity and mortality are, the numbers that we are going to see are probably much higher than we are going to see in actual clinical practice.

This is important, because if you use that to create your N for what you need to see for efficacy of a new antivenom, you are going to underestimate the number of cases that you need to treat in order to see efficacy.

In addition, this difference between Florida and Texas snakes is important because if you include Texas snakes in your clinical trial and then you are comparing it to historical controls that used mostly Florida snakes, again you are overestimating the efficacy and you are going to come up with a product that looks efficacious, but may or may not be.

Everyone agreed that there were currently no tools that can definitively identify in a clinically relevant timeframe the coral snake bite. This really doesn't give people a lot of problems. We treat things that sound like they are coral snake bites, and we don't worry too much about it. But in a clinical trial, the implication is that you may be including some patients that are not true coral snake bites. Perhaps as much as half of

your population might be dry bites and non-coral snakes, and in a small clinical trial, that certainly can skew your results significantly.

Certainly for managing the cases, it's adequate.

Venom effects can occur rapidly. The average time from a bite to the onset of neurologic effects, if they are going to occur, is somewhere around three hours. They can progress very rapidly. Within an hour or two, you may have complete respiratory arrest and the patient on a ventilator. This is important because if you give antivenom at any point along the way, you can stop progression of those neurologic findings, but you don't rapidly reverse them. So if someone is already on a ventilator by the time antivenom gets to the hospital and gets into the patient, they may very well remain on a ventilator for days or weeks, while you are waiting for the effects of the venom to wear off. This has implications both in terms of treatment and in terms of a clinical trial where you are trying to assess the efficacy and you have to give the antivenom at a point where you can still see whether there is cessation of progression.

As you heard from Dr. Borron, therefore, a distributed system of antivenom is necessary. You have to have antivenom either at the hospital when the patient arrives, so that you can give it in advance of or at the

very, very signs of neurotoxicity, or some way to get the antivenom to the hospitals very quickly. Although we do have rapid-transport means with helicopters and the like, there is virtually no precedent for any sort of a clinical trial involving poison centers like the scenario that he mentioned.

There are over 1,000 hospitals in the endemic area between Florida, Alabama, Georgia, Louisiana, and Texas, where these bites occur. That creates a variety of logistical issues, if you are either trying to distribute antivenom under a clinical treatment protocol or trying to do anything else that involves coordination of multiple sites.

There are at least two foreign antivenoms where we have animal data that suggests that they should be effective against U.S. coral snakes. Keep in mind that these are relatively crude animal assays. They are just basic survival studies, and you can't really get to the fine points of clinical management of cases until you actually start treating human beings with antivenom. In fact, there are no cases in which these antivenoms have been used in patients that are bitten by U.S. coral snakes.

We have an exotic antivenom delivery system that we have looked at as a possible way of dealing with coral snakes in the interim, because there is going to be some

gap here between having the existing antivenom and getting a new product. This is a bit of a problem. Zoos maintain antivenom stocks for exotics in their collection. If a zoo has a South American coral snake or a Mexican coral snake and they have one of these foreign products in their inventory, you can call up and get that shipped to a patient. Of course, the zoo could be in Seattle and the patient might be in Texas, and so there are going to be issues of time delay.

In addition, there are currently only 30 vials of unexpired foreign coral snake antivenom in zoos in the U.S.

There are some alternative models. You could have something like the Miami-Dade Antivenom Bank accumulate a large number of vials and rush this to patients within a region. This could effectively, perhaps, treat the majority of your snake-bite victims. But it still has a variety of issues, especially if you are trying to do this in the context of a clinical trial. There are issues of medical control, the long-term stability of that system -- their funding right now is under threat -- and there are other logistical issues.

Everyone agreed that, regardless of the ultimate mechanism chosen for approval of a new antivenom, substantial financial incentives are going to be required to encourage or even allow a firm to want to come in and

test their antivenom and try to get approval. It's not a problem if you can charge $1 million for a vial. Ultimately, they would get their money back. But it's not going to happen.

On the critical question of what the appropriate pathway to licensure is, there was no support among any of the respondents for a conventional licensure pathway. The accelerated approval pathway was considered possible by some questionnaire respondents -- about a quarter. But the timeframe of a clinical trial and some of the other logistical elements were not specifically addressed in the questionnaire. This, as I said, came out before the issue summary by this committee. But there were a couple of issues that were addressed.

If an accelerated pathway was approved, a comparative trial against the existing antivenom was considered impractical because of dwindling supplies. Also there should be some strategy for distribution of the experimental antivenom throughout the endemic areas, perhaps through treatment INDs, so that we did not have an absence of antivenom available to treat patients showing up in this interim period between the exhaustion of the existing supply and the approval of a new product.

However, low numbers, the wide geographic distribution in toxicity that you have heard of, the

absence of validated endpoints all serve to make a clinical trial impractical. That was a statement that was approved by a supermajority, more than 75 percent of respondents to the questionnaire.

By the same number, the supermajority felt that an animal efficacy pathway that combined human PK and human safety studies was the only feasible and ethical alternative for initial drug approval.

I took a look at the issue summary that was sent around. I want to just make a few points on what I consider to be the key issues.

Multiple sources of overestimation of your historical controls result in a significant undercount of what you are going to need in order to show clinical efficacy. This suits the purposes of the clinician who would like to see something validated, but I think, in the absence of statistical rigor, you are going to show that your antivenom is effective with a very small trial, whether it is or it isn't.

In addition, the geographic variability is a major factor. If you look at the two studies that were used to calculate that 15 percent rate, Parrish's had six Florida snakes and five Texas snakes. All five of the Texas snakes were dry bites. So 100 percent of the morbidity and mortality from that study is entirely based

on Florida snakes. Fifteen out of 17 from the Neill report were Florida snakes.

In addition, basic medical care has improved significantly since 1967. Every hospital emergency department is staffed by well-trained emergency physicians, who can intubate. You have improvements in critical care. You are going to see a much lower mortality rate than you did in 1967.

All of these are going to falsely -- actually, they are going to increase the number you would have to include in a clinical trial to see a true clinical benefit.

The surrogate markers -- I just want to make one little comment. I'm sure that ultimately we will find validated clinical markers, and it may very well be the disappearance of venom antigens. But I want to point out a recent example where that is not the case. We have CroFab for the treatment of viperid envenomations, and local recurrence -- the tendency for the local injury to stop progressing and then start back up again -- occurs in the absence of detectable serum antigens. So you would be doing the surrogate marker, you would see no recurrence of those venom antigens, but you would have a clinical recurrence of effect. Of course, it can work the other way around as well. You might not see disappearance of the venom antigens, but you may see a perfectly valid clinical

effect.

So there are pitfalls, and we are not yet at a point where we have a truly validated marker.

The challenges to a clinical trial I think are extreme. The bites are spread out over the largest area you can imagine. Most hospitals in this area see fewer than one bite a year, and few, more than two. To use a more reasonable number, you probably need 75 to 100 patients to do a study. If you have a single death, for any reason whatsoever, in your clinical trial, your number jumps to close to 200, to see what I think is a truly significant clinical efficacy.

But even if you accept that 20 patients will show efficacy -- and this is where I need to modify my slide just a shade. We have two recent examples -- more than two, but I'm going to just use these two to show you what the real-world example is of trying to do a clinical trial with envenomations.

We had the CroFab clinical trials which were done in the late 1990s and early 2000s. We had 100 times the available patient population. There are roughly 8,000 snake bites from viperids a year in the U.S. There are numerous hospitals around the country that see multiple bites -- 10, 15, 20, 30 bites a year or more. We were able to do this study at 20 different hospitals. It took three

years to enroll 42 patients. If you apply the same set of limitations to a 20-patient coral snake study, that works out to about 115 years.

There was a scorpion study just published in the *New England Journal of Medicine*. I took the data right from that published study. This is about 10 times the number of patient population -- there are 8,000 scorpion stings, but about 600 to 800 neurotoxic scorpion envenomations, ones that would be appropriate for antivenom, so about 10 times the study population. They indicated that it took them a year and a half to accumulate 15 patients. If you apply that to a similar 20-patient coral snake antivenom study, you are looking at about 20 years.

Okay, you are going to make some adjustments. You are going to try some novel things to try to reduce that. But I think this is your starting point. This shows the difficulties that you are likely to face in creating a coral snake clinical trial.

My answers to the questions that were posed to the committee would be: No. No. Possibly, but many pitfalls involved even in animal extrapolation. Possibly, but the answer should be postmarketing trial only. An animal efficacy model, with human PK and human safety studies, is the only feasible and ethical alternative to

initial licensure.

Thank you.

DR. SIEGEL: Thank you, Dr. Seifert.

Finally, let's hear from the FDA, from Jessica Kim, Ph.D., on statistical considerations.

Agenda Item: FDA Statistical Considerations

DR. KIM: Good morning. My name is Jessica Kim. I'm a statistician at FDA.

I'm going to present the statistical considerations on study design to demonstrate efficacy for coral snake antivenom products. My presentation focuses on providing the statistical background related to the FDA question number 1 to the committee members.

This is the outline of my presentation. As a background, I'm going to briefly point out typical clinical-trial designs under FDA's regulation and then clinical-trial designs for antivenom products, by showing three different studies: first, under the name of the historical-control study; the second one is an active-control study; and the third one is called a threshold study. Then I'm going to summarize my presentation.

Here's my background. Clinical trials under FDA regulations require well-controlled requires showing safety and efficacy prior to marketing approval. Typically, double-blinded, randomized, and placebo- or active-control

studies are performed. However, under special circumstances, such as where a placebo control is unethical or the active control is not available or a clinical trial for a rare disease where patient recruitment can be a challenge -- under those special circumstances, flexible clinical-trial designs are considered. The coral snake bites are one of the unusual conditions that need more creative trial design.

As far as flexible clinical-trial designs are concerned, under statistical umbrella, to show the efficacy for coral snake antivenom products, the following three studies will be discussed:

- The first one is a historical-control study. Here, the historical control means mortality rate without the treatment. It is to demonstrate the superiority of the new treatment over the historical control.
- The second one is an active-control study. The active control in this case -- FDA approved a product, which is the Wyeth product. It is to demonstrate non-inferiority of the new treatment versus active control.
- The third one is a threshold study. It is to satisfy a threshold recommended by the FDA. One of the well-known examples you can find is the beltline approach for the IGIV products.

For the first study design, which is the

historical control, we need to be aware of the general caveats, the general limitations, of using a historical control, which is estimated by the peer-reviewed scientific literature:

- The non-comparability of studies from published literature may impact on estimating the treatment effect.
- Internal validity and quality of the published studies may to be reliable in estimating the treatment effect.
- Innovations in medical treatments or changes in medical practice over time may impact on estimating the treatment effect.
- Comparability of the specific target patient population can be an issue.

Many of these items can be an issue in estimating the historical control of the mortality of the coral snake bites.

The historical-control study: The study objective is to demonstrate the superiority of the new treatment over the historical control, which is based on two published case series using untreated patients. Table 1 shows the weighted means to estimate the overall historical control, applying three different weighting methods. Using a study-size weighting method, historical control point estimate, we obtain 18 percent. The equal-

weight weighting method produces 16 percent. The reciprocal-of-variance weighting method produces 15 percent.

To take the variability consideration, the third column, which is the 95-confidence-interval estimate, is also provided. As you can see, it may vary from 2 percent to 32 percent, due to the variability of the small studies.

All of these numbers can be used for historical control, but the judgments depends on clinical benefit, as well as practical reasons. The numbers in yellow, which are 18 percent and 15 percent and 4 percent, are considered in calculation of the sample size to conduct the superiority trial.

The study hypothesis to show the superiority -- H_0 is the new antivenom treatment mortality rate, is greater than or equal to the selected historical control, versus it is less than the selected historical control.

Table 2 shows the estimated sample size with 80 percent power and 0.025 significance level. Depending on the assumption, the first column, the information about the new product predicted mortality rate, the sample size required for each selected historical control -- the second column is 4 percent and the third column is 15 percent and the fourth column is 18 percent. Those are the minimum required sample size to satisfy the assumptions with 80

power and the significance level to test this superiority.

As you can see, the last row in the second column shows, if you assume the new product predicted mortality rate is 3 percent, to show the superiority of the selected historical control, 4 percent, the study size becomes 2,756.

The second one is the active-control study. It is to compare a new product to the Wyeth mortality rate, which is known as very low, known as less than or equal to 1 percent. It needs to be compared to the Wyeth mortality rate, with a pre-specified non-inferiority margin.

The active-control study to show the non-inferiority -- the study hypothesis becomes, the difference between the new antivenom treatment mortality rate and the active-control mortality rate is greater than or equal to delta versus less than delta -- delta in this case, pre-specified non-inferiority margin, which means the new antivenom treatment mortality rate is no worse than the active-control mortality rate by no more than delta.

Table 3 shows the estimated study size, again with 80 percent power and 0.025 significance level. For example, to understand this Table 3, the first row -- if we pre-specify the non-inferiority margin delta as 0.5 percent -- the difference between that is no more than the 0.5 percent -- then the sample size required per arm would

be 6,383. If you relax the non-inferiority margin to larger than 0.5 percent, the study size could be relaxed as smaller than the 6,383.

The third study that we considered is a threshold study. It is to satisfy a threshold recommended by the FDA, such as, a one-sided and 97.5 percent upper confidence limit for the mortality rate is less than 4 percent. Here the choice of the one-sided and 97.5 percent confidence level is typically used in the clinical trials under FDA regulation. The 4 percent is arbitrarily chosen, but again it is based on the 95 percent lower confidence level of the historical control.

Table 4 shows the estimated study size to satisfy the FDA recommended threshold. To satisfy the recommendation, the minimum sample size required, without showing no deaths, will be 91. Again, if we relax the number of failures -- meaning that we allow more deaths -- in the new treatment, the sample size becomes large. To meet all the preset FDA's recommended threshold by allowing more mortality, the study size becomes large.

To summarize my presentation, I would like to point out the statistical issues related to each one of the study designs.

The first one, historical control: The validity of the historical control can be disputable. Selecting

which estimate for the historical control, such as point estimate versus confidence limit, can also be disputable.

For the second, the active-control study, the FDA-approved active control is almost exhausted, so the active control won't be available. The selection of a non-inferiority margin can also be disputable. Depending on which non-inferiority margin is selected, the study size becomes large.

For the third study, the threshold study, the scientific basis of determining acceptance criterion can be an issue.

So these are the statistical approaches to be considered as a feasible study for the efficacy of the coral snake antivenom products. These are the issues that I would like to point out.

Thank you very much.

DR. SIEGEL: Thank you, Dr. Kim.

We are a little bit early, but I would propose that we take a break now and then come back for discussion. As of now, there are no scheduled open public hearing speakers, so we'll have plenty of time for discussion.

Let's be back in 15 minutes. Thank you.

(Brief recess)

Agenda Item: Open Committee Discussion

DR. SIEGEL: I'm sure there are lots of questions

for the speakers from the committee. So let's get started. Who would like to begin? Perhaps Dr. Fleming?

DR. FLEMING: I don't know if I planned to begin, but let me at least throw out an initial question.

Obviously, this is a very challenging circumstance, because we have an extremely limited number of people -- fortunately, an extremely small number of people -- who are impacted. Yet it's extremely important to ensure that what we are offering to them provides the level of benefit that we are indicating. I'm persuaded that looking at an historical-control type of approach is going to be strongly motivated when you consider feasibility.

The issues, though, get to, in essence, understanding, as Dr. Seifert was saying, what is an unbiased estimate of that control? He pointed out some really key issues about publication bias and about the true rates of mortality differing very much in Texas and Florida, as previous speakers did. He was challenging more, though, the ability of using a surrogate that would be based on, for example, consistent decrease in venom levels if we try to use an accelerated approval strategy and advocated animal studies.

I guess my question for him is -- I appreciate very much his concern, the points he well articulated, that

we are at risk of declaring efficacy when you have, in fact, an ineffective or relatively ineffective intervention because of all these reasons. But why are animal studies not free of that risk? What specific validation do we have for animal studies for their enhanced reliability in this setting?

DR. SEIFERT: The reliability of animal studies in this setting is equally open to question. I think that the risk of not having antivenom available is at least as great -- we know these antivenoms work against analogous coral snake envenomations in their native environment, and they appear to have the same sorts of animal efficacy that we see against the snakes to which they were raised. So it's as good a surrogate as we have.

DR. SIEGEL: Dr. Zimrin.

DR. ZIMRIN: I have another question for Dr. Seifert. What is thought to be the mechanism behind the recurrence of local symptoms after the venom level decreases with the pit viper antivenom? Are there any thoughts about why that happens?

DR. SEIFERT: There probably needs to be a certain level of circulating antivenom to meet the locally destructive venom components at the margin of envenomated tissue and normal tissue. You have to have a certain amount of antivenom constantly arriving at that junction.

Because CroFab is a Fab antivenom, it distributes to a larger volume than the older IgG product, and so antivenom levels drop fairly rapidly from that distribution. That's probably the mechanism by which we lose protective levels of circulating antivenom for that purpose.

It was an example of how you couldn't necessarily predict that. We didn't anticipate it when we began the trial. It was just an observation. All of a sudden we were seeing recurrences where we had never seen that before.

DR. ZIMRIN: Thank you.

DR. ALAGON: Could I comment on that? What Dr. Seifert showed was the recurrence in local symptomatology in the case of rattlesnakes. But as I mentioned in my presentation, most of the venom remains at the local site. But there is a big difference rattlesnake venoms and coral snake venoms. Rattlesnake venoms produce lots of coagulation problems and sometimes they also produce lots of local tissue damage, whereas in the case of coral snake, you don't have that.

The way I can interpret his findings, even if you have the venom antigens neutralized in blood, you still have active venom on the injury site, which is producing that local damage. But in the case of coral snakes, you may think that you have active neurotoxin in the active

site, but that's not the target place. So whenever the toxins leave to the blood, to distribute to the body, then they will meet the antibodies and will be neutralized.

I think his observations are very interesting, but I don't think they do apply directly to what we are now here about.

DR. SIEGEL: Dr. Kulkarni.

DR. KULKARNI: I just have a question. Is there any data, since most of these venoms cause paralysis, about the levels of these venoms in the spinal fluid? Do they correlate with clinical manifestations? You are measuring it in the blood, but I'm just wondering whether --

DR. ALAGON: We don't know in patients -- that's why I used several examples, experimental as well as clinical examples, in my talk, to show that there is a correlation between venom levels and symptomatology. But for the particular case of coral snake, we still don't know. We know in our large-animal model that we start seeing the development of neurological symptoms when you have big amounts of venom in blood and that you prevent the development of that neurological symptomatology if you inject antivenom one hour after. Of course, now we are playing with this window. We will apply antivenom after two hours, three hours, and four hours, and see how much time we have, at least in that animal model.

DR. SIEGEL: Dr. Lewis-Younger.

DR. LEWIS-YOUNGER: I was just going to comment on that from the perspective of the clinician treating the coral snake antivenin. We have no clinical assessment/laboratory techniques. There is nothing that we have available, except for -- basically, we make the decision to treat or not to treat based on history.

DR. PEREZ: On the surrogate animals, if that was used, would you use the traditional ED_{50} or would you use maybe some rescue studies to do the comparison?

DR. ALAGON: In the case of the large-animal model, we are using rescue studies. We envenom the animal at zero time and then we provide the antivenom 60 minutes after. We still haven't tried longer periods of time after the venom injection. You may consider these experiments as being rescue experiments.

DR. HOLLINGER: Along those same lines, in the large-animal model -- again, in the dog -- when the venom comes back -- you may have said this and I just missed it -- when the venom comes back by ELISA, do the animals die? What is the outcome to the animals in terms of looking at a surrogate model?

DR. ALAGON: All I can tell you is, when you keep the venom levels low -- that means when you give a sufficient dose of antivenom -- the animals don't even have

neurological symptoms, so they survive very well. We haven't tested many different doses. I don't know exactly what will be the outcome if we use suboptimal amounts of antivenom. But if we use optimal amounts of antivenom, they don't develop any neurological symptoms and they survive.

DR. HOLLINGER: You showed one animal that had a venom level that came down when you gave a small dose of antivenom and then, subsequently, the venom level came up to the same level it was initially. My question is, what happened to that animal?

DR. ALAGON: The problem with that animal is, that animal was under anesthesia, and then we sacrificed the animal at the very end to measure the venom in the tissues. So I don't know that answer. We still need to carry out that experiment.

But in the case of the 10 vials, we allowed the animal to survive -- well, we woke up the animal from the anesthesia, and it was perfectly normal.

DR. PEREZ: In large animals, do any of those have any natural immunity, such as is found in opossums and wood rats?

DR. ALAGON: No. They die with 5 mg. When you use an animal of between 50 and 60 kg and you give 5 mg of *Micrurus fulvius* venom, they develop neurological

symptomatology. And this is very interesting. The symptoms start after 12 hours. So it takes, really, a very long time. But they die between 24 and 36 hours after you inject the venom. You also see muscle paralysis, and they die of respiratory distress.

DR. SIEGEL: Dr. Bianco.

DR. BIANCO: Another question: Are the immune complexes totally neutralized? If you compare immune complexes *in vitro* and then you inject, sometimes do you still have some toxicity?

DR. ALAGON: No. That's exactly what you do when you do your ED_{50}s in mice. You pre-incubate the venom and the antivenom, and when you have titrated all the venom antigens and you inject a mouse with that, they will not develop any symptomatology. They will just do fine.

DR. TRUNKEY: I have several questions. It was mentioned this morning that the venom is similar to curare. I would wonder, from the emergency physicians -- you can reverse curare -- has this ever been tried, once you get the paralytic symptoms?

DR. BORRON: There are some papers from the South American literature where edrophonium and, I believe, neostigmine have been used, with some effect, in patients who presented late, and improvement in paralytic symptoms. But this is a handful of case reports, one or two cases.

Steve, are you aware of any others?

DR. SEIFERT: It has to do with binding affinities when you are talking about postsynaptic receptors. Also there are some presynaptic receptors that traditionally are considered irreversible. There is some evidence that patients treated with antivenom get better faster than they would if they had not -- but there aren't good controlled clinical trials.

DR. TRUNKEY: My second question: On page 6 of the issue summary that was sent to us, you are comparing three products made outside the United States. But the flaw in this study, it seems to me, is that you are mixing the venom and the antivenom and then injecting it into mice, and so you don't have any tissue levels. I guess my question here is, if proper animal studies were carried out, can we license one of those three products?

DR. SEIFERT: You are asking whom?

DR. TRUNKEY: I don't know the answer to that.

DR. MICHAUD: I think what you are asking is one of the pivotal issues that we are discussing now, more or less, which is whether or not the animal efficacy rule can be invoked to license this kind of product, or whether, if it's considered feasible, we can license with clinical trials.

DR. TRUNKEY: I wouldn't accept these three

compounds based on the way it was modeled on page 6.

DR. MICHAUD: That's absolutely right. These studies would not meet the needs of the animal efficacy rule, first of all, because they don't really simulate what happens in people very well. What they are used for usually are studies of potency to release a product. The reason they were done is as sort of a beginning standpoint to figure out whether you are or aren't likely to have cross-reactivity.

But you're right. That's all there is that we are aware of. That's why we gave it to you.

DR. TRUNKEY: My third question is, why did Wyeth quit making it?

DR. MICHAUD: I believe it was a business decision. We do have people from Wyeth here who could speak to this.

We are not aware of any safety or efficacy specific concerns, if that was your question.

DR. TRUNKEY: I understand, but it's still a question.

The fourth thing -- this is just an aside -- if an adult drunk male picks up knowingly a coral snake, is it desirable to keep that person in the gene pool? (Laughter)

DR. CRYER: I have a question, too. If I heard it right, basically it wouldn't make any difference whether

you had antivenom or not; as long as you make it to a hospital, you are going to survive if you have a coral snake bite. You just support the patient until the stuff wears off, after a week or two, and they get better. So the only thing they die of is the complications of being on a ventilator or being in the hospital.

It seems to me, if that's true, that there is no reason to have mortality as an outcome in any study that we do here, because they are not going to die. None of them should, unless the drug itself is toxic.

Any comment on that?

DR. GOLDING: Can I just say -- I think it's in the issue summary -- that we realize that? This is obviously a valid point. What we would think of as an endpoint -- instead of using mortality, you would use an endpoint as the need to be intubated and placed on a ventilator as the endpoint rather than mortality, to take care of that problem.

DR. CRYER: In that case, then, would a placebo study be a better study design?

DR. GOLDING: We don't think a placebo study would be regarded as ethical, because we think the preponderance of the data suggests that the antivenom does prevent serious outcomes.

DR. CRYER: I thought I heard that in Texas, 70

to 80 percent of the patients aren't treated with antivenom anyway.

DR. GOLDING: I think the focus of the study would be on envenomations that were such that you would expect a clinical outcome. Listening, as you were, to the presentations -- and we knew this beforehand -- probably the focus of the study should be in Florida, with the *Micrurus fulvius*, and not with the other snakes, which are much less poisonous.

DR. SIEGEL: Dr. Lewis-Younger.

DR. LEWIS-YOUNGER: I get a little bit nervous when you start talking about placebo treatment of coral snake bites in Florida, and concerned. They are not all drunk men out there getting bitten by the coral snake. We have children. We have people just doing their normal, everyday activity, like painting a house or whatever. Not only do we have an issue that is pertinent for outcomes -- being on a ventilator for a prolonged period of time is not a benign outcome; potentially life-threatening -- but there is a huge economic burden as well. If you can treat someone with an antivenom and discharge them in less than 24 hours versus having somebody in an intensive-care unit, with all those risks and costs, it's a lot bigger problem.

DR. SIEGEL: Dr. Zimrin.

DR. ZIMRIN: If we erased drunk men doing stupid

things from the gene pool, I think that would have tremendous implications. I would not advocate that.

I do have a question about the -- we heard a lot in both Dr. Kitchens' presentation and his paper about the side effects associated with the Wyeth product. But I didn't hear much about the side effects associated with the alternatives. Do they have the same rate of anaphylaxis or serum sickness? Does anyone have any idea about that?

DR. BORRON: I don't have any human data or experience on these particular antivenoms. But in general, based on other Fab and F(ab)2 antivenoms -- for example, digitalis Fab and some of the others -- the incidence of allergic reactions is much, much smaller because the Fc fragment, which is cleaved off, is the source of most of the allergic reactions. So, in general, anaphylaxis is an extraordinarily rare event with Fab and F(ab)2 antibodies.

DR. KITCHENS: The closest death I had was the anaphylactic reaction. Essentially, if no one has died since 1967, whoever got treated with this -- I wonder how many people died because they got treated with the Wyeth product. I don't have those data. But the severe anaphylaxis, about 1 in 100, some number like that -- hopefully not fatal. But if we do a study, we have to make sure that, if we get some efficacy, we don't erode the efficacy by any reactions.

DR. BOYER: I'm Leslie Boyer. I'm the medical director of the Arizona Poison Center and I'm the principal investigator on clinical studies of the scorpion antivenom that you heard described earlier. It's an F(ab)2 product made by the same manufacturer in Mexico that makes one of the products that is listed in your handout.

We have to date enrolled over 900 children in a clinical study -- a series of them -- in Arizona. The safety rate with this antivenom has been phenomenally better than what we saw in the past with whole IgG. The rate of true serum sickness is hard to calculate because we haven't seen a single case that met the full-blown serum sickness definition. We have seen a handful of rashes that were indistinguishable from what you would expect in a normal population of rashes. The acute reactions that we have seen with the F(ab)2 have included a less-than-5-percent rate of any kind of acute reaction and 1 out of 900 with anaphylaxis. In that case, there was a second agent that may have caused the reaction.

So it's a very low rate.

DR. SIEGEL: Thank you, Dr. Boyer.

Is there, in fact, someone from Wyeth who could answer Dr. Trunkey's question?

PARTICIPANT: I am from Wyeth. Are you referring to the question about why Wyeth discontinued manufacture?

DR. SIEGEL: Yes.

PARTICIPANT: That was a decision based on a series of business decisions about closing the plant where we made the antivenin. It was not a decision about antivenin itself.

DR. TRUNKEY: Can you not transfer that license to another pharmaceutical company?

PARTICIPANT: Yes. Another pharmaceutical company has to be interested in manufacturing the product. One of the big concerns everybody would have, I'm sure, is what it will take to get that licensed. That's why this meeting is very important.

DR. SIEGEL: Questions from the committee?

DR. MCCOMAS: We have heard some things about the antivenom available to zoos, and that there is a limited amount left, but it has been transported out, presumably to coral snake bites. I'm wondering, have there been follow-ups on how much of this has been used and the efficacy of those products? Is this the Wyeth product or another product? Or are these data not comparable because they are not the Florida coral snake and they may be more exotic snakes?

DR. SEIFERT: As far as I'm aware, there have been no instances where the non-U.S. coral snake antivenom has been used to treat a U.S. coral snake bite in this

country. Roughly 50 times every year, we have an exotic bite by a cobra or a Taipan or a boomslang or some other exotic species. That's the system that has been set up to treat those -- to acquire zoo supplies of those exotic antivenoms and get then to the victims, wherever they happen to be. Since these are mostly private collectors, they tend to be scattered all over the landscape.

My point was simply that that system could be used to supply coral snake antivenom that zoos have for their exotic coral snakes, if we didn't have any other source, such as a clinical trial or a product that's approved on the basis of the animal rule.

Can I make a comment about the surrogate marker question? The key issue here is, if we had a validated surrogate marker, we might require very many fewer patients in a clinical trial to show efficacy. If I'm a betting person, I'm betting that Alejandro is right and that disappearance of detectable venom will be associated with clinical improvement. But we don't have a validated marker at the moment, and so I don't think it can substitute as a pathway to licensure. It's something that can be and should be developed, once we have a product that we can test in postmarketing surveillance.

Finally, let me just mention -- I forgot; I was going to start out with a foolproof way to distinguish

venomous from non-venomous snakes. You have heard lots of different rhymes, and you are probably already confused. Red on yellow, leave it alone; red on black, leave it alone. If it slithers on the ground, leave it alone. (Laughter)

DR. SIEGEL: Dr. Epstein.

DR. EPSTEIN: I just want to clarify, if you have a valid surrogate marker, you can do a conventional trial without a clinical endpoint, because you already know that the surrogate correlates with the clinical outcome. If you have a likely valid surrogate marker which is not proven to correlate with clinical benefit but is likely to do so, that's the standard for doing an accelerated approval, which is a conditional approval, subject to the requirement for a Phase 4 study to demonstrate that the product actually has clinical benefit.

So the absence of certain knowledge of a valid surrogate is not a barrier to an approval process under accelerated approval.

Is that clear to everyone?

DR. SIEGEL: Dr. Hollinger?

DR. HOLLINGER: Along those same lines, there is enough product to do comparison studies with the current product, the Wyeth product if there were a good surrogate test, like a large-animal model or a small-animal model,

you could do those surrogate tests … is that correct?

DR. EPSTEIN: Yes, a study in animals could be done directly comparative with the Wyeth product. FDA is not convinced that a prospective, randomized study against the Wyeth product as an active control is feasible. There is only about one year of product availability, and you have heard about the difficulties of the trial. But a study in animals could certainly be done comparative to the Wyeth product, yes.

DR. HOLLINGER: And then postmarketing studies after that to follow up with effectiveness and so on.

DR. EPSTEIN: Yes. If FDA were to accept anything short of a conventional study demonstrating clinical benefit, we would want a Phase 4 study to validate that.

DR. HOLLINGER: I wanted to ask Dr. Borron -- he mentioned a study that would be a helicopter and have you looked at the cost for doing the study? I don’t think we can stay on this committee for 115 years.

DR. BORRON: I think the cost would be enormous. I think the cost of any clinical trial would be absolutely mind-boggling. The geographic difficulties, the stocking, training, maintaining people up to date on the protocol who haven’t seen a snake bite in a year -- they are not going to remember what to do. They are not going to remember who

to call.

I propose that, because I think it's almost the only way that it could potentially be done, to be able to get things quickly and not have it stocked in a million places. But I really don't think it's logical. I think a well-controlled animal study in a couple of different species makes a lot more sense and is a lot more feasible.

DR. HOLLINGER: When someone gets a coral snake bite somewhere out in the community, who pays for the transport of the antivenom to that site? Does that come out of insurance or does that come out of …

DR. BORRON: Generally, up to now, as I understand it, most of the hospitals, even small hospitals, have kept a few vials of antivenom available. If they don't have it, they usually know another community hospital a few miles away, 20 or 30 miles away, that does have it, or they call the poison center and find out who might have it. The police transport gratuitously. So there has not been a charge for that.

DR. TRUNKEY: It's a good model as far as having a few centers -- poison centers, what have you. That way you can concentrate your vials and you could rapidly transport them. Helicopter transport averages about $10,000 right now. But keeping one or two vials at a hospital makes no sense to me at all.

DR. BORRON: I would agree with that. Having three or four vials in 1,000 hospitals -- I don't even know if the companies could produce that much antivenom in a reasonable period of time.

DR. PEREZ: (Off mic - inaudible)

DR. CRYER: I would agree. I brought it up because I noticed the study on the scorpion was a placebo-controlled trial. So obviously somebody has actually done it before.

But the other question I have regarding the proposed trial with the helicopter delivery is, I still don't see how you can legitimately get consent for a trial. You essentially have a drug that works 100 percent of the time, and now you are going to try to compare … I just don't see how it's feasible to do any study until what we have now, which apparently works so well, runs out.

DR. BORRON: That's the point. It is going to run out. By the time you got this study through 10 dozen IRBs and the FDA, I promise you, all of it will be gone.

DR. LEWIS-YOUNGER: Dr. Alagon, I have a question for you, sir. One of the concerns that we raised here is the issue of the confusion that is caused by the fact that we have a certain percentage of dry bites. You have developed an ELISA to analyze the venom. I would like to know, what kinds of false positives and false negatives do

you have in your ELISA test?

DR. ALAGON: We have optimized the ELISA so as to minimize, first of all, the background, the noise. We have tested serum samples from many human volunteers. So far there has not been a false positive coming from plasma. Of course, there is always a risk of having someone who will test positive. But when you are doing, actually, a study, you are taking at least two samples, a sample before, for measuring the venom levels before antivenom, and then after antivenom. If there is a patient that will give a false positive, then you will have a strong signal even after antivenom.

So there are ways to tell when you are having a particular problem with that particular sample. You can dilute the sample more and see how the background or the signal behaves. You can tell if you find a case like that.

DR. LEWIS-YOUNGER: How long after envenomation would you have to wait to get a venom sample?

DR. ALAGON: Our ELISA takes two hours to perform. For example, for the scorpion clinical trial, we measure the venom levels several months after samples were collected. Of course, we have tested the stability of venom under the conditions that the plasma samples were kept, and we know that the venom samples are stable for at least a year and a half at -20 degrees and things like

that. So we have tested all that for the scorpion, as well as for the rattlesnake clinical trials.

We are still testing some of these variables for our coral snake ELISA. We still haven't finished validating it for a laboratory viewpoint. But we are very close to finishing it.

DR. BALLOW: Some of this discussion assumes that there is going to be a U.S. company that may step in for Wyeth, but we haven't heard any discussion of that. Let me go back to our colleagues in Central America. Mexico and Costa Rica have antivenom … how close are those species to a coral snake in Florida? Have there been any studies outside of the United States?

DR. ALAGON: I will try to respond to your first question.

In terms of DNA analysis, there is a wonderful study being carried out by Eric Smith in Texas. In fact, he sent me his phylogenetic tree last night. If you are interested, we can look at it. It's very interesting. Apparently, the *Micrurus tener* from Texas may include at least two different species, according to Eric Smith.

What we know is -- first I need to say something else. The first lots of Coralmyn were meant only for Mexico. We use a certain venom species to develop that antivenom. We know that even if we are not using *Micrurus*

tener or *Micrurus fulvius*, that antivenom neutralizes very well. There have been at least two or three different laboratories or investigators throughout the world who have really tested it. In fact, Coralmyn is as potent in terms of neutralization capacity as the Wyeth product, using the mouse model. There is a limitation there. But we know already that *Micrurus fulvius*, using the large-animal model -- the Coralmyn neutralizes very well the *Micrurus fulvius* venom.

This is a very common situation with antivenoms. For example, for the scorpion antivenom, the Mexican scorpion antivenom doesn't use the Arizona scorpion venom for immunizing the horses. But the potency of the Mexican antivenom against the Arizona scorpion venom is very high. I can tell you that it's even higher as compared to -- it works better against the Arizona species, to one of the species that is used for immunizing the horses.

So I think that once you can prove that you have cross-neutralization in animals, you can feel quite confident that it will also work in humans.

DR. KULKARNI: I was wondering whether there was a global antivenom against neurotoxins from different species so you can increase the number of individuals who can receive this antivenom?…

DR. ALAGON: In fact, in Mexico we have something

like 20 different species of coral snakes. We know that Coralmyn is very good for some of those species and is good for some others. Bioclon is doing that based on basic research, doing these cross-reactivity studies. They are now changing the immunization mix for the second generation of Coralmyn. They are really working very hard on developing the second generation of Coralmyn, which includes *Micrurus tener*, as well as *Micrurus fulvius*, as the immunogens.

I'm very happy about that, as a Mexican, because *tener* is also in Mexico. We know that that will expand the potency for these other Mexican species that are really well neutralized. So the second Coralmyn generation will have a wider capacity to neutralize venoms of many different Mexican and North American species.

DR. PEREZ: I think a universal anti-venom worldwide would be very difficult to produce. The other question I have is, what about the other molecules in anti-venom such as phospholipase?

DR. ALAGON: That's a very good question. There is only one paper -- actually, there are two papers, one that I published in 1979 characterizing a non-neurotoxin phospholipase from *Micrurus tener* and a recent paper in which they did a proteomic study of *Micrurus fulvius* venom. But they only characterized the alpha neurotoxins, which

are small proteins, as you know very well, 7,000 daltons.

But we know already that there are beta neurotoxins, and those are phospholipases, 14,000 daltons, that also act on the neuromuscular junction, but presynaptically. They inhibit the release of acetylcholine. I think they have been underestimated in terms of the mechanism of action in human patients. I believe that, actually, we are dealing with two different types of toxins when we talk about coral snake envenomation, the well-known alpha neurotoxins, as well as the beta neurotoxins, these neurotoxic phospholipases. I agree with you 100 percent.

DR. SIEGEL: Dr. Rentas.

DR. RENTAS: Dr. Alagon, a quick question for you. I don't know if you work with this company or not, Bioclon.

DR. ALAGON: Yes.

DR. RENTAS: Are you willing to say that this company is going to put all the resources that it may take to get this through FDA license approval? If you are not, is there anyone else in the audience that could comment on that?

DR. ALAGON: I can comment on that. Pauline (phonetic) is also here, representing Bioclon.

Of course, I know that they are interested in

producing Coralmyn for the USA. But what I know for sure is that they will not really -- they don't have the resources to pay for any type of clinical trial. It will be just too expensive for them, especially considering that the market also will be very small. But they are interested in producing an improved Coralmyn -- and they are working as we speak on that -- that will have a better potency in terms of amount of protein, higher specific activity for the *tener* and the *fulvius* species.

In that regard, they are doing their part, in terms of improving the Coralmyn for the USA coral snake species. Perhaps they will be able to provide free antivenom for doing a limited study or things like that. But certainly they don't have the resources for doing a formal clinical trial of any type.

DR. ZIMRIN: This is a question for anyone in the FDA. Just to understand what we are considering, if we agree that a clinical efficacy trial is not feasible, are we considering sort of distributing this alternative … having it available and then having either poison control centers or an academic center, like the University of Florida, trying to coordinate the results, the clinical efficacy results, and the ELISAs?

DR. GOLDING: There is obviously a public health need here, and there needs to be some way of getting this

product to the market. The question that you have heard already is that the accelerated approval pathway does allow you to use surrogates, which could be animal models. That's probably the most likely way you are going to be able to get premarketing approval. If you can collect some clinical data before premarketing approval, that would be helpful. But the actual pathway requires that, because this is not a validated endpoint, you would have to do some postmarketing studies.

What we also envisage as a requirement for approval is to at least do some pharmacokinetic studies and some safety studies in a small number of individuals.

So that is our view. We haven't entered, as far as I know, into major discussions with any of the companies. We would have to think about it. But what we are asking, I think, the committee to answer is, because of the limitations that you are hearing, is that kind of approach reasonable to proceed with under these circumstances?

DR. EPSTEIN: If I could just comment a little bit further, if the committee were to recommend or advise that human studies are impractical and infeasible, then we would need a threshold for the animal efficacy rule, which means that we can license based solely on a study in animals. Alternatively, we must do a human study. If we

do a human study, the question is whether we have either a valid surrogate -- and I'm not sure I heard one -- or a likely valid surrogate, in which case we could do a small study based on a likely valid surrogate and grant an accelerated approval. That could be a small study. But I think we have heard that even a small study could be quite difficult.

So what we are really trying to figure out is whether we have met the standard for the animal efficacy rule. Dr. Golding's comment that we might still like to see some human data still stands. But the question is, what are the approval criteria? Are we approving based on the animal model or not? That's really the decisive question.

DR. FLEMING: Basically, I have been wanting to query along the same lines. I'm trying to think of what is the most practical and feasible, and yet still meets the regulatory and scientific and clinical insight that is needed. The most user-friendly approach that I can see -- and it seems to be the tone of where the question leads us, particularly that second question -- is to think through a possible accelerated approval strategy. The first step in that accelerated approval strategy would be to either rely on animal studies that are reasonably likely to predict clinical benefit or a measure such as disappearance of

detectable venom antigen in, as the FDA was suggesting, 10 patients.

An appealing aspect of this is that it doesn't require the surrogate or the animal studies to be viewed as truly validated. It just requires them to be reasonably likely. If we could hit that hurdle, then that would allow marketing approval and this would then be provided to all people. But there would be the need for the traditional validation trial, although that validation trial -- we would like it to be done in as timely a way as possible, but there are many examples where these validation trials can take a number of years.

An example of that validation trial would then be an historical comparison -- again, trying to go as far as we can to make this feasible -- an historical comparison where everyone would be provided the new intervention and essentially would be compared with historical rates. This requires us to think -- and potentially with an endpoint that would be mortality or intubation, to allow for enhanced quality of care.

What we would need to have is a sense of what a successful level of that rate would be. Could that rate be … the Wyeth vaccine, for example? By the way, this would presumably be enriched by, if not exclusively conducted in, the eastern region, particularly Florida, that is about 60

percent of all coral snake bites, because that seems to be where the risk is high.

If this were done in a reasonably screened population for people that did have coral snake bites in that region, two questions: In such a setting, what is the mortality/intubation rate of the Wyeth vaccine? Is it in the 1 to 4 percent range? If it's in the 1 to 4 percent range, we could hope that that would be the range for a new intervention. What would that rate be if you didn't treat, in the eastern region? We are concerned that we couldn't presume a 15 percent mortality rate. But could we at least presume a 15 percent mortality/intubation rate? If we could -- if the answers to those questions are yes -- then this could be done with a study of 35 to 55 people.

Then, if we went to the concept of a certified region poison center, if that would allow us to capture 15 to 30 people per year, this is a study that could be done in two to four years.

Those are the practical issues. If we believe that the intubation/mortality rate in the Wyeth vaccine is in the 1 to 4 percent range and that a new intervention could be similar … and that we could justify a 15 percent rate in the eastern region for untreated, then we can do this, in a historical postmarketing setting, where everybody is getting the intervention, with 35 to 55

people. Then we would just need to have a sense that the certified region poison center strategy or some other variation of that could give us at least 15 people a year.

DR. TRUNKEY: But that assumes that the Bioclon and the Costa Rican can come up with a DNA equivalent to the coral snake in Florida. Then you test it on either a large- or small-animal model, preferably, in my opinion, large. If they pass that, can that be done in this one-year period that the Wyeth is still available?

DR. HOLLINGER: I don't think we should ignore the opportunity here to move away from horse antiserum, with its potential adverse events of anaphylaxis and serum sickness, to a product that would be much better and cause fewer adverse events.

Also I like what Dr. Alagon and his institute did in looking at a surrogate model, the large model that he mentioned, in which they gave the antiserum after the venom was given. In most studies they just add the two together … and that's not what happens normally. So I like that approach.

I'm also not sure that I agree, Jay, with what you mentioned about the fact that this is not a good surrogate model. I think it's a good surrogate model. Maybe I missed what you said. But I think the animal model is a good surrogate model. Did I misinterpret?

DR. EPSTEIN: The FDA thinks that the model is a good model. We were concurring with your point, though, that premixing antivenom with venom in the typical mouse potency assay doesn't replicate the clinical situation. But it's still informative about the potency of an antivenom.

DR. ZIMRIN: There was a suggestion of using intubation as an endpoint. That's not a hard endpoint. I think Dr. Kitchens' description of someone who was intubated prophylactically with an airway illustrates the difficulty. There might be other people at other hospitals, who would normally intubate with that failure. In a small study, the physician's decision about whether or not somebody who has borderline condition should be intubated I think could really cloud the issue.

DR. FLEMING: So it's essentially addressing Dr. Cryer's observation about whether we could keep people alive without antivenom, just by enhanced quality of care. What do we do to achieve that? What we are doing to achieve that, if that were part of the endpoint, would address that concern. Would intubation be addressing that issue?

DR. ZIMRIN: I just think that it's a difficult endpoint to define. I don't know if Dr. Kitchens wants to …

DR. KITCHENS: It could be one endpoint.

DR. PEREZ: I have a question. What was the reasoning that you don't want to do that?

DR. FLEMING: The data that was presented by a couple of the presentations suggested that, based on the epidemiology, the rate for mortality and serious consequences seems to be much higher in eastern states … some of the data indicated the mortality would be in the vicinity of 20 percent, whereas in the Texas setting, the mortality could be closer to zero, even without having received the antivenom.

DR. PEREZ: But there is no assurance that we know what is out there.

DR. FLEMING: You are saying in the Texas region. In fact, there could be. But I think Dr. Seifert pointed out a number of issues of concern that could lead to a false conclusion of efficacy, and one of those was if you are bringing in a cohort of people that in fact would have inherently been at a very low risk, even without receiving the antivenom, then there isn't an effective way to establish efficacy.

DR. PEREZ: But that was based on a very few studies.

DR. FLEMING: Unfortunately, everything that we know here is based on very few studies.

DR. PEREZ: Then the other questions I had was on licensing -- can you give us an example?

DR. GOLDING: There's another topic today. It's alpha-1 antitrypsin. You are going to hear a lot about that in the afternoon. Essentially the four licenses that are being held today are based on serum levels of alpha-1 PI in the patients. In the threshold there was 11 µM. So that was a surrogate marker. All of those licenses required postmarketing studies. Some of them were done prior to accelerated approval. But the ones that were done under accelerated approval required postmarketing studies to validate that that endpoint was correlated with clinically meaningful endpoints -- to validate clinical benefit.

DR. BALLOW: Actually, it's an important question, because I think we are getting to the point where some of us realize that, at least for approval, an appropriate animal model may be the way to go, with a postmarketing arm. We really haven't talked about these surrogate markers very much at all. I have to agree with my colleague at the other end of the table that the intubation may not be a good surrogate because there is so much variability in the decision points to intubate someone.

I think we need more discussion on surrogate

marker. Perhaps length of hospital stay might be another surrogate marker or other symptoms that patients develop … I'm worried about the intubation. Mortality can be low … I'm worried about the intubation as a surrogate marker.

DR. GOLDING: Can I just add to that? We are also worried about that. We have had that in other scenarios. But you can try to establish criteria that would trigger the intubation. You can in a clinical trial try to make sure that that is consistent in all the centers. For example, you could look at features of respiratory failure using blood gases and so on to establish that the person is in respiratory failure before you intubate.

DR. TRUNKEY: I think we have to be very careful in using length of stay, because there are regional differences. I think you look at mortality, but that's not going to be a great one. But certainly if you develop cranial nerve signs or if you develop a clear-cut indication to intubate the patient, those would be two very good ones.

PARTICIPANT: I was going to comment, as one who has done antivenom studies where there were transports involved, there is a significant potential problem with intubation, as having a different indication, if you have patients that are an hour from your critical-care unit that

are intubated prophylactically for the helicopter transport rather than risk losing their airway on the way, and by the time you can evaluate them pretreatment, they are already heavily sedated and ventilated.

DR. CRYER: I agree with all those comments. Somehow or other, the way that I hear it, and logically so, is that the most common symptom of this venom is a muscular weakness. So you just have to come up with some score for evaluating that before they are sedated.

But that point aside, I think that a lot of what we are talking about is moot, because if we don't come up with some mechanism for approving the drug, there isn't going to be one, and you will have your uncontrolled control group, placebo group, within a year's period of time, with no drug available to the vast majority of people who get envenomated.

So it seems to me that what we really need to talk about is how to design the right approval process rather than whether or not there is going to be a particular kind of study. I don't see how any study can do this before the drug that we currently have runs out.

DR. FLEMING: That's why this discussion is key, because it's empowering the ability to be as flexible as possible in terms of what you require premarketing. If there is a plan for an informative postmarketing component,

then the premarketing animal study or surrogate isn't held to as high a standard of rigor. It just has to be reasonably likely. So the idea is, if we can formulate a plan that we feel would be informative postmarketing, when everybody is receiving this intervention, then the animal study/surrogate endpoint would only have to be reasonably likely to predict clinical benefit. Then in the postmarketing, if we … a failure endpoint as mortality or cranial nerve signs or clear-cut criteria for intubation as an endpoint, as long as we can say, with some reasonable confidence, this is what the rate would be for that endpoint without an intervention, this is what we hope it would be with … then we have a way forward.

DR. CRYER: I agree with that 100 percent. I guess where I had some difficulty was, if we decide we are only going to do this in Florida, for instance, then presumably people in Arizona and Texas wouldn't have drug available.

DR. FLEMING: No, no. They would. Under this plan, it would be broadly marketed. The idea would be, how do we do a postmarketing study? A postmarketing study is very standard, under accelerated approval. It is, in fact, the law. If we do a validation trial, that validation trial doesn't have to be in everybody, but it should be in the representative population that gives us the most

sensitive and timely assessment of efficacy and safety. Everybody could be included for safety. Everybody could be included for efficacy, if we had an endpoint that we believed we could conclude would occur at a high rate in both Texas and Florida.

DR. PEREZ: One of the things … the Texas coral snake is harder to neutralize … the Wyeth antivenom doesn't do a very good job on the Texas coral snake at all.

If I heard right, I think Alejandro said that he had some information he got last night about coral snakes. Do we have time to hear a little bit of that?

DR. SIEGEL: I think we do. But we are up against our noon deadline.

DR. GOLDING: Just regarding availability, it's true that wide distribution would only occur after licensure. But once a sponsor comes in and is under IND, there is a trial, and the trial has to follow the protocol. But there is a way of getting a product to snake-bitten people in Texas or anywhere else, if they are not part of the trial, through the emergency-use IND mechanism. So it wouldn't be as if it would only be available for people in Florida or, depending on the trial --

DR. FLEMING: But what I was asking was, if you took the accelerated approval strategy and you had a lower bar for what it takes to get it out there, the FDA now is

empowered to enable that marketing uniformly to people in need. The validation trial would then be done in a subgroup of those people that would give us the most timely and reliable validation of efficacy and safety. But, meanwhile, it's available to everyone.

The nice thing about this validation trial is that it's not randomizing people to a placebo or even to something else. Everybody is still getting this same antiserum.

DR. SIEGEL: Dr. Bianco.

DR. BIANCO: It's just to bring a perspective … I think what's being discussed now is very positive as a fast way to get licensure and then the Phase IV trial. But let's not forget that there is nobody that wants to invest resources. So whatever we do for this small population of people at risk, it has to be very reasonable and affordable. Otherwise, it's not going to happen.

DR. BORRON: I have a question. Under the scenario that was proposed over here with a surrogate and then postmarketing study, would the animal work that is done to support it have to be done GLP, as it would under the animal rule? That's going to make a huge difference in the cost of being able to do an animal study or animal studies to support this. Probably the best lab in the country doing venom work is not GLP, if I understand

correctly.

Dr. Perez, your laboratory is not GLP?

DR. PEREZ: No.

DR. BORRON: That's the best venom laboratory in the country. You are talking about difficulties in a lab that is not used to working with snakes, having to ramp up snake work when they don't usually do that.

Just an observation.

DR. SIEGEL: Before we address the questions, we need to afford the people outside the committee an opportunity to speak. If there is anyone in the audience who wants to address the committee as part of the open public hearing, please speak up now.

(No response)

Lacking anyone, let's proceed to the questions.

DR. KO: Mr. Chairman, committee members, I'm here again just to reiterate the questions for your discussion.

The first question: Is a clinical trial to assess efficacy of a new coral snake antivenom feasible and practical within a suitable timeframe, using the licensed product as an active control or using a modeled historical control of no antivenom treatment?

Mr. Chairman, would you like me to continue or do you want to --

DR. SIEGEL: Yes, please.

DR. KO: The second question: If the answers to questions 1(a) and 1(b) are no, would the committee agree that the following data are reasonably likely to predict clinical efficacy of coral snake antivenom:

- First, dose determination -- that is, to determine the relative potency with venom neutralization of the new product against the Wyeth product to inhibit a lethal challenge in an established animal model -- then basing the dose of a new product on a proportional adjustment of the dose of the Wyeth antivenom using relative potency in the animal model.
- With the dose obtained that way, do clinical studies on a small number of envenomated patients treated with the new product, 10 or more, to show a point estimate of improved clinical outcome compared to historical controls with no treatment -- for example, 15 percent mortality -- and a consistent decrease in venom levels after treatment.

Also we note that PK, pharmacokinetic, and safety data would be obtained from normal volunteers premarket and then a postmarket study would be needed to confirm product safety and clinical efficacy. The postmarket would take place as a continuation of the pre-licensure clinical trial.

Thank you.

DR. SIEGEL: Thank you very much.

DR. FREAS: We are going to ask for yes votes, we are going to ask for no votes, and we are going to ask for those who abstain. I'm asking you to please raise your hands, keep your hands held up. I'm going to go around and count the number of hands and then, if needed, call the names.

DR. BALLOW: The second option was not exactly framed the way we were thinking, that the product would be approved for distribution based on, say, an animal model. Then postmarketing studies could be performed. That's not the way framed it here. Is that the way I read that?

DR. FLEMING: I believe what we were discussing was a somewhat more flexible approach. It was still the same fundamental approach for approval as we were discussing. Part 2(b) is specifically indicating that the accelerated approval step would be based on looking at this overall rate of effects on venom levels. So it's more based on the surrogate endpoint in humans, while we were discussing an alternative option for that that could be based on an animal study that was reasonably likely to predict clinical benefit. So we were saying in the discussion that there could be two ways of establishing that accelerated approval step.

Otherwise, I think … fairly consistent with what we were discussing, except that question 1(b) -- I think it was raised by Dr. Glynn right at the beginning -- is confusing the way it's stated. The FDA helped to clarify that what they meant by that was, if you were using the 4 percent … confidence interval, it's that type of historical trial not feasible? In fact, what we have been talking about in part 2(b) is a postmarketing study, which … would be an historical trial.

DR. KO: I think, Dr. Fleming, you have explained it very well. As to the question about the pharmacokinetic and safety data, we would need that premarketing. The postmarketing study is basically to verify the clinical benefits.

DR. HOLLINGER: I just have a couple of questions. There are several questions here. One has to think further down the line … let's say a company is approved, has FDA approval for a product, but then that company doesn't make the product anymore or something happens to the company and the product is no longer available. Then the question is, what is your active product? What is your active control? What is the thing that is going to compare the dosing and a variety of other things?

My question is whether the FDA is going to have a

venom sample or something of that nature with which other products can be assessed. I can see a problem down the line here. Let's say the Australian product wants to come in and be approved here, but there is no more Wyeth product available. You almost have to have a standard -- and the FDA has standards for a variety of other things -- either a standard venom or something like this to which other products can be compared.

DR. TRUNKEY: I personally think there are … in regards to Texas, because we were told that sometimes clinicians prefer not to give the antivenom. You are saying that the DNA may be different in your snakes. It may be a different venom, because it doesn't neutralize … but the other thing is, we should essentially have zero percent mortality if the patient gets to an appropriate facility, because even if the antivenom didn't work, you are going to be able to put them on a ventilator and keep them there. So I think you have to use clinical symptoms. I'm not really that concerned that we are going to lose patients with new treatment modalities, provided we have standard …

I would just put that out there for our consideration. The FDA has to be very flexible if they move forward with this within this one-year period.

DR. PEREZ: (I would also agree. I think

traditionally … it's whether we use pit viper venom or coral snake venom. It's still horse serum … so I think that … based on what you said about proper medical treatment …

DR. EPSTEIN: I think what is bothering some committee members is that you would like to be able to vote on a third alternative, which is whether efficacy of coral snake antivenom can be shown in a suitable animal model. I think logically you would vote no to question 2 and yes to an additional question that we would add.

Just to make sure that the issue is crystal-clear, a vote yes on question 2 means that you are accepting the clinical study of a point estimate of failure, whatever we call failure, less than 15 percent and reduction in serum venom level as the surrogate for efficacy under accelerated approval. If you then say that efficacy can be established solely based on an animal model, you are saying that you believe that the animal efficacy rule can be applied, that it's scientifically valid, and that there is feasible or practical alternative.

DR. FLEMING: It seems like there is a middle ground, if I'm understanding this. That middle ground is to say we would like to be able to use animal studies. Animal studies can provide insight. It's really problematic, though, to say they are going to reliably

fully tell us about efficacy and safety. But if those animal studies give you a rapid way to get this product through accelerated approval into the market, and still have the validation trials you have for accelerated approval, then that provides a reassurance that you are going to get enhanced understanding about benefit and risk.

It is loosening up somewhat the second question to say that the accelerated approval step could be based not just on a clinical human study that shows that you are reducing the serum levels, but you could also have an animal study that is reasonably likely to predict clinical benefit that serves as the basis of getting the product into the market. But there is some protection there, because you would then have a number of years available to then assess safety and efficacy in humans directly.

Could we have that looser formulation for question 2?

DR. EPSTEIN: You may have to leave it to FDA to decide whether the regs allow us to use data solely in animals as a basis for accelerated approval. What we are really asking is whether a study in 10 patients looking at the markers as proposed is practical and feasible. If it's practical and feasible, then we actually can't apply the animal efficacy rule.

DR. FLEMING: I understood Subpart H does allow

for animal studies as a source of information. Is that not correct?

DR. EPSTEIN: Are you saying the sole source?

DR. FLEMING: I understood that it did. Is that not correct?

DR. EPSTEIN: I may not be able to answer that on my feet right now. I think we have to look back at the regs.

DR. FLEMING: If you took the approach of the animal studies, is that going to have a postmarketing --

DR. EPSTEIN: Yes, FDA would certainly want that in any case.

DR. FLEMING: I think the second and third options are really not that different.

DR. EPSTEIN: There is a crucial difference, though. The crucial difference is whether, pre-approval, we want a human study of at least 10 subjects. That's the difference.

DR. CRYER: That doesn't seem to be possible.

DR. EPSTEIN: It may not be.

DR. CRYER: There is no ethical way that you can actually ask somebody who just got snake-bit consent on taking a new drug. There's just no way you can do that.

DR. GLYNN: If you want to somehow recommend to the FDA that you would like to have these products marketed

and then a postmarketing surveillance study, how do you vote? I'm confused. Do you vote no or yes to 2(b)?

DR. FLEMING: While FDA is waiting to answer that, it was my understanding that Subpart H -- it's rare, but this is an extreme case -- it's my understanding that in such an extreme case Subpart H doesn't exclude the possibility of having the accelerated approval information coming from animal studies.

DR. EPSTEIN: Could I just suggest -- the advisory committee advises us on the science. The FDA has to determine what the legal pathway is here. What we really wanted the committee to advise on was, what would be sufficient information for the product to become widely available? I think you can advise on that and let us deal with which mechanism we are actually applying.

But again, the crux of the question is whether you want to see human data before an approval. What we have proposed is that you could look at ELISA levels and you could look at a point estimate of serious adverse events. Either you feel that's a necessary step or you feel that it is not. That's scientific advice. Then let us worry about the framework.

DR. ZIMRIN: Do we have to include both parts of Part (b)? … I can see a difference between venom levels. But I think looking at 10 patients is not very likely to

produce useful information.

DR. EPSTEIN: I would agree. But, of course, if you had two failures out of 10, you would know you are not dealing with an effective product.

This is a point for discussion, obviously. That's why we have put it up there.

DR. FLEMING: Personally, I would rather have the human data, if it were feasible, practical, and timely. If we found 10 patients, do we even know that we would be able to get reportable levels pre and post to determine the venom levels?

So there is one issue as to whether it's even a valid or reasonable outcome. Let's say it is. Do we even know that we could get pre and post assessments that would allow us to assess that?

DR. COLVIN: One of the other difficulties, it seems like, looking at a small trial like this, looking at clinical outcomes, is that the timing of the patient getting from the bite to the hospital, then maybe not getting any kind of antivenom at all, then progressing neurologically, then maybe they would require intubation. But based on everything we heard you still would have a problem.

DR. KULKARNI: (Off mic - inaudible)

DR. CRYER: The practicality is, when you use

these things, you basically have somebody who is developing symptoms that they develop when you give them more. If they develop, then you give them moreand the dose is going to be irrelevant in the actual clinical situation. As soon as that patient develops symptoms on the drug, are you going to give them more.

DR. COLVIN: One way to sort of answer it -- it's not a great way -- if the equivalency trial was done in two different species. If they are the same, if the equivalencies are the same in two different species, that makes it more likely that the third species … rather than being just one species.

DR. FREAS: For the 1(a), could I have a show of hands from all those who vote yes to question 1(a)?

(No response)

That is zero hands raised.

Could I have a show of hands from all those voting no to question 1(a)?

(Show of hands)

It is a no vote.

On question 1(b), could I have a show of hands … be specific, because it's too vaguely stated or this is the model of historical control using lower limit of the confidence interval of 4 percent. It's a different question-- that is what you need to ask.

There are 15 voting people at the table right now. A unanimous vote, 15 yes votes (**CORRECTION TO TRANSCTIPTS,** THE CORRECT COUNT WAS 15 "**NO**" VOTES, W.F. 8/11/2009).

On question 1(b), could I have a show of hands for all the yes votes?

(No response)

Could I have a show of hands for all the no votes?

(Show of hands)

As with the first question a, question b is unanimous no votes.

Now we are going to go to question 2(a). Could I have a show of hands for a yes vote to question 2(a)?

(Show of hands)

That's 14 yes votes -- raise your hands one more time. Yes vote for Dr. Cryer, Dr. Zimrin, Dr. Younger, Dr. Ballow, Dr. McComas, Dr. Colvin, Colonel Rentas, Dr. Siegal, Dr. Trunkey, Dr. Glynn, Dr. Fleming, Dr. Perez, and Dr. Bower.

Could I have a show of hands for the no votes.

(Show of one hand)

There is one no vote, Dr. Kulkarni.

Could I have a show of hands for abstaining?

(Show of one hand)

There is one abstaining, and that's Dr. Hollinger.

Now we are at question 2(b). Could I have a show of hands for yes votes for question 2(b)?

(Show of one hand)

There is one yes vote for question 2(b).

DR. FLEMING: For 2(b), are we interpreting this in a more inclusive way?

DR. EPSTEIN: 2(b) means a clinical study pre-approval. That is accelerated approval.

We have reviewed the reg on accelerated approval, Subpart H. It's silent on animal data. I don't know if there are any precedents for using solely animal data.

But that point aside, what we are asking the committee is whether there should be a human study, however small, before approval. We think that the votes in the negative on question 1 open the door both to an accelerated approval and to the animal efficacy rule. We think we could go either way. We are asking the committee, do you think we need human data before we approve, pre-approval? That's what 2(b) means.

If the answer is yes, then you have to say what data -- and maybe you want to comment on that -- and/or vote separately on the sub-points, because, as Dr. Zimrin pointed out, there may be different opinions, whether a

point estimate on 10 patients is of any meaning whatever. You might, in fact, take one opinion on that point and independently either accept or not accept venom levels as a useful metric.

Just two points here. One, 2(b) is asking the committee whether there should be human trial data before approval. That would be compatible with accelerated approval. It doesn't preclude asking the third question, which we will pose, which is whether efficacy of coral snake antivenom can be shown solely in an animal model. That's the animal efficacy rule.

Again, having been advised that a clinical study, conventional study, is neither feasible nor practical, we do believe the door is open to an animal efficacy study.

DR. BALLOW: I don't think it's unreasonable to do a small study. The first paragraph bothers me, whether 10 patients is going to be enough to answer any kind of efficacy question.

DR. EPSTEIN: Let me just be clear. The footnote is that FDA would require a PK study and indeed a safety study in normal volunteers in any case.

Are you saying a PK study in envenomed people? We would do a PK study in normals, regardless, and a safety study in normals, regardless.

We are really only asking whether we should have

a small efficacy study in humans using a surrogate marker. That's the accelerated approval question.

DR. KO: Let me add to what Dr. Epstein just said. He is perfectly correct that the regulations are silent about using animal data for accelerated approval. However, there has been precedence, including the approval of ciprofloxacin for anthrax. So it is not unprecedented.

DR. HOLLINGER: If I understand this correctly, since you mentioned to get the human side that you want us to answer this based on scientific evidence that a human study will provide this information before you are going to approve it, then I'm not sure how we can do that.

DR. EPSTEIN: Then you would vote no.

Perhaps the point is that the committee may need to discuss whether there are likely valid clinical endpoints that could be used in a small trial. We have given you the two we think you could look at, but maybe there are others.

But the bottom-line question is whether we should require even a small human trial. Maybe we should just ask that question, independently of the subparts, and have an essay question on what surrogate markers you would use.

Let's vote on 2(b), clinical studies on a small number of envenomated patients treated with the new product, 10 or more. Strike the word "showing." Then, for

those who vote yes, ask, what markers would you use?

Is that clear enough for the record?

DR. FREAS: We will vote now on a modified question 2(b), as given by Dr. Epstein. Can I have a show of hands for all those voting yes on modified question 2(b)?

(No response)

There are no yes votes for question 2(b).

Could I have a show of hands for the no votes for question 2(b)?

(Show of hands)

There should be 14 yes votes. They are Dr. Cryer, Dr. Zimrin, Dr. Younger, Dr. Ballow, Dr. McComas, Dr. Colvin, Dr. Kulkarni, Dr. Rentas, Dr. Siegal, Dr. Trunkey, Dr. Glynn, Dr. Hollinger, Dr. Perez, Dr. Bower.

Could I have a show of hands for no vote.

(Show of one hand)

Dr. Tom Fleming abstains.

DR. EPSTEIN: I propose that the committee be offered the following question to vote. Can efficacy of coral snake antivenom be shown solely in an animal model?

We are really just asking, should FDA apply the animal efficacy rule?

DR. FLEMING: For clarification, with a commitment for a validation trial postmarketing, is that

correct?

DR. EPSTEIN: That would be our intent. I'm not sure that's required by the animal efficacy rule. It is required? Okay. So it's implicit.

DR. FLEMING: That makes a big difference. So for clarification you would still have the validation trial … just as you would the accelerated approval.

DR. EPSTEIN: The postmarket trial would be the same.

DR. FREAS: For question 2(c), modified by Dr. Epstein, could I have a show of hands for the yes votes?

(Show of hands)

That is a unanimous yes vote.

DR. SIEGEL: Thank you all for your attention. We will adjourn for lunch.

(Whereupon, at 12:20, the meeting was recessed for lunch.)

AFTERNOON SESSION

DR. SIEGEL: Good afternoon.

Bill is going to make a couple of announcements before we start.

DR. FREAS: I would like to welcome three new people to the table for this afternoon's session. Please raise your hand as I call your name.

Dr. James Stoller, Jean Wall Bennett professor of medicine, Cleveland Clinic College of Medicine.

Dr. Peter Terry, professor of medicine, Division of Pulmonary and Critical Care Medicine, Johns Hopkins University.

Sitting next to Dr. Terry, Dr. Peter Choyke, chief, Molecular Imaging Program, National Institutes of Health.

My welcome to all three of you.

DR. SIEGEL: The topic for this afternoon is "Clinical and Surrogate Endpoints for Evaluating Efficacy of Alpha-1 Proteinase Inhibitor (Human) Augmentation Therapy."

The first speaker will be L. Ross Pierce, from OBRR, FDA, who will introduce the topic.

Agenda Item: Clinical and Surrogate Endpoints for Evaluating Efficacy of Alpha-1 Proteinase Inhibitor (Human) Augmentation Therapy - Introduction

DR. PIERCE: Thank you.

The title of my presentation is “Clinical and Surrogate Endpoints for Evaluating Efficacy of Alpha-1 Proteinase Inhibitor (Human) Augmentation Therapy.”

Data accumulated since the original approval of the first alpha-1 proteinase inhibitor product, Prolastin, in 1987 leads us to reconsider whether there is an adequate basis to continue to use the historical serum trough target level to assess efficacy of augmentation therapy in emphysema due to alpha-1 antitrypsin deficiency. Epidemiologic data leads us to question whether the postulated historical target of 11 µM is necessarily an optimal protective threshold. Several laboratories have demonstrated that severely alpha-1 antitrypsin-deficient patients have abnormally high lung neutrophils and neutrophil elastase concentrations, and this finding appears not to have been taken into account by Gadek and Crystal when they originally formulated the 11-µM threshold, which came out of comparing different phenotypes and their associated risk of emphysema.

The questions that BPAC will consider during this session include:

CBER has identified serial high-resolution computerized tomography (HRCT) lung-density measurements as an appropriate, clinically meaningful endpoint to assess

the efficacy of augmentation therapy with intravenous alpha-1 proteinase inhibitor products on emphysema disease progression. The question is, does the committee agree that the rate of change of lung density as measured by serial HRCT could potentially be used as a primary endpoint in pivotal studies of efficacy of alpha-1 proteinase inhibitor augmentation for inhalation therapy?

The sub-question 1(a) is, before embarking on pivotal studies, should sponsors first establish to what extent CT density measurements are confounded by, one, inhalation therapy itself, and two, exacerbations? You might ask, how could inhalation therapy confound CT measurements? We have seen instances where some investigational inhaled alpha-1 products have been associated with a picture of hypersensitivity pneumonitis. There is also the possibility of a direct irritant function. We have had products that are both dry powder inhaled and aerosol inhaled, liquid aerosol, and there is theoretically the possibility of irritation and inflammation and, were that to progress over a long period of time, even possibly fibrosis.

These are theoretical questions.

1(b): Does the committee recommend that any additional information regarding HRCT lung-density measurements be obtained prior to sponsors initiating

pivotal studies of efficacy of alpha-1 PI augmentation for inhalation therapy?

Question number 2: Does the committee recommend that FDA reconsider the use of biochemical surrogate endpoints of serum and epithelial lining fluid (lung) antigenic and functional alpha-1 PI levels, to provide substantial evidence of efficacy pre-licensure of new IV therapy products in favor of more clinically meaningful endpoints -- i.e., HRCT lung density, FEV_1, pulmonary exacerbations, or mortality?

I should define new alpha-1 PI products. This would not include products of any sponsor that has already come to us at FDA and negotiated with us a path to licensure, but rather totally new products that have not come to us as yet, by the intravenous route.

Question number 3: Does the committee recommend any other alternatives as primary endpoints for alpha-1 proteinase inhibitor premarketing clinical trials, either (a) for inhalation therapy, or (b) for new submissions of IV therapy (new as I just defined it)?

Question number 4: Does the committee recommend that studies of intravenous alpha-1 PI augmentation therapy include higher doses than previously approved, assuming adequate safety?

Today’s meeting occurs in the context of a series

of meetings that FDA has been involved with regarding the alpha-1 proteinase inhibitor therapeutic story. Not the first, but the first one that I would like to highlight is a meeting in 1985, a workshop jointly sponsored by FDA and the NHLBI, where the participants recommended the use of biochemical surrogate endpoints, serum and lung epithelial lining fluid, alpha-1 proteinase inhibitor levels, both antigenic and functional, to evaluate the efficacy of alpha-1 proteinase inhibitor augmentation therapy products.

So there are really four surrogate endpoints: the antigenic, the functional, or antineutrophil elastase capacity measurements, both in the serum and in epithelial lining fluid taken during bronchopulmonary lavage.

In 1987, Prolastin, the first product in this class, was licensed using biochemical surrogate endpoints -- somewhat of a different level of those endpoints compared to what had been recommended at the workshop in 1985, as I'll go into in my second presentation today -- including the theoretical protective threshold serum alpha-1 proteinase inhibitor level proposed by Gadek and Crystal, as well as epithelial lining fluid alpha-1 proteinase inhibitor levels.

Eleven years later, in 1998, the majority of BPAC members voted to recommend that FDA continue to accept the biochemical surrogate endpoints of serum and epithelial

lining fluid concentrations of antigenic and functional alpha-1 PI to provide substantial evidence of efficacy to support product licensure of new alpha-1 PI intravenous products. Eleven voted yes, three no, and one abstention.

Over that 11-year period, only Prolastin was on the market in the United States. We convened that BPAC after the results of the large NHLBI registry study and epidemiologic study of alpha-1 PI, including some patients on augmentation therapy, as well as the natural history in unaugmented patients, was published.

In 1999, the Alpha One Foundation held a workshop at which the opinion was voiced that the licensure of aerosol alpha-1 PI products would require clinical trials evaluating clinically meaningful endpoints rather than relying only on biochemical surrogates. At the time, the latter were deemed inadequate for aerosol products' efficacy evaluation.

In 2001, the Alpha One Foundation held the first of two workshops on lung CT. At this workshop, the summary paper that came out of it recommended the density corresponding to the 15^{th} percentile in the distribution of lung voxel -- a voxel is a three-dimensional pixel -- as the primary endpoint of HRCT clinical trials evaluating products for emphysema.

Then in 2005, there was a BPAC that evaluated the

biochemical heterogeneity among the licensed alpha-1 PI products. It came to light that using isoelectric focusing, one of the approved products -- two-thirds of the molecules in the vial were not corresponding to the primary sequence of alpha-1 proteinase inhibitor, but had a single amino acid truncation that did not appear to affect the activity of the product *in vitro*.

At that BPAC, designs of Phase 4 randomized studies of newer alpha-1 proteinase inhibitor products using clinically meaningful endpoints -- HRCT, FEV_1, pulmonary exacerbations, and/or mortality -- were discussed. The BPAC members supported FDA's plan for the postmarketing commitment studies of clinically meaningful endpoints for the licensed alpha-1 PI products.

In 2008, the second workshop sponsored by the Alpha One Foundation on lung CT was held. A variety of details with respect to study methodology, possible endpoints, and future applications of that technology for clinical trials were discussed.

In March of this year, FDA cosponsored a workshop with the Alpha One Foundation and HHS. The title of the workshop was "Improving Endpoints, Improving Care: Alpha-1 Proteinase Inhibitor Augmentation Therapy and Clinical Trials." I'll highlight some of the outcomes of that workshop and some of the questions which were considered by

workshop participants.

The workshop participants agreed that HRCT lung density appeared to be a sensitive and clinically meaningful endpoint to evaluate the efficacy of therapeutic products for emphysema due to severe AAT deficiency, although there was no formal vote. Information regarding a number of older and newer surrogate biochemical endpoint was presented, and the consensus appeared to be that these biochemical endpoints still need further evaluation and may be used to help with dose selection of new products, particularly when used in combination, for example.

Topics considered at the workshop included:

- Whom to enroll in clinical trials for evaluation of AAT augmentation therapy.
- What has been the clinical-trial endpoints experience today?
- What are current and potential endpoints for clinical trials, divided into the categories of functional endpoints, radiological endpoints, and biochemical endpoints?

Selected questions that were considered at the workshop included:

- How should disease severity and rate of progress impact patient subset selection for enrollment in augmentation therapy trials?

- What have been the major challenges to development of endpoints for clinical trials of alpha-1 PI augmentation therapy, and how might these be ameliorated?
- How strong is the need for dose-ranging studies in evaluating alpha-1 PI augmentation therapy?
- What are thought to be the most useful currently available functional predictors of clinical efficacy in alpha-1 proteinase inhibitor deficiency? What are the advantages and disadvantages of these measurements?
- What has been learned from ongoing and recently completed studies that is useful for future HRCT endpoint studies?

With this, we can go to the next presentation, unless there are any questions.

DR. SIEGEL: Unless there are some burning questions, let’s go to the natural history and pathogenesis of AAT deficiency, presented by Dr. Kenneth Chapman from the University of Toronto.

Agenda Item: AAT Deficiency: Natural History and Pathogenesis

DR. CHAPMAN: Thank you very much.

I believe in this afternoon’s series of presentations, I represent the ghost of endpoints past. As you deliberate on evolving endpoints and hear from experts at the leading edge, it may be useful to put some of the

current information into a context. I'll try not to duplicate what's already in your wonderful briefing document, which is certainly very detailed and very accurate and up-to-date, but to share with you some of the background that may not appear or be highlighted in the briefing document.

I'll touch briefly on the discovery of this abnormality, the pulmonary consequences of alpha-1 antitrypsin deficiency, the natural history studies, and a meta-analysis of studies looking at the impact of augmentation therapy. I'll offer a word or two of summary.

This is probably a slide that's well understood by the members of the panel -- perhaps better understood by them than by me. This is work of people doing electrophoresis. This is the sort of slide that Lurell and Erickson (phonetic) looked at in the 1960s. The topmost slide is the abnormal one. The electrophoresis pattern is missing a band here that is present on the bottom. This is a normal protein electrophoresis. Lurell and Erickson understood, after back-checking where these samples came from, that they were seeing an abnormality that was present in 1 in 1,500, 1 in 4,000 of the Scandinavian population. The missing glycoprotein was a 52-kilodalton protein. When it was missing, people had a very high risk of developing emphysema. The pivotal paper describing this deficiency

was published in 1963, and since then, the glycoprotein has been characterized and the genetics have been characterized much more fully.

We have a diagram here from a paper of Dr. Brantly's. He will be presenting a little later this afternoon. The key message to pulmonologists became that the glycoprotein had a major function of inhibiting neutrophil elastase by binding to it -- something of great importance. We will be talking for the most part this afternoon, of course, about the ZZ abnormality of alpha-1 antitrypsin deficiency, but there are more than 100 genetic variations associated with deficiency.

This little cartoon is meant to remind us of the pulmonologist's view of alpha-1 antitrypsin, which is that its job is to be present in the lung and to bind to neutrophil elastase to prevent that neutrophil elastase from damaging the lung. It does so in this way.

It's a worthwhile picture to pause on. This afternoon you will be hearing about voxels and Hounsfield units and 15th percentiles. But clearly the picture on the left is a dense, spongy lung parenchyma; that on the right is the loose-knit lung parenchyma of someone with emphysema. The discussion is around an endpoint to measure the progression from left to right and to see if we can detect a slowing in the loss of lung density with our

various therapies.

I want to pause with this slide. It is a slide that's meant to talk about the number of individuals in Western nations who have alpha-1 antitrypsin deficiency. This is an estimate -- a guestimate, if you will -- of the numbers of individuals with alpha-1 antitrypsin deficiency in North American and Europe, 150,000. The very smallest of the columns, just over 4,000, would be the number of those individuals currently treated with augmentation therapy.

This slide is usually put up in a teaching session or discussion session to talk about physicians failing to screen in appropriate populations for deficient individuals. But it may be worth pointing out in this context that this slide also hints at the tremendous range of natural histories of this disorder. The second-most column, estimated at 54,000, would be the estimated symptomatic individuals. Presumably there are a great many individuals with the deficiency who don't have clinically overt or manifest disease, at least not to the extent that it's bringing them to the attention of physicians.

We understand one obvious reason for the tremendous natural-history range -- smoking and other injurious exposures, but most of all, tobacco smoking -- but we have yet to unravel some of the other cofactors. We

suspect that having the genes associated with ATP and asthma may be a bad set of cofactors for somebody who is also alpha-1 antitrypsin-deficient, but we have to prove that in any concrete way or to show in a concrete way any other genetic cofactor that may worsen prognosis in alpha-1 antitrypsin deficiency.

One thing we do, of course, understand is that there are different phenotypes. We have a different electrophoresis now. We all hope that our own phenotypic background is that shown on the left. We see the heavy bands of M alpha-1 protein present in this individual. We don’t see the Z bands in the individual diagrammed on the far right. This would be a ZZ homozygous individual with alpha-1 antitrypsin deficiency. The pattern on the left is the individual with one M and one Z allele, the carrier or heterozygote.

This is something that Dr. Sandhaus will spend more time with. We talk about the risk of developing alpha-1 antitrypsin deficiency relative to one’s phenotype and the range of serum concentrations associated with that. On the y-axis, the estimated range of serum alpha-1 antitrypsin concentrations at steady state. We have the MM, or healthy, normal individual, with two wild-type alleles. We have over here the individual we’ll spend most of our time talking about, the individual with two Z

alleles, or deficiency alleles. We won't, perhaps, talk directly about these individuals, the individuals with null alleles and no measurable levels of alpha-1 antitrypsin protein. Dr. Sandhaus, I'm sure, will talk about this individual, the SZ individual. You see on this standard teaching-type slide the electro-micromolar protective threshold, as it's labeled here. That is based on earlier estimates that are now up for discussion, the suggestion being that, given the increased risk of emphysema in this population -- I think in your briefing document it's given as a relative risk of 3.5 -- perhaps the protective level shouldn't be where these individuals spend most of their time, but should be somewhat higher.

Monitoring disease progress and the endpoint of alpha-1 research past would, of course, be based on spirometry, a technology that is about a century or so old. It's very simple and reproducible. But, of course, it has its limits, and the FDA and its advisory committees have been working on developing better endpoints in the field of COPD in general. Of course, the discussion is about developing or improving upon endpoints in alpha-1 antitrypsin deficiency emphysema in particular.

What I would like to review now is the outcome of a couple of studies and a meta-analysis we did, a meta-analysis of studies that tracked the natural history and

the treated history of alpha-1 antitrypsin deficiency using that endpoint, FEV_1. You will see that we start with a great deal of literature, as meta-analyses are wont to -- 247 papers that listed the relevant search terms. After weeding out redundant and data-free publications, we ended up with just five studies, plus some additional unpublished data of our own, to put into our meta-analysis. You will see the single biggest of these studies would be the data of the NHLBI registry from this country.

Listed here are the studies themselves, these published studies. I'll come to the NHLBI registry study in a moment. You will hear from Professor Dirksen this afternoon. I'll show you some Canadian data and some natural-history data from the U.K. registry of Rob Stockley's.

The single largest study -- a natural-history study, if you will -- is from this country. I don't believe Dr. Sandhaus is going to be showing these tables, so I'll do this, if I may.

This is a table from the registry established in this country, tracking individuals with alpha-1 antitrypsin deficiency for several years -- a total of just under 1,000 subjects. For the purposes of this afternoon's discussion, the registry researchers compared the individuals in this leftmost column -- almost 300 individuals -- who were not

treated for their alpha-1 antitrypsin deficiency with augmentation therapy, to people in these two columns, the middle column and the right-hand column, people who either received augmentation therapy some of the time or continuously through the monitoring period.

I wanted, in showing this table, to highlight one of the problems with the natural-history data. We note, for example, that the individuals who weren't receiving therapy -- this was, remember, a natural-history study; this was not a randomized, controlled trial -- were individuals who were much more likely to be nonsmokers. Forty-one percent of them were never smokers, versus 15 or 11 percent of those who did receive augmentation therapy. This is clearly, in that perspective, a comparison of apples and oranges.

Or this other major variable: The presence of pulmonary symptoms. Present in not quite half the individuals not on augmentation therapy versus more than 80 percent of those who were receiving augmentation therapy.

This speaks to the natural-history variability. This appears to be a different population. Most likely these are the individuals in the registry who are identified by means of family screening. These are probably the individuals identified on the basis of symptoms.

We have more of the same on the following. If we look at lung-function measurements -- and we are now back to our FEV_1 -- more than half of the individuals not on augmentation therapy had normal levels of FEV_1 -- that is, greater than 80 percent of predicted -- versus a very small handful of individuals in the augmentation therapy groups.

Bronchodilator response: 18 percent versus 36 and 34 percent.

Mean FEV_1: 74 percent, just under the normal cutoff value, versus 41 and 37 percent.

Clearly, very different populations. You might argue that these are the individuals who probably didn't need augmentation therapy -- or not as many of them did -- and were perhaps those with a very different and perhaps more benign natural history.

Nonetheless, if we look at the outcomes -- and we see diagrammed here the rate of FEV_1 decline in millimeters per year for different subpopulations based on their FEV_1s -- the leftmost people are those with very low baseline FEV_1s, less than 35 percent of predicted, those with moderate or even mild impairment, and those with normal or nearly normal FEV_1s. The difference between patients not receiving and receiving augmentation therapy is not apparent in the severely obstructed patients. It is apparent in those with moderate obstruction. It's

apparent, but it's a very small number of individuals, up at this high end of the scale.

This is perhaps a better representation. This is from the paper itself, a scan of the outcomes on a continuous scale. This is mean baseline FEV_1 as a percentage of predicted across the x-axis. Here is the FEV_1 rate of decline. We see that the difference between not being augmented and being augmented is evident in the middle of this range, the moderately obstructed patients.

Those in the alpha-1 community view the two ends of the spectrum in very different ways. These are perhaps patients who have different natural histories. These are the people with nearly normal FEV_1s and benign courses. The numbers are also very small. Over here, the severely obstructed patients, we assume, are people where we have problems with sensor data -- that is, much loss of lung function at that end of the scale -- that eliminates the patient from further study and further measurement.

We have another way of looking at FEV_1s. Dr. Wenker's group in Germany looked at patients who were not declining rapidly before augmentation and they continued their benign course after or patients -- and these are individual patients -- who declined rapidly, received augmentation therapy, and their rate of decline slowed.

The Canadian registry was an attempt to gather up

patients who were treated and non-treated in Canada, where access to health care is more universal. We identified 21 patients receiving augmentation therapy, 42 matched patients who were not. They were very well matched by the characteristics shown on this table. We'll just show the summary slide. The average FEV_1 decline was 30 mL per year in the patients who received augmentation therapy. This, by the way, is a commonly quoted normal rate of FEV_1 decline for a healthy nonsmoker without COPD. In non-augmented patients, or control patients, the rate of decline was twice that.

The U.K. registry data are simply untreated data or natural-history data. In this case, the rate of FEV_1 decline is shown moving upward rather than downward. It's worth noting, perhaps, that we again have this variability in disease progression, patients with very severe obstruction not showing much change in rate of FEV_1 decline -- again, there's not much room to move at that end of the scale -- similarly, a low rate of annual change in FEV_1 for the much less obstructed or nearly normal patients. The differences we see tend to be in this middle, FEV_1s ranging between 30 and 80 percent of predicted.

So our meta-analysis, putting all of the data together: I'll show it by baseline FEV_1, the most severely obstructed patients. The data points on this forest plot

to the right of center suggest a benefit of augmentation therapy. This is the number in terms of FEV_1 decline in mL per minute. It's very small and non-significant in these severely obstructed patients.

However, in the patients in that middle, 30 to 65 percent of predicted, the average slowing in FEV_1 decline, based on this meta-analysis, is 18 mL per year. In your briefing document there is an attempt to put this into context. It's fair to pause at this point and say, despite more than a century of using spirometry and more than half a century of trying to track the natural history of COPD in general and almost half a century in this field, we have not defined, as pulmonologists, the MCID for differences in FEV_1 decline.

This is the nearly normal patients, with FEV_1 greater that 65 percent of predicted, a trend that is non-significant.

Here, overall from our meta-analysis, using all patients and all available data: Slowing the rate of FEV_1 decline of just over 13 mL per year, favoring augmentation therapy versus the lack of augmentation therapy.

I'll just pause with a slight detour at the end. I was asked at the meeting in March to talk about Canadian perspectives on endpoints and the FEV_1 calculations. I'll just remind you, the American Thoracic Society and ERS

statement about augmentation therapy suggests it can be used in deficient individuals who have an FEV_1 between the benchmarks we have been talking about. That's the indication for alpha-1 antitrypsin augmentation therapy. But there is a footnote which says, "And in those who are declining rapidly, greater than 120 mL per year, one might consider augmentation therapy."

The Canadian Thoracic Society's most recent statement, which is now somewhat dated, suggests that we want to augment only those individuals who are deficient, who have otherwise optimized their care by quitting smoking and receiving other less specific medical therapies, who are in a range -- this is a very narrow range -- of FEV_1 at baseline, and who have a rapid rate of decline. This time the benchmark, somewhat arbitrarily, is given as greater than 80 mL per year rather than 120 mL per year.

I want to pause and say that it's actually very difficult in a clinical context, and perhaps in a clinical-trial context, to operationalize this sort of definition. Our staff at the registry has been looking at how to do this. The recommendations in the Canadian Thoracic Society's statement suggest at least three post-bronchodilator measurements of FEV_1, preferably at certain intervals and preferably over at least two years.

Here is one of our patients in the registry. We

see FEV_1 tracked over a dozen years. Here are post-bronchodilator FEV_1s. We have a nice, tight line of regression. This certainly seems to be a stable patient in terms of rate of lung-function loss.

Our staff tried to apply with these measurements calculations of rates of FEV_1 decline. What you see here are moving time averages over two years. We see increase or decrease in FEV_1. That dashed red line would be the Canadian Thoracic Society's benchmark of rapid decline -- that is, more than 80 mL per year. We see that the moving time average, if a clinician or clinical researcher were to average over two years, is all over the map -- people, presumably, gaining lung function or losing lung function dramatically. We are just going through our database to try to understand what the sensitivity/specificity of certain calculations is.

Things improve a bit, tighten up a bit, predictably, with three-year moving time averages. Then a continuous moving average incorporating all the available data, of course, is the best possible outcome.

I won't show other examples, but simply say, as one attempts to understand the optimal patient for augmentation therapy, the FEV_1 measurement becomes a very tricky measurement when applied in the individual case.

To summarize, augmentation therapy, in terms of

the traditional endpoint, lung-function measurement, FEV_1, reduces the rate of FEV_1 decline. Most of the data, as you know from your briefing document, is natural-history, non-randomly assigned therapy, with a couple of exceptions.

The baseline FEV_1 or baseline rate of FEV_1 decline appears to be an important predictor of benefit. But calculating this baseline rate of FEV_1 decline is a challenge in individual patients, and baseline FEV_1 seems to be the best surrogate.

The limitations of our analysis of FEV_1 decline include:

- As mentioned, the lack of randomized, controlled data.
- The endpoint, which is certainly up for discussion.
- As we should mention, the variable dosing schedules that have been used in previous studies, all of which used different dosing schedules -- one infusion per week, infusions every two weeks, or, very early on, infusion on a monthly basis.

Thank you.

DR. SIEGEL: Thank you very much, Dr. Chapman.

Next we will hear a “Review of Epidemiological Studies Using Augmentation Therapy and Serum AAT Levels as a Surrogate Endpoint to Evaluation Efficacy of Augmentation

Therapy," Robert Sandhaus, M.D., Ph.D., National Jewish.

Agenda Item: Review of Epidemiological Studies Using Augmentation Therapy and Serum AAT Levels as a Surrogate Endpoint to Evaluation Efficacy of Augmentation Therapy

DR. SANDHAUS: Thank you very much. I appreciate the opportunity to address you on this interesting question.

I have provided a more compact title for my talk: "Do we actually know what dose of augmentation therapy is correct?"

I'm going to discuss the story of our magic protective level of 11 µM and its history, review the studies that have been done using this as the endpoint for the clinical trials, focus in on one particular group of papers that have been published, the story of the individuals with the PiSZ genotype, and then summarize my conclusions from these.

Basically, once we understood the mechanism of disease in alpha-1 antitrypsin deficiency, the logical next step was to develop therapies based on that understanding. Through work that was primarily done in Ron Crystal's group at the NIH, we were able to see a proof of concept for purification methods from plasma of alpha-1 antitrypsin concentrate and to develop biochemical endpoints based on

blood levels of alpha-1 antitrypsin and, as was mentioned, the epithelial lining fluid levels.

Clinical efficacy was evaluated in the 1985 meeting that Ross Pierce mentioned. Essentially the group decided that it was an impossibility at that time to do clinical efficacy studies because of the N required versus the number of patients who had been identified with alpha-1 antitrypsin deficiency at that time and the duration of therapy required -- or at least expected to be required -- to look at a clinical endpoint like FEV_1 spirometry.

Here's a slide familiar to you already from Dr. Chapman's talk. I put this up only to make the broad statement that it has been generally assumed -- the current dogma that guides our therapy is that the risk of lung disease is directly related to the serum levels of alpha-1 antitrypsin. As you'll see, some data exists that calls that central dogma into some question.

To embark on the history of our magic 11-µM protective level, in fact the very first publication regarding the use of augmentation therapy in alpha-1 antitrypsin deficiency was the 1981 publication with the first author Jim Gadek. Basically, at the NIH, they infused five subjects with weekly infusions of purified alpha-1 antitrypsin. Three of the five subjects underwent bronchoscopy to measure epithelial lining fluid levels as

well. They gave a fixed dose of 4 grams of alpha-1 antitrypsin protein per week intravenously. They decided that the protective level they were trying to achieve was 70 mg/dL.

This allows me to introduce the concept that you will hear a number of different scales for the measurement of alpha-1 antitrypsin. Many clinical laboratories use and continue to use the milligrams per deciliter in the United States, grams per liter are used in Canada and in Europe, and most of the reference laboratories here in the U.S. use micromoles or micromoles per liter.

So they decided that the protective level was 70 mg/dL, based on the observation that the patients that had reported to the NIH with a PiSS phenotype never seemed to get emphysema and no PiSS patient that they had identified had a level that was less than 70 mg/dL. Now, that was a small number of patients, but at least it provided some rationale for attempting to achieve a protective level.

All five subjects reached this trough level target on the 4 grams per week. All that were tested had an increase in their epithelial lining fluid levels.

Interestingly, and without written explanation, subsequent studies from the same group changed their protective level to 80 mg/dL, or 15.38 µM, based on the molecular weight of alpha-1 antitrypsin. They started

basing their dosing on weight. Instead of giving a fixed dose of 4 grams, they gave 60 mg/kg. My presumption is that they found some patients who didn't reach the projected trough level on the fixed dosing of 4 grams per dose and that they felt it was necessary to change to a dosing based on weight.

Prolastin was approved in 1987 based on maintaining trough levels above 80 mg/dL, after the studies done at the NIH. In 1989, when the NIH registry was getting off the ground, the one of the inclusion criteria for enrollment in the registry was that individuals considered for enrollment had to have an alpha-1 antitrypsin level below 11 µM. This was decided upon because an evaluation of PiZZ subjects showed that there were no individuals whose level at any measurement was greater than 11 µM. So it was a way of ensuring that patients enrolled in the registry had severe deficiency of alpha-1 antitrypsin.

New IV therapies began development in the 1990s. All of the studies leading to approval of the drugs that are currently on the market, after Prolastin, used a target trough level of 11 µM, or 57 mg/dL. There has been a post-hoc rationalization for how that change occurred. It's basically that the purity of alpha-1 antitrypsin standards was so poor in the 1980s that 80 mg/dL in the 1980s

represented 57 mg/dL in the 1990s, and that these were actually identical targets. In fact, this may actually be true. There has been a gradual improvement in the quality of the standards, and that has been reflected in clinical laboratories with a dropping of the normal range that has been published for normal laboratory determination of alpha-1 antitrypsin deficiency.

To switch over to studies that use these target levels, first, of course, was the approval of the three drugs currently marketed in the U.S., Prolastin, Aralast, and Zemaira. While the studies on these three drugs took place over the course of a couple of decades, their design was essentially the same:

- A small number of subjects, always fewer that 50, were given weekly infusions of short duration -- about six months -- at 60 mg/kg.
- Blood levels of alpha-1 antitrypsin were measured in all subjects.
- ELF levels of alpha-1 antitrypsin were measured usually in a subset of subjects.
- Various other analyses, such as alpha-1 elastase complex, this free elastase activity, were also measured.
- In the case of Aralast and Zemaira, a subgroup of subjects received Prolastin for about half the study

period, in a blinded fashion, as a comparator, to show either non-inferiority or equivalence.

- The primary endpoint in all cases was achieving a blood level of antigenic alpha-1 antitrypsin of 11 µM or 80 mg/dL in the case of Prolastin.

You have already heard a lot about the NIH registry:

- The enrollment in the registry was 1,129 subjects. Many of the scientists here in the room were participants in this registry.
- It was not designed as an efficacy trial of augmentation therapy. It was a natural-history study.
- The study compared, as you saw in the table from Dr. Chapman, individuals enrolled on therapy or started on therapy after enrollment -- and that number was 368 patients -- and they were compared with case controls never on therapy throughout the enrollment in the registry.
- They found that mortality and FEV_1 decline improvements were noted in the treated group compared to that group that never had any therapy. The FEV_1 decline improvement, as was just shown, was only statistically significant in those with moderate obstruction at baseline, the 35 to 49 percent of predicted.

Also mentioned -- and I'll go through this quickly -- was the German-Danish study, where 198 German

patients on augmentation therapy were compared with 97 Danish patients not on augmentation therapy. Overall, patients on augmentation therapy had a significant reduction in FEV_1 decline. This difference was almost entirely accounted for by those with baseline FEV_1 in the 31 to 65 percent of predicted range.

In the study by Dr. Wenker that was mentioned -- and Dr. Wenker is in the audience right now -- patients had longitudinal spirometry data before starting augmentation therapy and then continued their spirometry following augmentation therapy. The rate of decline of FEV_1 was compared before and after therapy. As mentioned, the largest effect was seen in individuals with rapidly declining FEV_1 prior to the start of therapy. These were virtually all in the group with an FEV_1 at baseline that was better than 65 percent of predicted.

Finally, to do the absurd, which is to summarize the data of a presenter who just spoke, the Canadian retrospective analysis in 2009 looked at 21 patients in the Canadian international registry receiving augmentation therapy and compared those with 42 matched controls not receiving augmentation therapy. As the graphic shows, the group on augmentation therapy had a rate of decline of FEV_1, in terms of mL per year, of about 30 mL per year, compared to the 64-mL-per-year decline that those not on therapy

had.

Now I'm going to relate this to the story of PiSZ individuals. PiSZ individuals tend to have higher alpha-1 antitrypsin levels than those that are PiZZ, who have the two Z genes. But they have lower alpha-1 levels than those with the PiMZ genotype. Their AAT levels tend to straddle the 11 μM magical threshold. Their alpha-1 antitrypsin levels tend to be more labile than those that have the PiZZ genotype, in that on any given day, their levels can change by a larger percentage, especially in response to inflammation, infection -- anything that causes rises in acute-phase reactants.

Studies have demonstrated less lung disease in those with PiSZ genotype compared to ZZ individuals. Most interesting, often lung disease did not directly correlate with the alpha-1 antitrypsin level measured in that individual.

I'm going to review some of the literature that those general statements are based on.

Jerry Turino and many of the investigators in the NIH registry looked at the PiSZ subjects who had been evaluated and accepted into the registry and those PiSZ subjects who were evaluated and did not meet criteria for enrollment in the registry because their alpha-1 antitrypsin levels were too high. They found, I think, 58

or 59 subjects, of which 50 had complete enough data to analyze. In that group, 10 of the patients had alpha-1 antitrypsin levels at baseline that were less than 11 µM and 40 had alpha-1 antitrypsin levels that were greater than 11 µM.

They compared these groups with PiZZ subjects. The PiSZ subjects, in comparison with the PiZZ subjects, had less emphysema on their chest x-ray -- there were no CTs done as part of the NHLBI registry -- and they had better lung function, judged by spirometry. Most PiSZ subjects who did not smoke had normal lung function and a normal chest x-ray. This was distinct from the PiZZ subjects, where many smokers enrolled in the registry had abnormal lung function and/or abnormal chest x-rays. They concluded that PiSZ nonsmokers were at little risk of disease.

But they also made an interesting observation. The alpha-1 antitrypsin levels, when comparing the group with levels greater than 11 µM with those with less than 11 µM -- there was not a correlation between the alpha-1 antitrypsin level and the disease risk. In fact, there were more patients with lung disease in the group that had levels greater than 11 µM compared to the those with levels less than 11 µM.

I was also privileged to be able to present to

you a paper that has not yet been published, although it has been accepted for publication, Rob Stockley's group, with Jayne Holme as the primary author. That study, from the U.K., looked at patients with SZ in the U.K. registry for alpha-1 antitrypsin. They found that patients with SZ had less emphysema and more upper-zone distribution of emphysema than those with PiZZ. PiZZ subjects, in general, are thought to have primarily lower-zone distribution emphysema, although there are some data suggesting that emphysema might start in the upper zones in early emphysema and then become more severe in the lower zones, through work that was actually done in the U.K. group.

Both groups had emphysema that was panacinar, not the usual type of central lobular emphysema that you see with typical smoking-related emphysema.

Despite these differences, there was no difference in the health status of the PiZZ and PiSZ subject groups. They were each equally affected in their daily lives, quality of life, and functional capacities.

The PiSZ individuals with levels of alpha-1 antitrypsin greater than 11 μM had more breathlessness, as judged by the NRC dyspnea score, but a better short form 36 physical summary score, than those with less than 11 μM. PiSZ individuals with alpha-1 antitrypsin levels above and below 11 μM were not different in their CT densitometry, CT

scan appearance, C-reactive protein levels, as a measure of acute-phase reactants, or any other variables that they looked at.

So what are the problems with drawing major conclusions from these studies?

The very first one, of course, is that the N, the number of patients evaluated, even in these studies as a whole, is very small, especially when you consider that the prevalence of the S gene should be significantly higher than the prevalence of the Z gene. Very few individuals with SZ are identified, presumably because they don't have clinical symptoms that warrant testing for alpha-1 antitrypsin deficiency.

All these studies compared PiSZ subjects with PiZZ subjects. There is no data making comparison with MM subjects, normal individuals with alpha-1, although it would be clearly expected that these patients would do worse clinically than MM subjects who had a similar smoking history.

Because of lability in the AAT levels in PiSZ, a single level might not be representative of what the integrated level of someone is over time in PiSZ.

Alpha-1 antitrypsin levels greater than 11 μM in SZ may represent ongoing inflammatory stress in the sicker patients. Since alpha-1 is an acute-phase reactant, there

is the possibility that what we are seeing is a reaction to the fact that these patients are sick, and manifesting that as an elevated SZ level above baseline. However, in the Stockley study that I mentioned, when CRP levels were looked at, since the CRP levels were not different between the SZ subjects with low alpha-1 levels compared to higher-level alpha-1s -- if CRP directly correlates with other acute-phase reactants, that would argue against this as an explanation.

We have ignored the possibility that the PiS protein has biochemical attributes that are distinct from the PiZ patients. Perhaps there are some deleterious effects of having a slightly higher S protein level that we are not yet appreciating.

In summary:

- There is some question whether alpha-1 antitrypsin levels directly correlate with disease risk in all subjects.
- We don’t know what the correct protective level is for a given patient.
- We don’t know if current weight-based dosing is appropriate, especially at weight extremes, very small individuals or very heavy individuals.
- Therefore, we don’t really know what the appropriate dose is to a given patient. That is manifest

clinically, to those of us who treat a lot of alpha-1 patients, in the setting that there are some patients on appropriate dose of alpha-1 antitrypsin augmentation therapy whose lung function seems to continue to decline at a relatively rapid rate. There are those of us in the clinic who, in spite of the package insert, raise the dosing of drug that we give to patients in that kind of a setting.

- This inability to know what the appropriate dose is for a given patient has led to a variety of unapproved dosing regimens and to dosing based on repeated trough level measurements and adjustment of dose based on trough levels, or changing dosing based on clinical response.

This leads me to suggest that dose-finding studies should be done in alpha-1 antitrypsin deficiency on any new drug moving forward in clinical trials to treat alpha-1. In order to do these, however, we must have pragmatically achievable clinical efficacy endpoints. Clearly, the gold standard from the past has been FEV_1. But other measures of lung function, such as well-done diffusing capacity measurements or Kco measurements, and various ratios of Kco to other values, might be a possibility. CT lung densitometry is one that you will be discussing in detail today. Clinical endpoints like

exacerbation and mortality maybe should be looked at a little more closely by companies moving forward in alpha-1 antitrypsin deficiency.

A discussion can reasonably be made regarding the use of surrogates for dose-ranging studies, including:

- Various measures of elastin breakdown. Many of those elastin-breakdown assays have been called into question in terms of their utility, but there are still some new techniques that are still moving forward.
- The possible surrogate endpoint of the elimination of free elastase activity on bronchoalveolar lavage or the measurement of alpha-1 antitrypsin elastase complexes and various other inflammatory mediators and markers that have been shown to be affected by the administration of augmentation therapy to alpha-1 antitrypsin-deficient patients.

That’s my take on dosing as we currently know it in alpha-1 antitrypsin deficiency.

DR. SIEGEL: Thank you, Dr. Sandhaus.

We will next hear a discussion of “Inhalation Therapy for Emphysema Due to AAT Deficiency,” with Mark Brantly, from the University of Florida.

Agenda Item: Inhalation Therapy for Emphysema Due to AAT Deficiency

DR. BRANTLY: I would like to thank the agency

for inviting me.

Committee members, I'm tasked with providing the scientific basis for inhalation therapy for the treatment of emphysema in alpha-1 antitrypsin-deficient individuals.

As mentioned, I have some conflicts, which are listed above.

Let me give you some background from a biochemist, as far as alpha-1 antitrypsin. I think there has been an accumulation of data over the last 10 years regarding a broader concept of alpha-1 antitrypsin as a therapeutic molecule. There is accumulating evidence that it modulates inflammation in general.

It's a 52-kilodalton glycoprotein. Concentrations typically are between 20 and 53, but oftentimes the alpha-1 antitrypsin levels can double with acute-phase response. The normal concentration in the lung is approximately 2 to 5 μM.

It has many functions, including as a broad-spectrum antiprotease. It inhibits neutrophil defensins, cytotoxicity and proinflammatory factors. It's an antioxidant. It blocks LPS-mediated inflammation. it reduces CRP concentrations in alpha antitrypsin-deficient individuals. Indeed, we have done a study of approximately 500 individuals that are ZZ and demonstrated that augmentation therapy actually reduces the CRP by about 50

percent.

In thinking about aerosolized alpha antitrypsin, one of the background things you have to think about is what's going on in the lung in an alpha antitrypsin-deficient individual. This is one of the ways I like to conceptualize alpha antitrypsin deficiency in lung. We have an initiation phase, most predominantly caused by smoking, but indeed infections. There is now early data that suggests that polymers of alpha antitrypsin may -- and there is now some early data that suggests that the cellular response to unfolded Z alpha antitrypsin may also play a role in causing inflammation.

These triggers basically activate the inflammatory system within the lung, involving virtually all the inflammatory cells at one point or another. These targets include the airway, the alveolus, and, obviously, the interstitium. There is a dynamic interaction between these proinflammatory cells and proinflammatory factors, which are important in basically ramping up and expanding inflammation.

The effects of injury from these cells include oxidant proteases and alpha defensins. As I mentioned, alpha antitrypsin specifically inhibits many of these molecules.

What about the potential of aerosolized alpha

antitrypsin? This is a cartoon of the size of the aerosol we need to get down, which is about 2 microns in size.

Why should we be pursuing aerosolized alpha antitrypsin?

- Number one, IV augmentation therapy is not the magic bullet, and it requires IV access once a week.
- Most of the IV alpha antitrypsin does not reach the lung, and IV alpha antitrypsin that does reach the lung is in the low-normal range in the epithelial lining fluid.
- Somebody that has inflammation may require, actually, larger amounts.

What are some of the potential advantages of aerosolized alpha antitrypsin?

- Ease of use is one, for sure.
- Number two is direct delivery to the airway and the lower respiratory tract.
- Number three is the potential to deliver high doses, pharmacologic doses, into the lung itself.
- The possibility of using aerosolized alpha antitrypsin in other inflammatory diseases besides alpha antitrypsin deficiency.

What are the components necessary to have a successful clinical program for alpha antitrypsin?

- We need highly purified alpha antitrypsin. I'll show you some data as to why that's important. We

have some of these on the market now -- Aralast NP, Zemaira, Prolastin NP, and a Comet API(?), and a RIVA(?) recombinant alpha antitrypsin.

- We often need high-efficiency deep-lung delivery devices as well.
- We need to determine the safe and effective dose, appropriate dose.
- We need robust outcome variables appropriate for rare-disease studies. Let me emphasize again, I think this committee is very attuned to rare diseases. We can't have the same study designs as we have for large disease populations.
- We need robust surrogate markers for the Phase 1 and 2.
- We need to demonstrate that aerosolized alpha antitrypsin reaches the interstitial space.

Let's talk just for a minute about aerosol characteristics. This is from a paper from Brand in the *European Respiratory Journal* from 2003. The upper panel right there is the total deposition, using various types of nebulizers. You can see that there are now possibilities of getting nebulization in high amounts -- 60, 70 percent total deposition. You can see here a decline over time with the various different nebulizers. What's an important thing is, obviously, the decline is associated with how

much you initially deliver.

The amount of peripheral deposition is located right here. Basically, you can see that you get more and more peripheral deposition as you have a more efficient device.

Finally, the total amount of material -- the time is also an important function. We know from our aerosol studies that patients tend to not tolerate aerosolizations more than about 20 minutes. They become less comfortable over that time. So we have to have an efficient nebulizer that will deliver the drug to the lower respiratory tract in a relatively short time.

This is a more recent paper from the same group, showing the amount of drug dose in the various types of compartments. Interestingly enough, when you have efficient devices -- and this is the AKITA apex device -- the compartments are nearly the same as healthy individuals as compared to alpha-1 antitrypsin-deficient individuals.

Here you can see also that, remarkably -- an important characteristic that is necessary is, we have to have uniform deposition over a large FEV_1 as well. These are the challenges that have already been overcome by different devices.

One of the other key things that we have to be able to do is -- when you do aerosolized alpha antitrypsin,

you are delivering it in an unphysiologic manner. Normally alpha antitrypsin is secreted by the liver. It goes into the vascular system and leaks into the lung. In this approach, we are basically delivering the alpha-1 antitrypsin to the epithelial side. The question is, does it get into the interstitial?

This is a study from Rick Hubbard (phonetic) in 1989, in sheep, where he took recombinant alpha-1 antitrypsin. You can see here that the aerosolized alpha-1 antitrypsin actually appears in the lymph and also in the blood over time.

I'll later show you some animal studies that demonstrate similar things.

I would like to review with you the background of some of the aerosol studies that have been done in preparation for aerosolized alpha-1 antitrypsin. We'll talk about a study demonstrating airway inflammation and lung inflammation in alpha-1 antitrypsin-deficient individuals. We'll talk about a Phase 1/Phase 2 study using recombinant sheep alpha-1 antitrypsin, the PPL study. We'll talk about a Phase 1 study using the dry powder form, which is highly efficient. Finally, we'll talk briefly about a Phase 1 study using recombinant yeast alpha-1 antitrypsin.

As some background, this is a study evaluating

individuals with alpha-1 antitrypsin deficiency compared to normals. You can see that the alpha-1 antitrypsin levels are five times lower than the normal group of individuals. This is also reflected in their lung. There is an approximately tenfold difference.

If you look at these individuals -- and I just want to point out that this group of individuals had FEV_1s in the near-normal range -- their neutrophils are already starting to accumulate as compared to normal individuals. So even in the early stage of lung disease in alpha-1 antitrypsin-deficient individuals, they are developing an increased number of neutrophils in their lungs.

If you look at the cellular features in these individuals, you can see here, when you compare 14 normals to 22 alpha-1 antitrypsin-deficient individuals, the returns on the amount of fluid that comes back from bronchoalveolar lavage is very similar. The total number of cells is relatively similar as well. The alveolar macrophage content is approximately the same as well. You will see that there is a significant difference in the number of neutrophils. Lymphocytes and ciliated cells are relatively the same as well.

If you look at other inflammatory factors, besides the cells, you can see that here, what we call the ELF volume -- and for those who are non-aficionados, “ELF”

stands for “epithelial lining fluid” volume -- we basically use a methodology to correct from BAL by using urea as a dilution factor and then we calculate back what the actual concentrations will be in the lower respiratory tract, in the epithelial lining fluid.

As you can see, some of the dramatic things -- again, these are individuals who have very mild lung-function abnormalities -- their neutrophil elastase is approximately 30 times higher. Their alpha-1 antitrypsin levels are approximately 10 times lower than in a normal individual. The complexes are, interestingly enough, exactly the same. The reason why they are the same is, in deficient individuals, the limiting reagent is alpha-1 antitrypsin and in the normal individuals the limiting reagent is neutrophil elastase. Alpha defensins are 42-fold higher in alpha-1 antitrypsin-deficient individuals. You can see that some of these classic proinflammatory factors, such as IL-1beta, IL-6, IL-8, and LTB-4, are also substantially increased in this group of individuals.

But it’s important to understand that some of these surrogate markers, particularly in the individuals that have mild lung disease, correlate very well with some of the clinical factors that we know about as well. Here is an example of neutrophil elastase correlating with IL-8. If you look at neutrophil percentage, you can see that it

correlates with the FEV_1, as well as the rate of decline in lung function and the DLCO. The same is true also of neutrophil defensins in this group, where you can see that neutrophil defensins correlate with some of the other neutrophil products and the percent of neutrophils, as well as the FEV_1 and DLCO and the rate of decline.

So these surrogate markers correlate, at least in patients that have mild disease.

While we think of alpha-1 antitrypsin-deficient individuals as having predominantly emphysema, they actually also have airway disease. I know this for a fact, because I visit their lungs on a regular basis. This is a picture of an individual that was a nonsmoker with alpha-1 antitrypsin deficiency, that has areas of bronchiectasis, which is a component of alpha-1 antitrypsin. You have significant airway damage.

What is some of the evidence that they do? On CT scan, approximately 90 percent of individuals have some areas of focal bronchiectasis. Sixty-five percent of all the alpha-1 antitrypsin-deficient individuals have a positive bronchodilator response at one point during our study. Animal models of airway hyperactivity are reduced by alpha-1 antitrypsin. There is anecdotal evidence from patients, in the form of studies by Jack Lieberman, which indicate that augmentation therapy actually helps control

their asthmatic component. Finally, there is a large fraction of individuals that present predominantly with asthma symptoms as their major factor.

When you biopsy these individuals, they oftentimes will have classic damage. They have a thickened basement membrane. They can have airway metaplasia. They have goblet cell hyperplasia, as well as airway fibrosis, and importantly, have an increase in the number of inflammatory cells that they have.

This is a summary of one of the studies we have done in this mild group of individuals. You can see that the total inflammatory cells under the basement membrane per millimeter-squared is not quite half as much. Interestingly enough, there is an increase in the number of CD4 and CD8 cells, as well as an increase in mast cells. Interestingly enough, neutrophils, at least in these mild individuals, were not significantly positive in these individuals. So they clearly have airway inflammation, as well as lower respiratory tract inflammation.

I hope I have convinced you that these patients, even early on, have airway inflammation. So let’s talk about one of the first studies using aerosolized alpha-1 antitrypsin.

This is a study using aerosolized recombinant alpha-1 antitrypsin. This is a transgenic form from sheep,

Dolly. It was a 250-mg dose once a day using a PARI LC Star nebulizer. This is the alpha-1 antitrypsin concentration. This extends over an eight-week period, where the patients are given a daily dose.

You can see here that the alpha-1 antitrypsin levels went up to approximately the normal range, which is this green bar right here. It was a little bit lower at eight weeks, indicating, number one, that alpha-1 antitrypsin didn't accumulate over time. It was eliminated quickly. The antineutrophil elastase capacity -- that is, the ability to raise the defenses against it -- was increased in all these individuals as well.

Importantly, one of the things that I would like to point out is that the alpha-1 antitrypsin neutrophil elastase complexes in these individuals that had free neutrophil elastase went up substantially over time. Interestingly enough, and importantly, when you give inhaled alpha-1 antitrypsin to this group of individuals, you see a drop in the number of neutrophils in the lungs of these individuals.

Proinflammatory factors may trend toward the lower areas, but they did not reach statistical significance.

So what are the big questions? We have shown that alpha-1 antitrypsin passes into the lymph of animals

and also into the bloodstream. But in humans we have not been able to do that, since humans tend to not like to have their lymphatics cannulated.

Normally -- again reminding you -- we give it intravenously. It goes from the endothelial side to the epithelial side. Here we are giving it in a reverse manner. We are taking to the epithelial side. What is that concentration that we really need?

We were able to create a monoclonal antibody that recognizes M alpha-1 antitrypsin, but not Z alpha-1 antitrypsin. Looking at the plasma of the previous study individuals, here is the baseline, which is sort of nonspecific stuff. You can see that there is the appearance of normal alpha-1 antitrypsin in the blood, indicating that the aerosolized alpha-1 antitrypsin reached the plasma by crossing through the interstitium in this particular case.

Here is a study using the dry powder form of alpha-1 antitrypsin, Zemaira. It was a Phase 1 dose-escalation study. I'm just taking a few of the doses and putting them here. The dosage was once a day for two weeks. Bronchoalveolar lavage was done before and then following two weeks. There are no SAEs in this particular study. The top dose extended the normal alpha-1 antitrypsin level substantially. Here is the normal alpha-

1 antitrypsin level right here. You can see where we are seeing 6- and 8-μM concentrations and a nice dose response associated with this. This nebulizer is predicted to be about 75 percent efficient.

We have done similar things using yeast alpha-1 antitrypsin, where we use a nebulizer, an Aero-Eclipse, which is not nearly as efficient. It's a breath-actuated device, giving 200 mg/day. You can see, however, that there are some challenges. You can see that the alpha-1 antitrypsin concentrations vary by lobe. The lung is a multi-compartmented area, and it has both variation in the inflammatory factors in each of the lobes and the amount of alpha-1 antitrypsin that gets into these lobes.

This is the dose-escalation portion of that particular study as well, one looking at 100 mg once a day, one looking at 100 mg twice a day, and 200 mg. You can see that there is actually no difference between the 200-mg doses. So we have a good idea from this particular study what kinds of dosing ranges we might need.

In summary, inflammation is present in alpha-1 antitrypsin-deficient individuals. Alpha-1 antitrypsin is a natural anti-inflammatory molecule, based on a great deal of work. Alpha-1 antitrypsin decreases inflammation in the lower respiratory tract. Presumably, because it decreases inflammation, it should be associated with less injury to

the lung.

We have excellent deep delivery devices, with short nebulization times. Aerosolized alpha-1 antitrypsin clearly crosses into the interstitial space. Once-a-day dosing is possible. BAL outcome variables for dose finding and anti-inflammatory effect definitely are feasible.

Long aerosolized -- we have some history with other proteins that have been inhaled, including dornase alpha, with the cystic fibrosis group, which is out to 96 weeks in studies, and obviously for much longer. There is an important point to make regarding that. The amount of dornase alpha that we typically give is about .25 mg. That's between five and 10 times less than we would expect to deliver to the lungs of alpha-1 antitrypsin-deficient individuals.

It is unlikely that aerosolized alpha-1 antitrypsin will alter the lung density. The lung at any given time has approximately 30 percent of the blood volume. Typically, when we are looking at aerosolizing -- at least the liquid types -- we are talking about putting 4 mL into the lung at the time.

Thank you very much.

DR. SIEGEL: Thank you, Dr. Brantly.

Next we will hear from the FDA, from L. Ross Pierce, M.D., from OBRR, "Trial Design Considerations for

Clinically Meaningful Endpoint Trials in AAT Deficiency."

Agenda Item: Trial Design Considerations for Clinically Meaningful Endpoint Trials in AAT Deficiency

DR. PIERCE: Thank you very much.

The title of this presentation is "Clinical and Surrogate Endpoints for Evaluating Efficacy of Alpha-1 Proteinase Inhibitor (Human) Augmentation Therapy."

The outline of the presentation today:

- Covering the history and basis of using the biochemical surrogates for evaluation of intravenous alpha-1 PI augmentation therapy products. Some of this has been gone into by Dr. Sandhaus.
- The design and results of biochemical surrogate endpoint studies of alpha-1 antitrypsin products, intravenous, in a bit more detail than what was presented this morning, will be gone into.
- Finally, design options for trials using clinically meaningful endpoints.

We actually have four of them now. The licensed alpha-1 proteinase inhibitor products, all intravenous, are dosed to augment but not normalize the serum and lung levels of alpha-1 antitrypsin in severely alpha-1 antitrypsin-deficient patients who have emphysema. Achieving the historical target serum threshold of 11 μM alpha-1 antitrypsin was hypothesized to restore the

protease-antiprotease balance in the lung and prevent accelerated lung elastolysis. I want to highlight that term: "protease-antiprotease balance."

As mentioned earlier, in 1985, there was a joint workshop with NHLBI where participants recommended the use of both serum and lung epithelial lining fluid levels to evaluate the efficacy of these products.

Specifically, they recommended that the products should achieve serum antigenic and functional AAT levels in the serum, in the range of PiMZ heterozygotes. We believe these levels to be in the range of about 17 to 33 µM.

The workshop participants also recommended that appropriate increases from baseline be seen in antigenic and functional AAT levels in the ELF.

Bayer used for Prolastin, and FDA accepted for licensure, a somewhat lower target trough level of 11 µM for both the antigenic and functional alpha-1 antitrypsin levels. This had been proposed in the literature, as was described by Dr. Sandhaus. This target was based on the differential risk of emphysema among various AAT phenotypes and their associated typical AAT concentrations. But at the time that this 11-µM threshold was developed, there was actually limited data as to the range of serum AAT levels for various of the key phenotypes, including SZ and SS patients. We now understand that the range of the levels

that those patients have is somewhat broader, particularly to the up-side in the case of SZs.

When we talk about a therapeutic threshold and we talk about a therapeutic product, we have to recognize that the levels are not static. When we give a product once a week, we have a pharmacokinetic curve. In this trial, day 1 is what we normally consider day 0. It's the day of administration of the product. This represents -- taken from the literature -- for Prolastin a PK curve for serum antigenic alpha-1 antitrypsin levels after four weekly doses of the labeled recommended dose of 60 mg/kg/week. Four weekly doses get us close to steady-state values. So this would be a steady-state PK curve.

The peak level is in the area of 50 or so µM, which is at the top end of the normal range. Immediately after giving the infusion, you have levels which are around the top of the normal range. But then there is a rapid fall over the first two days, and you are down to levels a little over 20, which is around the region of the upper range for SZ patients. Then it falls progressively.

This just represents the mean values. I haven't plotted the range of the standard errors. There is variability.

The basis of the 11 µM, again, reflected an observation that there appeared to be a stepwise

progression of emphysema risk, looking at the different AAT phenotypes. But the available epidemiologic data do not identify an emphysema risk gradient as a function of the serum AAT level within any one AAT phenotype, as was described, particularly for the SZ patients, by Dr. Sandhaus.

This table represents a subset of the phenotypes you saw displayed in Dr. Chapman's and Dr. Sandhaus' presentations earlier. You can see that the PiSZ patients straddle the level of 11 µM. Not depicted on this particular slide are the SS subjects.

I would point out that the original formulation of 11 µM as a therapeutic threshold sought to make a separation between, as a group, the null-nulls, the ZZs, and the SZ patients, who were considered to be at some risk -- albeit the SZs at less risk than the ZZs. They wanted to establish a level that would separate them from the SS patients. But even as recently as 2005, a meta-analysis by Dahl concluded that there were an insufficient number of studies or data with the SS subjects to evaluate what their emphysema risk actually was.

Dahl also in that paper did a meta-analysis of SZ risk and concluded that it was approximately 3.3 times that of the normal MM individuals, but with a wide confidence interval range, although the result was statistically

significant in terms of greater risk than normals.

The MZ patients, whose levels can range from about 17 to 33 µM, also are concluded to be at increased risk as a group for emphysema by the European Respiratory Society and American Thoracic Society statement on alpha-1 antitrypsin deficiency. In one study where the rate of hospitalization in a Scandinavian was compared among MZs and MM patients, the risk of hospitalization for COPD was 2.2 times increased in the MZ population.

So you can see, based on these levels and the fact that MZ is associated with a low absolute but nevertheless significant increase in the risk of COPD, it means that if you achieve a level of 11 µM as a trough level, you have not necessarily protected the individual throughout the entire seven-day period that dosing occurs.

If we go back for a moment to this level here, you can see that if it turns out that the optimal protective level is on the order of 40 µM, for example, you would be protected for maybe a day and a half out of that seven-day inter-dosing interval. But this also would depend on how quickly new neutrophils are coming into the lung to replace those that have been neutralized.

If, on the other hand, the protective level were around 20, then you can see that the duration of protection out of the seven-day inter-dosing interval would be

greater.

The conclusion that we have reached is that the minimum and optimal therapeutic targets for alpha-1 proteinase inhibitor cannot be discerned directly from the literature and that the postulated therapeutic target levels either of 11 μM or the 17 to 33 μM, proposed, respectively, by Gadek and Crystal and the 1985 NHLBI workshop participants, is not extremely well-grounded by epidemiology in terms of identifying a precise optimal therapeutic target.

In addition, we note that the severely AAT-deficient individuals have greater than normal lung neutrophils and neutrophil elastase burden. Again, when you compare these different phenotypes and you attempt to, not go to normal levels, but go to the level of MZ heterozygotes, and you recognize that the original attempt with augmentation therapy was to basically convert somebody to normal -- to stop the progression of emphysema dead in its tracks and convert somebody into a normal individual who has some small rate of FEV_1 decline per year -- this degree of effect has not actually been seen in the epidemiologic studies and clinical trials to date, as we will go into.

So if you have an increased level of neutrophils and neutrophil elastase in the severely deficient patients,

then the whole premise on which to identify a therapeutic target based on a comparison of the different phenotypes comes into question. In order to respect the protease-antiprotease balance theory, you have to know how much proteinase is in the lung or neutrophil elastase that you are talking about.

Data show that some SZ individuals with serum AAT levels greater than 11 µM do, nevertheless, have emphysema.

The studies that have suggested that alpha-1 PI may be what I would term partially effective in slowing emphysema progression are listed here. These are exactly the same studies that were presented by Dr. Chapman in the meta-analysis:

- The largest being the NHLBI registry study.
- The two-country study published by Seersholm, and also in another publication by Wenker, that compared the rate of decline of patients in Germany who were augmented to the rate of decline in another cohort of patients in another country, Denmark, who were not augmented. Of course, any time you do that type of cross-country comparison, you need to assure yourself that the underlying medical care in the two countries is quite comparable.
- The study by Wenker, which was a mixed retrospective/prospective epidemiologic study. I should

mention, by the way, that in the discussion section of Dr. Chapman's meta-analysis, they mention some caveats with respect to interpretation of that meta-analysis, including that some of the individuals in the augmented group in the Wenker study were actually also in the two-country Seersholm study. Those two studies comprise 45 percent of the weight of the Chapman meta-analysis. It's not clear from my reading of that paper whether the double-counting of patients has been taken into account in the weighting in that meta-analysis.

- In addition, the Dirksen clinical trial, the only clinical trial out of the group of studies in the meta-analysis by Chapman and colleagues. It's also interesting to note that in that study, the primary endpoint was, in fact, FEV_1, but it was a different measure of FEV_1. It included both patient self-assessment and the laboratory assessment. So the primary endpoint of that study was different from the FEV_1 that was used for the Chapman meta-analysis.
- Then we have most recently a study, again by Dirksen and colleagues, published just this year, which will be described by our last speaker today.

These studies, while suggestive of benefit, do not identify any particular serum AAT level as partially or fully protective from further emphysema progression. In

fact, the largest of these studies, the NHLBI registry, did not measure the AAT levels in the patients during the course of the trial as part of the study. They studied those levels only at entry.

In 1998, after the NHLBI registry study was published, the FDA Blood Products Advisory Committee recommended that the same original criteria of serum and ELF AAT levels as had been used for the licensure of Prolastin be used for the approval of subsequent products. As I mentioned, we now have four products. The Aralast exists in two forms, Aralast with the substantial proportion of molecules in the vial having a single amino acid truncation, as well as Aralast NP, which is 98 percent the native sequence. All four of these products were based on the combined surrogate endpoints of serum and lung ELF levels, both antigenic levels and functional levels.

In addition to the requirement that the serum antigenic trough level be greater than 11 µM, we also, in all of the pivotal trials for the newer products, required that there be a non-inferiority comparison as part of a co-primary endpoint, so that the antigenic level in the blood was not inferior to Prolastin. The non-inferiority margin was on the order of 20 percent, or 3 µM.

We also looked to see a statistically significant rise from baseline in the ELF of the various analytes.

The pivotal studies actually included anywhere from 14 to 30 subjects per treatment arm. In the Prolastin study, subjects served as their own controls. As was mentioned, the newer products were studied in a parallel dosing fashion, randomized and blinded against Prolastin for the first 10 weeks or so. Then the patients who were originally randomized to Prolastin crossed over to the new test product.

Here you see the actual levels that were achieved in the various trials. In gray you see the results for Prolastin in a single-arm study. These values are all antigenic levels.

I should mention that the functional levels are going to be less than this -- in the case of Prolastin, often 90 percent less, or slightly over 90 percent. But because each lot can vary in the ratio of active to inactive product and the specific activity, the values can be lower than this for the antineutrophil elastase or functional measure of alpha-1 PI.

In any one Prolastin vial, you have about 20 percent of the protein that is other proteins and, of that, 100 percent of alpha-1 PI molecules -- you can have anywhere from about 60-some percent up to 90 percent being active Prolastin that you will be able to measure in the A and the C assay.

In the case of the reddish-brown analytes, you see the result for Zemaira, being nearly 18 for a mean value, a range from 14 up to 23 for antigenic AAT. In the parallel Prolastin arm, the values were nominally greater than that, 19, with values ranging from 14 to 28.

You can see in the Aralast study, depicted in blue, that the mean values for Aralast were somewhat lower than for Prolastin, but still within the predefined PK equivalence limits -- a mean value of a little over 15 and, in the case of Prolastin, a value of 17.

Here we see the results of epithelial lining fluid. Again, we are looking only at antigenic levels here. If you talk about the functional levels in the epithelial lining fluid, one of the papers published in the original Prolastin registration study indicated that six days after giving the product in repeated weekly dose, you achieved a level in the epithelial lining fluid of active alpha-1 PI -- that is to say, the ANEC -- that was about 57 percent of that in normal individuals, again, emphasizing that this is not full replacement therapy. This is something less than that, even though the neutrophil elastase burden in these patients who have the severe deficiency we know to be higher than normal, and also higher than people with established COPD, matched for the severity of COPD, but without the alpha-1 antitrypsin

deficiency gene.

Here we are talking about sub-studies. The numbers are very small. Nevertheless, we are seeing values in the same range in AAT mean level change from baseline for these various products, and substantially above the baseline values.

Key points for the licensed alpha-1 PI products:

- The recommended dose is not based on dose-ranging studies and may not be optimal.
- Neither epidemiologic studies nor randomized, controlled trials provide an adequate basis for using the historical serum concentration target of 11 μM when dosing and evaluating alpha-1 PI products. And if you do happen to measure the alpha-1 PI concentration in a commercial lab, you are measuring only the total product, which is a mixture of active and functionally inactive product.
- The available data, while not conclusive, suggest partial efficacy, but the minimum and optimal therapeutic serum levels remain unknown.
- More recent epidemiologic data suggest that the historical target of 11 may be too low to discriminate those at risk from those not at risk of COPD emphysema in terms of our examination of SZ patient data.
- Severely A_1 PI-deficient patients have increased neutrophil and neutrophil elastase burden. In

the case of Dr. Brantly's study, he showed us that neutrophil elastase was 30 times higher in concentration in the lungs of those ZZ patients with mild airway obstruction as compared to normal individuals. So even a revised target of, perhaps, 20 µM may not necessarily be optimal.

Several years ago, we contacted the manufacturers of the licensed alpha-1 proteinase inhibitor products and requested that they commit to conducting postmarketing studies that would use clinically meaningful endpoints in order to verify the efficacy of the products. They committed to either do clinical trials which would look at serial lung-density changes by CT, pulmonary exacerbations of COPD, or serial pulmonary function testing, or a combination of these. We would also accept mortality as an endpoint for such studies.

The approach was two stages. In the first stage, the objective was to just estimate a magnitude of the treatment effect and to assist in sample-size determination for a follow-up study that would be adequately powered. The second stage of the two-stage process would involve doing a trial to provide substantial evidence of efficacy and additional long-term safety data.

In 2005, we had an opportunity to present to BPAC this plan for the Phase 4 postmarketing studies to look at clinically meaningful endpoints for the intravenous alpha-1

PI products. The committee supported the approach that we were taking, to go beyond the biochemical surrogates.

The specific design features that the companies have agreed to:

- For the initial phase, just to estimate the size of the treatment effect, to use a randomized, controlled, parallel, masked design. We did not mandate that these trials necessarily had to be placebo-controlled. An active control, such as dose-control study, could be an acceptable option.
- A minimum of 60 subjects, 30 per group. But the appropriateness of this number would depend on the endpoint chosen.
- A minimum one-year duration to avoid seasonal bias in pulmonary exacerbations. Again, the duration could be dependent on the endpoint chosen.
- It would be important to measure baseline and steady-state antigenic and functional alpha-1 PI blood levels in order to better understand their correlation with the more clinically meaningful outcomes.
- There would be an option for a post-trial follow-up.

There is currently one manufacturer out of the three sponsors who have licensed A_1 PI products in the United States that is conducting a two-year placebo-

controlled, randomized, controlled trial, using lung CT, in 100 subjects. The completion date is 2011. All the other sponsors have agreed to conduct such studies. The protocols are under development. Currently, FDA strongly encourages the inclusion of a higher-dosage arm in the study design because of the potential for greater benefit with higher dosing.

CBER has identified serial high-resolution computerized tomography lung-density measurements as an appropriate clinically meaningful endpoint to assess the efficacy of augmentation therapy with the IV products on emphysema disease progression. We seek BPAC's advice concerning the use of HRCT to evaluate efficacy of aerosol alpha-1 PI products.

But HRCT exists as an option among many potential endpoints, including:

- Mortality/lung transplantation.
- We have talked about FEV_1.
- Exacerbations.
- Exercise capacity. A timed-walk test is another option. But we don't have an estimate of what the treatment effect size might be for that particular endpoint.
- Diffusion capacity and Kco would be other things that could be considered.

Currently, we see as one attractive option for a licensure pathway -- first, for an inhalation therapy product, but also for newer IV products. We think that the same type of pathway could be appropriate, and we seek BPAC's advice in this regard. A typical pathway for an inhalation product might be to perform a clinical study with a primary endpoint of HRCT lung density, but to also include important secondary endpoints, such as PFTs, exercise testing, exacerbations, and ELF analytes and serum levels -- ELF in just a subset of the patients, most likely -- and secondly, basically do the same thing for intravenous products, to switch from relying, premarket, on only the biochemical surrogates and move to using HRCT lung density as a primary efficacy endpoint. You will see information later as to the validation and the types of numbers that we would need and the practicality of that, in the follow-up presentation by Dr. Dirksen.

Thank you very much.

DR. SIEGEL: Thank you, Dr. Pierce.

Finally, we'll hear from Dr. Asger Dirksen, from the Gentofte Hospital, University of Copenhagen, on "QCT and Disease Progression in AAT Deficiency: Results of EXACTLE and Danish-Dutch Studies."

Agenda Item: QCT and Disease Progression in AAT Deficiency: Results of EXACTLE and Danish-Dutch Studies

DR. DIRKSEN: Thank you for inviting me.

I shall talk about quantitative CT as an outcome measure in alpha-1 antitrypsin deficiency.

First, I would like to make a few historical points. The first one is, Hounsfield, who invented computer tomography, actually envisaged this scanner as a densitometer, because it's basically measuring the attenuation of the uptake that it creates an image of. This attenuation is closely correlated to density. So essentially CT measures density.

The next is the definition of emphysema, which goes back to 1985, by Gordon Snider, who actually defined emphysema as the permanent enlargement of air spaces distal to the terminal bronchioles, accompanied by destruction, but without fibrosis. In essence, this is disappearance of lung tissue. A better word for emphysema would probably be "pulmonary porosis," in comparison with osteoporosis, which is a disease where the bone disappears.

This has been known for a long time from autopsies. This was shown already by Professor Chapman. You have the normal lung on the left and the emphysematous lung on the right. It's quite obvious that the density of this lung is quite low. This is nowadays easily seen on CT, where you see all the holes that these patients have in their lungs.

In 1990, the Scottish chest physician, David Flenley suggested that CT might be a better measure of progression of emphysema than the more traditional lung-function measurements.

From a technical point of view, there are some difficulties with the CT measurement. The first one is that you need to find the lung on the image. This is nowadays done automatically by software that segments the lung. When you have done that, you can take out all the voxels, all the picture elements, which are essentially density values, and you can make a frequency distribution. This is actually a normal lung, where most of the voxels have a density at around minus 850, which corresponds to 150 grams per liter.

What happens when you develop emphysema? Here you have the distribution you saw before, just in another scale, because then there is more space for the alpha-1 patient. We have actually followed one such patient now for more than 10 years. This was his first scan in 1993. Here you see his scans during the next years, where you see that the frequency distribution gradually moves to the left, indicating that the density goes down. At the end, in 2004, you see that the density of this lung was very low. He actually got a lung transplantation one year ago. Otherwise, I think he would be dead today.

I just want to say a little bit about the technical issue, how to get one measure out of such a histogram. The first thing that you probably should do is to turn the frequency histogram into a cumulative histogram. You do that as you see it here. This shows how a frequency histogram was turned into a cumulative histogram. When you have done that, you can take out two measures from this cumulative histogram, which are the measures that have been used the most. There are many others, but they have not been very useful.

The most useful one is the so-called relative area of emphysema. The principle is that you choose a threshold -- for example, minus 910 -- and then you define all the voxels below this threshold as emphysema. In this way, from the cumulative histogram, you can get the percentage of emphysema. In this case, the percentage is almost 50, which is a quite high percentage, so this is a quite sick person. This percentage is called relative area below minus 910.

An alternative way to do this, which is essentially the same thing, but it’s just in the opposite direction -- here you choose a percentile -- in this case, the 15^{th} percentile -- and then, from the cumulative curve, you can read the density, the so-called percentile density, corresponding to the 15^{th} percentile. We call that

percentile density 15.

These are, in fact, very much the same, but they have different characteristics. One of the very nice things about relative area is that it can easily be visualized. You can simply highlight all the voxels in the lungs that are below your threshold. In this way, you can see the emphysema. This is not possible with the percentile density. But it has very nice properties when you do longitudinal studies.

This is illustrated in this slide, where you see one baseline cumulative histogram, and then, when the disease gets worse, this histogram moves to the left. That means that the relative area goes up. Here you have the relative area corresponding to minus 910. But the important point here is that the change in percentage is very dependent on your threshold.

You can probably see this better now. The arrows do have various lengths, depending on the threshold you choose.

More important, perhaps, the result is very different depending on baseline. If you have a more normal histogram, you have very, very small changes as compared to the more sick one, which was the first I showed. So the important point is that the relative area is very dependent on baseline and on severity of disease.

This is not true for the percentile density. It is much more robust. For the percentile density, the arrows are horizontal, and the length of these arrows is very much the same, no matter which percentile you choose or no matter where you started, so to speak. If the first curve was here, you will have the same length of the arrows. That is the reason that the percentile density is much more useful in longitudinal studies, because you get much more precise results.

This has also important implications for the major confounder of lung density, which is inspiratory level. It is quite obvious that the density varies a lot with inspiration. This has been known for many years. This is a very old graph, actually, showing the frequency histogram in expiration and inspiration in the same patient. This is obviously very different.

When you do efficacy studies, the problem is that there are two reasons for changes in density . one reason is inspiratory level, if that changes, and the other reason is disappearance of lung tissue. Both will influence density. So when you do these studies, it is very important to be able to discriminate between these two possibilities.

This is a clinical situation where you see a change. Is it due to inspiratory level changes or is it

due to loss of lung tissue? This is easily solved for the percentile density, because you can assume that the total weight of the lung is stable during inspiration. That means that when you breathe, the density goes up and down with volume in a hyperbola. This turns into a straight line in a log-log plot, with a slope of 1. That makes it very easy when you have two measurements. Say this is the baseline measurement of lung density and this is the measurement after treatment -- no, it's actually opposite. This was the baseline, with a relatively low density, and this was after treatment, where you seem to have a higher density. But this could simply be due to differences in inspiratory level.

But you can easily correct for that. You can calculate what the density would be if the inspiratory level was predicted. You can predict that from this sponge model, and then you can easily calculate the real change in density. In this case, you can see that there was a real loss of density, although a simple comparison of the values would have shown the opposite, due to this difference in inspiratory level.

These are, in fact, important considerations about noise in quantitative CT. There are, in principle, two sources of technical noise due to variation in scanning procedure and image analysis, and biological noise, which

is due to variation in the patient performance -- that is, inspiratory level -- and could also be due to other diseases, such as congestive heart failure. If you develop pulmonary edema, there will obviously be a dramatic change in lung density. The same goes for pneumonia. So there are, obviously, biologically sources of noise.

However, technical noise can usually be reduced to a minimum by careful attention to scanning procedures. The biological noise -- the most important one is the level of inspiration. You can minimize that by using the sponge model for calculating an adjusted density, so to speak.

We can go now to the clinical trials, the randomized trials.

The first one is an old one from the first half of the 1990s. These are just pictures from that trial, where they prepared the drug for treatment. Here is the patient getting alpha-1 antitrypsin. In that study, we included 26 Danish and 30 Dutch deficient patients with emphysema. They got four-week infusions. The primary endpoint at that time was lung function. In fact, we thought at that time that by measuring the lung function every day, it would be possible, in such a small trial, to see a difference. We knew when we started the trial that it had not power enough to show a difference with a more traditional lung-function measurement, every three months

or half-year.

This is the result of that study. The daily lung-function measurements proved to be a disaster. I don't want to go into that, because that's another story. I will just here show you the traditional lung-function measurements that did not show any difference between the treatment groups. We knew that, actually, from the beginning. This is FEV_1. Here you have the diffusion capacity. These are the standard errors of mean. You see there is absolutely no difference.

The positive surprise at that time was that the lung density by CT actually showed an interesting trend. The loss in the treated group seemed to be smaller than in the placebo group.

But you see here that there is space for improvement, because even the treated patients lost more than we think normal people do. The data we have seem to indicate that normally you lose almost nothing. So this had only half the effect, or even less, than what we would like.

Another important thing from this graph is that there was absolutely no effect for the first year and then, gradually, the curves diverged as you would like to see. The P value was .07.

The thing we learned was that the time trend for

change in lung density by CT seemed to be much stronger than the change in pulmonary physiology. So that might be a more appropriate measure. Furthermore, we found a positive correlation between decline in FEV_1 and loss of lung density. However, this was not significant. We did also find a positive correlation to pulmonary diffusion, and this was significant.

After doing this study, I have done a follow-up of the patients. This was the time when they were in the study. This was baseline inclusion in 1991. I have done a follow-up. After the finish of the study, they started to die. At this point, half of the patients have died, which is far more than you would expect, according to standard death rates -- “standardized mortality rate,” I think is the technical term.

The mean age at inclusion was 50 years. The median survival for these patients was 15 years, which is much less than expected.

I have done an analysis showing which of the baseline characteristics were most prognostic for mortality. Sex had no significance. Interestingly, age, which is often a very good prognostic factor for mortality, was actually not significant. BMI was not significant. Pack-years was borderline significant. FEV_1 was clearly significant. Lung density by CT was the best predictor of

mortality.

This has also been shown in larger studies, from Stockley's group, for example. I will show that later.

Then the more recent study, which is quite similar to the first one, was an exploratory study, where we now defined CT scan as our primary outcome. Three countries participated, Denmark, England, and Sweden. There is nothing exceptional in this randomized design. They were followed with annual CT for two or two and a half years. Now we used the treatment regimen that was according to the standard recommendations, with weekly infusions.

Study endpoint was defined as lung density by CT. In this trial, we also had other secondary endpoints, not only pulmonary physiology, but also exacerbations and quality of life, by St. George's questionnaire.

Here you have the result, which is very similar to the previous. Again, you have a quite small effect for the first year and then they diverge. Again you have space for improvement, so to speak. When we calculated the P value, it was just the same as the last time. But when you add these two P values together, you have a clearly significant result. If you pool the data, so to speak, then there is a clearly significant difference in lung of lost tissue between the two groups.

Furthermore, in this last study, we did a regional analysis. It's known that alpha-1 emphysema primarily develops in the lower parts of the lung. If you divide the lung into three regions, there was a significant difference in loss of lung tissue in the lower third, which is what you would expect from this treatment.

Again, we could calculate the sensitivity of various outcomes. Again, lung measurements were far less powerful than CT measurements. Here you see that the percentile density is much more powerful than the relative area. In longitudinal studies, this is a much better measurement.

Now we also have St. George's questionnaire, but the sensitivity of this is not very high.

We also have exacerbations. They are not very sensitive either.

In this study, we actually found a significant correlation between decline in FEV_1 and loss of lung tissue.

The last thing is, I will just show three studies that document a correlation between change in CT density, pulmonary physiology, quality of life, and mortality. These are the three studies.

The first one, by Pah (phonetic), shows a significant correlation between loss of lung tissue and pulmonary physiology. Here it's FEV_1, lung density measured

as relative area. Here's it's measured as percentile density -- or the opposite. I cannot really see it. Anyway, this correlation was significant in both cases. It was more significant, again, for the percentile density than for the relative area.

This is a study by Stahl's (phonetic) group in the Netherlands that showed a clear correlation between loss of lung tissue, by the percentile density, and change in St. George's questionnaire about health status. The open circles are actually normal people without symptoms. They should probably have been excluded. That would give an even better correlation. So this was also significant.

Then the last one was the survival study from Stockley's group, which was a larger sample of 256 alpha-1 patients. Here you have their mortality. They were followed for four years. At that time, 20 percent had died. When he did an analysis of lung function, he could show that survivors had better lung function than the non-survivors and that those that died from their lung disease had the lowest lung function, which is by no means surprising. He could show the same thing for CT density. Here it's relative area. Again, the survivors had the lowest percent of emphysema and those that died of their lung disease had the highest percentage. This was clearly significant.

Then he did a Cox survival analysis, where he compared all these variables. Again, he found that lung density by CT was by far the most prognostic factor.

I would like to conclude: Extensive observational studies of CT lung density in alpha-1 deficiency have demonstrated that this parameter not only relates to health status and exercise capacity, but is indeed a better predictor of mortality than FEV_1. Furthermore, CT lung density decline relates to progressive reduction in pulmonary physiology. This has been shown for FEV_1 and Kco and quality of life.

For these reasons, I think CT lung density is perhaps a better outcome than the more traditional FEV_1.

Thank you.

DR. SIEGEL: Thank you very much, Dr. Dirksen.

I think it’s exactly time for a break. Why don’t we take 15 minutes and come back?

(Brief recess)

Agenda Item: Open Public Hearing

DR. SIEGAL: While we are reconvening, I'd like to thank Dr. Freas for bringing the temperature in this room up a degree or two.

I am obligated to read the following text with respect to the open public hearing. Open public meeting announcement for general matters meetings. Both the Food and

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

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

That is the statement. Dr. Freas will introduce the speakers for the open public hearing. We have two on line, and perhaps others.

DR. FREAS: Would the first speaker, Mary Gufstason, please come to the microphone?

DR. GUFSTASON: Thank you. I am Mary Gufstason.

In terms of conflict of interest, I am a salaried employee of the Plasma Protein Therapeutics Association. PPTA is the trade association and standard setting organization for manufactures of plasma protein therapies, including collectors of source plasma.

PPTA member companies are committed to providing safe and efficacious augmentation therapy to people who have inheritable A1-PI deficiency. A1-PI therapies have been available since 1987 when Prolastin was licensed by the FDA, using biochemical surrogate end points as discussed by Dr. Ross Pierce earlier.

Recently interest has shifted to the use of clinically meaningful end points. FDA has requested that sponsors of currently licensed A1-PI therapies perform postmarketing studies using clinically meaningful end points, for example, HR CT pulmonary function tests, pulmonary exacerbations and mortality.

PPTA has used HR CT as a validated and operationally feasible primary clinical end point for the evaluation of A-PI products. HR CT has been utilized in studies outlined in CBER's briefing document, and has demonstrated consistent performance and better discrimination than pulmonary function tests, rate and severity of pulmonary exacerbations, or other clinical end points. HR CT studies can be accomplished with a reasonable number of study

participants, which is an essential consideration when studying a rare disease with a very small patient population.

For systemically administered A1-PI therapy, PPTA views that the effect of these products should continue to be studied using currently accepted biochemical surrogate end points, that is, trough levels. Demonstration of pharmacokinetic equivalents to marketed products should suffice as a surrogate of clinical effectiveness of augmentation therapy. Introducing additional licensing criteria pre or post market may have an inhibitory effect on introducing new or improved systemic A1-PI therapies to provide better patient care.

For inhalation A1-PI therapy, PPTA agrees that the serum A1-PI level cannot be used as a surrogate marker of efficacy due to the local mode of administration. PPTA views the current knowledge of HR CT provides assurance of obtaining a clinically meaningful end point for pivotal efficacy studies for approval of inhalation therapy products.

Data from retrospective and observational studies as well as exploratory clinical trials support that a therapeutic benefit is obtained with a dose of 60 milligrams per kilogram. It has become a recent topic of discussion as to whether higher doses provide additional benefit. Conventional dose ranging or dose comparison studies to address this issue would be difficult to conduct in this rare

disorder, as fully powered dose comparison studies require a large number of study subjects to discriminate a modest treatment difference between the approved dose and a higher dose level.

For practical reasons, randomized control clinical studies in A1-PI deficiency are limited to 100 to 200 patients. As an example, one PPTA member company is currently running a postlicensure study as required by FDA. The study opened for enrollment in March 2006. It has taken over three years to reach its current enrollment of 130 patients, despite significant efforts to enroll subjects. To get to this high number of patients, 120 sites were screened around the world and 32 sites are currently participating. Such recent efforts to investigate efficacy using clinically meaningful end points make it clear that large conventional dose ranging studies are impractical.

The total number of patients required to participate in studies and the length of time to enroll an adequate number of patients should be considered before recommending dose comparison studies for this small patient population. The slow progression of disease and lack of good biochemical markers currently make conventional dose ranging trials impossible in this rare disease.

While PPTA believes that alternative to conventional dosing studies should be considered, PPTA is

concerned about the potential utilization of biomarkers in dose ranging or dose comparison studies.

To date, biomarkers specific to severe A1-PI deficiency disease progression and treatment effect in patients have not been established, let alone validated as meaningful surrogate markers. These may present future potential, but at this time are an early exploratory phase and are not currently applicable to the therapeutic product development.

PPTA member companies would like to reiterate their comment to providing safe and efficacious A1-PI therapies for augmentation therapy in people with A1-PI deficiency, and look forward to working with FDA to establish the most practical and feasible way to improve therapies for A1-PI deficiency.

Thank you.

DR. FREAS: Thank you, Mary, appreciate it. Our next open public hearing is Dr. John Walsh, CEO of the Alpha-1 Foundation.

DR. WALSH: Thank you. I was diagnosed with alpha-1 antitrypsin deficiency in 1989. I have been on augmentation therapy since '93. Travel paid for by the NIH to participate in a Rare Disease Clinical Research Consortium meeting, and all other travel expenses paid for by the Alpha-1 Foundation.

Thank you for the opportunity to address the committee today. I have submitted a written copy of the Foundation's remarks for the record.

The Alpha-1 Foundation is a national not-for-profit organization whose mission is dedicated to providing leadership and resources that will result in increased research, improved health, worldwide detection and a cure for alpha-1.

The Alpha-1 Foundation has promoted therapeutic development as a service to the alpha-1 patient community since its inception. As the only surviving member of the original group of Foundation co-founders, I am proud that we have three products available for treatment of alpha-1 lung disease in the market, and excited that the next generation of therapies are on their way. I maintain a sense of urgency about the critical importance of more effective therapies and ultimately, cure, as many continue to perish too soon within the alpha-1 community.

In adults with alpha-1 antitrypsin deficiency, COPD is the leading clinical phenotype and for driver of morbidity and cause of premature death in the majority of patients. While alpha-1 antitrypsin deficiency currently is available and can slow the progression of lung disease, the intervention has not proven to be curative, presumably because of protease-antiprotease imbalance paradigm does not

fully explain the development of COPD.

Furthermore, augmentation therapy requires intravenous infusions, making this a sometimes prohibitively expensive therapy and keeping some patients from using it because they are reluctant to submit themselves to the inconvenience of intravenous drug administration.

New interventions are needed. This includes the formulations of alpha-1 antitrypsin and synthetic anti-elastase preparations, small molecules and protein folding chaperones, and perhaps gene and stem cell therapies. For these therapeutic interventions, reliable clinical end points are needed to assess the efficacy and data ranging studies are needed to find the optimal dose for clinical use.

As the Advisory Committee considers changes in regulatory policy regarding clinical and surrogate end points for evaluating efficacy of alpha-1 augmentation therapy, the Foundation wishes to express the following. The Alpha-1 Foundation strongly supports the use of quantitative high resolution chest CT as a primary outcome in the new therapeutic trials directed at lung disease of alpha-1 alpha-1 antitrypsin deficiency. Quantitative chest CT is a reliable means of assessing lung structure in COPD, but the methodologies should be standardized to permit comparison among different studies.

Quantitative chest CT will not only complement

FEV_1, but can replace it as a primary outcome. COPD is characterized by remodeled lung tissue that impairs lung function. Traditionally a measure of lung function, notably FEV_1, has been use to monitor disease progression and assess the effect of therapeutic interventions in COPD, including in patients with alpha-1 antitrypsin deficiency. However, this is an indirect measure of lung destruction and it is now well established that functional and structural changes do not correlate well in COPD. Inasmuch as lung remodeling is a hallmark of COPD, a quantitative reflection of this process by high resolution chest CT may be preferable to FEV_1 as a primary end point in interventional trials.

In 2008 the Alpha-1 Foundation sponsored a scientific meeting in the role of quantitative chest CT in COPD research, especially with a view to its potential as an outcome in clinical trials. The meeting was attended by scientists and clinicians from the pulmonary community, radiologists and key representatives from government and industry. The participants concluded that quantitative chest CT is a reliable means of assessing lung structure in COPD, but that the methodology should be standardized to permit comparisons among different studies.

A joint FDA-Alpha-1 Foundation follow-up conference in 2009 on new end points in alpha-1 antitrypsin deficiency related clinical studies reached a consensus on

proposing quantitative chest CT as an accessible primary outcome measure in clinical trials. This end point has already been successful and used in two European double blind placebo control augmentation studies in patients with alpha-1 antitrypsin deficiency, and Quantum, an ongoing NIH-Alpha-1 cosponsored natural history study of COPD in patients with alpha-1 antitrypsin deficiency, which also uses quantitative chest CT as a primary outcome.

We therefore believe quantitative chest CT will be invaluable in future clinical studies of COPD, and this includes study in alpha-1 antitrypsin deficiency. Quantitative chest CT will not only complement FEV_1, but can replace it as the primary outcome.

New drug development requires dose ranging studies. It is impractical to use chest CT as an end point in these studies, because structural changes in the lung occur slowly and require years to detect by lung imaging. Perhaps dose ranging assessments could be incorporated into proof of concept studies, in which the readouts are a reflection of the intended mode of drug action. Biomarkers of lung inflammation, unopposed elastase activity and elastin degradation in sputum, bronchoalveolar lavage fluid and serum or urine could be considered. In such short term investigation of aerosol alpha-1 antitrypsin administration to patients with cystic fibrosis, some of these sputum

markers were found to support the intended action of the inhaled protein. A fixed dose of alpha-1 antitrypsin was used in that study. It may be feasible to include different doses in similar protocols in patients with alpha-1 antitrypsin deficiency with the intent to optimize dose with respect to efficacy and safety.

The Alpha-1 Foundation recognizes that while providing critical pulmonary data, dose ranging studies might delay the start of pivotal new drug trials. A compromise approach, whereby fixed dose clinical trials are conducted in parallel with dose ranging studies could be considered.

The Alpha-1 Foundation community looks forward to the next generation of products and encourages the FDA to expedite licensure of new therapeutics for the treatment of alpha-1. The Alpha-1 Foundation and the alpha-1 community have made significant progress in the treatment and cure as a result of the work by many in this room, and we wish to commend the work of several of the investigators that have presented, and also the FDA and their focus on this issue.

Thank you.

DR. FREAS: Thank you very much, Mr. Walsh. Is there anyone else in the room at this time who would like to address the committee? Seeing no one, I turn the microphone back over to you, Dr. Siegal.

Agenda Item: Open Committee Discussion

DR. SIEGAL: Thank you, Dr. Freas. It is time for discussion. I believe we are open for discussion at this point. Dr. Fleming, you had expressed an interest in speaking at this point.

DR. FLEMING: Yes, I had mentioned to Dr. Siegal that we are dealing as we do in all disease areas with complicated issues about choosing proper end points. We are being asked today to talk about assessing appropriateness of biomarkers and surrogate end points.

There is a rich science for evidence based validation of biomarkers. I wish we had a half an hour to talk about this. I think it is important to do so, because we are not recreating wheels. These are issues that have been talked about over and over again. Yet I realize that we have to keep discussion as our background short so we can spend as much time for interaction as possible.

What I would like to do is just take a couple of minutes to highlight some of those principles that have been widely developed and implemented when it comes to determining validity of biomarkers as surrogates.

Dr. Temple at FDA was one of the first, although there have been countless others, that have indicated the definition of a clinical end point, ultimately clinical efficacy. Those are direct measures about how a patient

functions, feels or survives. In this setting as I can best see, there are many that would directly represent how a patient functions, feels and survives, exacerbations, overall survival, lung transplantation, FEV_1, six minute walk test, quality of life by validated PROs, those are all such measures.

We have often however recognized that to show an effect on a valid clinical end point would require large trials or long term trials. So there has been an interest in looking at whether we can have surrogates or biomarkers that reliably tell us how a treatment will affect outcome in a more efficient way.

The classic quote that we often have to remind ourselves about is that a correlate does not a surrogate make. If we are looking at trying to improve a clinical end point, it is weak evidence to show that a measure is correlated. So for example if we are looking at high resolution CT and you wanted to see whether effects on that represent effects on mortality, it is very weak evidence to show that that measure correlates with mortality. The essence of the information for a valid surrogate is, a treatment effect on that measure has to predict effect on that outcome.

I'll just give a few examples. In fact, it would be good to take time to talk about. There are 50 examples I

could give where we have correlates that have given us very, very misleading results. Just to mention a few, in osteoporosis, bone mineral density is of interest to enhance density and reduce fracture rates. Yet it doesn't represent bone quality. So interventions can affect bone mineral density but adversely affect bone quality and give you the wrong result.

In HIV, CD4 count is strongly correlated with risk of AIDS and death. Yet you can give IL-2 and spike CD4 and in essence create higher cells that are functional and you have no impact on AIDS defining events around mortality.

In MS there has been a long interest in CT and MRI. The idea is, can we look at MRI in contrast enhancing lesions as a way to predict whether interventions affecting progression to major disability. Yet it has consistently been shown as we probe that those MRI measurements are not reliably predicting whether treatments are truly affecting what the patient truly cares about.

We have 500,000 Americans a year who are being treated on the basis that arrhythmias post MI are strongly correlated with sudden death. Yet the currently suppressed arrhythmias triple the death rate. Half a million Americans a year are being treated with an agent that triples the death rate, because we believed in a correlate.

There have been many recent examples where we have

gone awry because of off-target effects. We know that in lipids we want to lower LDL, we want to raise HDL. Yet Torcetrapib, a torvastatin, was recently studied because for the first time we would do both, and this should improve our MI death rate. Yet the study was terminated early with an increase in mortality with Torcetrapib, even though it was doing the intention of increasing HDL.

There have been major recent concerns in end stage renal disease and cancer induced anemia, where we know that hematocrit is strongly correlated with risk of survival; we need to normalize hematocrit. But erythropoietic stimulating agents have turned out to potentially increase the death rate and have other unintended effects when you try to use them to fully normalize hematocrit.

In type two diabetic, an advisory committee for FDA met last year and said, we now have to have large scale cardiovascular clinical end point trials for all agents, because in the past we have relied on hemoglobin A1C, which is strongly correlated with microvascular complications. Yet the ACORD trial, which was targeting complete normalization of hemoglobin A1C, was stopped early when there was an increase in mortality.

Other examples that have existed with Rosiglitazone and Eroglittizar(?) that also showed adverse effects on macrovascular complications, cardiovascular deaths, stroke

and MI, when it was thought that this would be positive.

So essentially when you have a correlate, it may be that that isn't capturing the essence of the causal mechanism. Just one quick example, mother-child transmission of HIV. A pregnant mother with HIV, the lower her CD4 count, the more strongly likely it is she is going to transmit HIV to her infant. You can give her IL-2, spike that CD4 count 100 cells in the ninth month of pregnancy. It will do nothing about transmission, because it is not the causal mechanism.

So what is essential is for us in the disease setting to understand what are the causal mechanisms. Part of the problem is, there may be causal mechanisms that we are capturing in part, but there may be other causal mechanisms of the disease process that we are missing.

So I am just asking questions here. It may be that high resolution CT of density is an important component of how this disease process leads to clinical outcomes. But beyond lung density there is lung inflammation, there is the other elements captured by FEV_1. There are morphologic features, there is functional measures, there are pathologic sources of noise that we heard about, edema and pneumonia, and I'm sure a number of others that we could identify for people that have a richer understanding of this setting than I do.

But if there are multiple pathways through which a disease process influences outcome, we can miss how interventions truly are affecting outcome by just looking at certain of those pathways. Even when you understand the right pathway, you have to understand the magnitude and duration of effect. So you might say post MI with a thrombolytic, it is patency, stupid. You want to restore blood flow to prevent heart damage and improve 30-day mortality. But mistakes have been made because people have used the wrong timing. They have used establishing blood flow at the wrong time for what really mattered for protecting the patient.

So even if we have the right causal mechanism, what is the magnitude of the effect that we have to see and the durability on HR CT in order to reliably know how our treatments are affecting clinical outcomes.

Then the other aspect that I mentioned is off target effects. The biomarkers are not representing the totality of the effect. They are typically going after the on target effects that we are intending to achieve and they are not capturing the off target effects.

So how do we proceed? How do you validate? There is a hierarchy of biomarkers. The highest level are measures that truly are clinical end points. Again, a clinical end point, going back t bob Temple's definition, is a direct

measure of how a patient functions, feels and survives. A patient comes to you as a caregiver, what are they asking specifically to have done? Are they asking you to be able to feel better, to function better, to survive longer? Those are the direct measures at the highest level. Then there are validated surrogates. They are rare. To validate a surrogate, and I saw only two sources of information that really get at validity of surrogate for high resolution CT. A valid surrogate is one where studies show a treatment effect on the biomarker predicts a treatment effect on the clinical end point.

So when the cardiovascular advisory committee, heart and renal advisory committee, was asked to determine whether we can now validate blood pressure in antihypertensives, we were given data on 500,000 patients from randomized trials as the basis to understanding the relationship between treatment effects on blood pressure lowering and treatment effects on clinical end points.

Other examples exist, but the two sources of trials that I have heard today are the Danish-Dutch setting and the EXACTLE study of the second trial. The first one actually shows a different result for the effect on high resolution CT compared to the effects of FEV_1. This is exactly what you don't want to have happen.

Now, the argument is, yes, Fleming, but it is more

sensitive to look at the effects on high resolution CT. But the kind of evidence that validates a surrogate is when you show disease induced effects on the biomarker predict accurately treatment induced effects on the clinical end points.

In the EXACTLE study, there was an indication that the correlation between the treatment effects on lung density in FEV_1 were not strong. The exacerbation frequency was unaltered in that trial, even though there was the suggestion of an effect on the HR CT.

So bottom line is, we do want sensitive measures. We want to be able to assess effects particularly in rare settings like this in as small a study as we can. But ultimately we don't want to compromise reliability of insight for getting a quicker answer. And ultimately the goal in general should be to show effects on direct measures of clinical benefit or at least on measures that are validated, and those are rare, validated surrogates.

But the third level as we talked about this morning are measures that are reasonably likely to predict clinical benefit. Then of course there is a whole array of the fourth level which are correlates that aren't reasonably likely to predict clinical benefit.

My own sense in summary is that the HR CT long density may well be much better than just a correlate. It

may well be a biomarker reasonably likely to predict clinical benefit. But the traditional evidence based strategy to move it up to the second level of being a validated surrogate, we have two studies, and they are in fact pointing in the wrong direction.

DR. TRUNKEY: As a follow-up on that, I would like to ask Dr. Dirksen some questions about repeated CT scans in these patients. There is preliminary evidence from the United States that if you give a single CT to a young person, one in 1,000 may develop a cancer in the future.

We know from treating young people with Hodgkin's Disease particularly in the chest that some will develop mesotheliomas and carcinomas of the lung even though they have been cured from their Hodgkin's, 20 years later.

So you are going to be doing CTs, I thought I saw, three times in the first year, and then yearly for up to four years. What are you going to do to make sure that this is safe for the patient in the long term? Because I have had concerns about repeated exposures to CT.

DR. DIRKSEN: I think this is a good question, because it is potentially dangerous to use X-rays. But when we do these measurements, we can use very low doses of X-rays. We do use what is called a low dose technique, which is only one-tenth of the usual dose when you do a CT of the lungs. So I think the risk is much smaller than what you

mentioned.

Another important point is age. Radiation is much more dangerous for children and much less for old people. In this context, people above 50 years have a reduced risk. These people, or at least those who participated in those trials, were actually above 50, most of them. I think in the last trial they were all above 50.

So the dose is low. One CT scan corresponds to three, four months background radiation. Due to the age, the risk is even lower. So I don't think this is a serious problem. Obviously there are rules for this. There are international rules about how much ionizing radiation you are allowed to use in these kinds of trials. It has been improved, these trials, by these committees. So I don't think this was a problem in the trials that we did.

DR. TRUNKEY: You had two groups that died. Those that died, as I recall, some died because they had not received the treatment, and how many of those that died without the AP died from carcinoma of the lung?

DR. DIRKSEN: That is a good question. Lung carcinoma is a very common disease, as you all know.

DR. TRUNKEY: You have true environmental toxins that basically could --

DR. DIRKSEN: Yes, but it is a very uncommon disease in alpha-1 deficiency. The reason is that these

people cannot smoke enough to get lung cancer, because before they have smoked that much they will die of their lung disease.

So I have never seen an alpha-1 patient with lung cancer. Lung cancer is a very common thing in usual smokers, but I have never seen it in that. It obviously can occur, but this is not a common problem. All the patients that died in the trials that I have participated in died of their lung disease, of COPD so to speak, or emphysema.

DR. ZIMREN: I treat patients with Hodgkin's Disease. The patients in those long term follow-up studies that develop lung cancer have received radiation therapy of the chest, not just CT scan. So their dose of radiation was quite a bit higher than would be received even with conventional CT, much less the lower dose.

DR. TRUNKEY: When I first started doing CT primarily in trauma patients, I had no idea that if I was doing them in children I might be setting them up for a problem, either.

DR. STOLLER: I appreciate Dr. Fleming's comments about the criteria for establishing surrogacy of end points. With the same clarity I would like to reframe the data to your point about the relationship FEV_1 and CT, given that FEV_1 has been the traditional outcome measure.

In fact, as Dr. Chapman pointed out, there are two

trials, observational albeit, that have suggested that often patient therapies associated with benefit in FEV_1, the largest of which was in the NHLBI registry, and which many of the people in this room including myself participated in.

That trial demonstrated that in the FEV_1 30 to 65 percent range, the so-called APS stage two, FEV_1 35 to 40 100 percent predicted for that trial, that there was a statistically significant difference between the two groups. In the trial they included 1,129 individuals, 747 of whom received alternate patient therapy at some point.

It did not use CAT scan. This was done in the early '80s. It wasn't funded for that, did not look at CT. Concordant data emerged from the Dutch-Danish study, much smaller study, 197 compared to 98 patients, but showed almost identical detriments in the FEV_1 in the recipient group compared to the non-recipients. Again, observational data.

We then fast forward to the data of Dirksen and colleagues' initial randomized trial, which included about 58 patients, which showed self collected flowometry and lab based flowometry no trends. But I would remind you, the noise of FEV_1 measurements was perhaps demonstrated convincingly for those of you as pulmonologists, who were noticing that the day to day variation, the test to test variation, may be as high as 100 ml, which far obscures the treatment effect that was observed in these large trials.

The decrement in FEV_1 smallest was 27 ml per year, which is four-fold lower than the test to test variation on serial spirometry measures, even carefully done.

So to your point, the interesting nature of these data is that the correlation between spirometry and CT does not emerge from within a single trial, that is to say, the EXACTLE trial, or in the Danish initial trial, partly because of the very small numbers of patients.

But in aggregate, given the two largest observational studies, both of which are concordant with regard to the subset of the respondents and the magnitude of effect on FEV_1, if one takes the totality of the data about FEV_1 and looks at it in the context of the more recent high res CT, I believe that your criterion of establishing correlation between this new surrogate and the conventional measure is in fact satisfied, although it is not satisfied within the context of a very small trial.

So that your points are fabulously articulated, but they need to be contextualized for the specific nature of the data. As someone who have spent, as have colleagues in this room, most of my academic career studying this, this is a disease constrained by the ability to have multi-center randomized control trials every time something is done. But there is no one center experience that provides adequate experience of which perhaps we have seen demonstrations of

that.

So I would just volunteer the notion from the perspective that in fact, your criterion is satisfied with the asterisk that the data don't emerge from the same trial. But I think it is overly ambitious, given this particular disease and the nature of FEV_1, to expect it to have done so with trials that accrue fewer than 100 patients.

DR. FLEMING: Great points, and I concur with what you are saying. While it is very valid to say that if you look at the Danish and Dutch experience, and the discordance was more than just minor, i.e., it went in the wrong direction on FEV_1, but your point is very well taken. That is, there is not strong evidence against concordance because of the smallness of the trial.

The issue of the burden of proof is not on me to disprove it, it is on those to prove it. While your points are well taken, I might argue the rigor of the data that you were mentioning to say that there is concordance. We really do need measurements across the same trials. We need randomized trials. We need to avoid over interpretation, et cetera. So your points are exactly on target.

I would say kind of a truce here is, the data don't truly refute a correlation, but they don't begin to truly reliably establish the correlation that we would need to say, a validated surrogate is level two. But this could be a

level three. I.e., this readily could be a measure that would justify an accelerated approval type decision because it is not reliably telling us about clinical benefit, but it is making it reasonably likely, that then could be purposed in a postmarketing validation trial.

DR. TERRY: I would like to ask Dr. Chapman and Dr. Pierce and any others in the room a question about that particular subset of patients who had an FEV_1 between 30 and 50 percent, who appeared to get benefit from the replacement therapy.

My assumption is that in spite of the fact that they appear to have a slowing of the deterioration of their FEV_1's, that nevertheless they ultimately did have a continued deterioration. A significant number of them must have ended up in a group below 30 percent. That was the subset that you don't haw evidence of efficacy.

My question is this. Have you looked at the data on the group that initially appeared to benefit, who subsequently fell into the lower half of the 30 percent to see if the slope of their deterioration in fact changed once they entered that group, or if the slope actually changed that would suggest there is some intrinsic property of the lung that is the issue, and it has nothing to do with the replacement therapy.

DR. PIERCE: Dr. Stoller can correct me if I'm

wrong, but one interesting feature of the analysis of the NHLBI registry study, when you are talking about the subgroup analysis by FEV_1 category, they didn't use the baseline values of FEV_1, but rather they used the sum total of all of the FEV_1's over the course of the trial, if you read the original paper very carefully.

This contrasts to many of the other studies that we have been talking about, where they were talking about baseline FEV_1. But I'll let other people answer the rest of your question.

DR. STOLLER: A couple of points in that regard. The duration in the NHLBI registry was to a maximum of seven years. As I recall, mean duration of follow-up was about four years. Given an average rate of decline of FEV_1 of about 54 mls per year, as I recall from the overall group, it would actually be a very small subset of patients who crossed stages from APS stage two to one.

So from a practical point of view, the analysis which you rightly point out would be interesting, would whittle down to such a trivial number of patients, that it could not provide meaningful data.

The other issue, and I think we have seen this alluded to in Dr. Chapman's comments, is that the rate of change of lung function is actually a function of baseline FEV_1. So as one's FEV_1 falls to 30 or 20 percent predicted

there is very little room to detect a signal. What happens in individuals whose FEV_1's decline from a 15 percent predicted below that they either [electronic interference] or expire.

So there is actually no statistical opportunity for the opposite of a healthy worker effect, there is no statistical opportunity to identify a difference in FEV_1 slopes about any intervention in individuals whose FEV_1's are below 30 percent predicted.

So in my own mind, the statistical reasons, and possibility of ascertaining a signal of the analysis about FEV_1 below 30 percent predicted, is confounded by the fact that it is going to be nothing to do with treatment, but have to do with the clinical state of affairs and natural history of individuals who fall into that very limited subset of individuals.

DR. CHAPMAN: To follow on those thoughts, I am certainly not aware of an analysis of patients who crossed that numerical threshold. As Dr. Stoller has pointed out, the rate of FEV_1 decline varies relative to baseline. I think that was shown in my presentation most clearly in the patients' natural history data from the UK registry.

Just to follow on that limitation of FEV_1 as an end point, I would like to get back to Dr. Fleming's comments, and point out that FEV_1 itself is a surrogate end point. It

is a time honored end point, and we think we know a great deal about it, but in fact it is only approximately correlated with mortality. A great deal of the FDA's current work in COPD in general is developing other end points that are more pertinent to patient clinical outcomes.

In the clinical research arena we are busy trying to devise better indices such as the Bode index, which would incorporate patient functionality, body mass index and so on that would be better associated with or predictive of outcome.

So I would reframe the discussion as comparing two somewhat validated end points that move in the same direction. The FEV_1 is not as clear a surrogate end point and certainly is not a validated one in the context of this discussion.

I will also point out that lung density is a much more direct measure of emphysema as we currently define it, that is, the permanent loss or destruction of the airways distal to the terminal bronchioles. FEV_1 is not even close to a direct measurement of that particular phenomenon. We all know of patients who have emphysema in the absence of spirometric findings or FEV_1 abnormalities.

DR. FLEMING: That is also a very good point. I would like to bring you along with me when I try to make that argument, as I try to make over many instances over the past

25 years.

There are different levels of direct clinical relevance. Without any doubt, mortality and exacerbations and lung transplantation and quality of life measured by validated PROs are direct measures of clinical benefit.

I have frequently raised for debate the issue, are PFT measures, FEV_1, FVC, are they sufficiently -- and I will put another one in that category two, the six minute walk test, which has been the traditional standard by which we have been approving agents in pulmonary arterial hypertension, and we have talked about it in idiopathic pulmonary fibrosis settings.

So I agree with you in part. I put them more on the edge. My personal view is, they are not as direct clinical efficacy measures as the other things that I have mentioned. I have had a whole lot of pushback on that, where many have argued that no, it really is very much getting at how a patient is functioning and feeling, very much in a tangible way.

So many have argued against us, saying that PFT measures like FEV_1 and FVC and the six minute walk test are in the category of direct measures and how a patient functions, feels and survives.

DR. ZIMREN: We haven't talked much about the use of inhalation therapy and its effect on high resolution CT.

One thing I don't know is, if someone develops fibrosis, for example, as a result of any kind of inhalation, what does that do to the lung density?

DR. BRANTLEY: Again, I want to remind you that these patients are deficient in alpha antitrypsins, so we are replacing them. So the likelihood that protein that is normally seen in these patients' lungs would cause fibrosis is pretty small.

Could you develop an interstitial lumenitis from the inhalation? I think it is possible. I think one of the safety concerns when we design these types of studies is, we can't pick up these things. I think it is one of the reasons why we need to do bronchoalveolar lavage as a portion of these clinical trials, because we know some of the subtleties are markers of hypersensitive lumenitis, and probably would detect them before we would proceed on to fibrosis, which would be a problem.

Obviously some of the same CT scans that we use for the detection of loss of lung tissue can be used also for the gain of lung tissue as well, like for instance fibrosis. That would be a qualitative measure rather than a quantitative measure in those particular cases.

DR. CRYER: I think the appealing thing of the CT scan as a measurement has to do with the concept of pulmonologic reserve. The idea that by the time you get any

kind of clinical abnormality you have to knock out 90 percent of the organ to actually see it, that is way too late to intervene in the disease process.

In fact, the data that we have heard today would make me think that the hypothesis is wrong. Just because you are missing a protein, it you get it back it doesn't do any good. The patient still deteriorates.

So the thing that I think would be appealing is, if you could make the connection -- and I agree with you that some of that data has to be mined to see if there is a correlation between how long it takes for overt clinical failure related to different CT scan slope changes or something along those lines -- I think really the potential benefit of all of this may be to identify a group of patients that you can intervene in a lot earlier. Taken when they have a normal CT scan of the lung, normal pulmonary function, start treating them then, and seeing if there is a difference in how long it takes for the CT abnormality to show, you have a lot better chance of making somebody survive a lot longer time than if you already have somebody who is on the C part of the curve, where their pulmonary function and their pulmonary units measured by whatever structural measure, are already deteriorating rapidly.

DR. BALLOW: I want to ask the question, in the time element of these clinical trials, let's say it is a year

that these clinical trials are done, do you actually see changes in the chest CT, where you expect changes in a short period of time, in one year, when a clinical trial would be going on? Did your sites answer that or look at that?

DR. DIRKSEN: That is easy to answer. You can definitely not see anything in the usual chest X-ray. In fact, it can be difficult to see emphysema at all in chest X-rays. So it will not be possible to see any change.

In some cases, the patient develops bullae and that can be seen. Did I misunderstand you?

DR. BALLOW: High resolution chest CT.

DR. DIRKSEN: Oh, you are talking about chest CT?

DR. BALLOW: Yes.

DR. DIRKSEN: Oh, I thought you were talking about the chest geography. So the question is whether you can see emphysema on chest CT? No, no, again the question is no, subjectively you will not be able to see a difference. You can only see it when you do these computer analyses.

DR. BALLOW: What is the time on it? In other words, when you do high resolution chest CT as a surrogate marker, and these trials say are only a year long. Is that enough time to see differences in the high resolution chest CT?

DR. DIRKSEN: This question depends on the measuring error compared to the change. If you take for

example FEV_1, the change is around 30 mls per year or something, the measuring error is 100 mls per year. So you haw to follow a patient for almost ten years before you can get a reliable measure of the slope of his FEV_1.

With HR CT it is better because the loss is larger compared to the measuring error. So in the individual patient you will be able to calculate precisely within three, four, five years. But this is absolute a minimum. You cannot see anything within one year due to the measuring error. Do you understand? Okay.

DR. TERRY: A major contributor to lung density on CT is the volume of the lung. I wanted to ask the question, because I noticed in several of the presentations that there was a suggestion that people with alpha-1 disease have airways disease, they have bronchospasm, and that bronchospasm can lead to air trapping, and therefore a reduced density of the lung. That might potentially be misinterpreted as emphysema when it is simply air trapping. Would you comment on that?

DR. BRANTLEY: That is a very good question. Indeed, when we run these trials we make sure that patients are fully broncho dilated before we do them so we have the same amount of air trapping consistency. For the Quantum 1 study, all the patients received Spiriva prior to, which is one of the most effective drugs in preventing air trapping,

so every patient gets that just prior to their treatment or their assay for their lung.

So we pay close attention to that, and try to minimize that effect. But obviously just like with FEV_1, it has some challenges. We have learned how to decrease the variability in the study.

DR. TERRY: Could I ask the follow-up question, at what lung volume do you these studies, and how you standardize those?

DR. BRANTLEY: It is usually TFC.

DR. DIRKSEN: It is usually after a full iteration. If you had trapping, that would influence the shape of the frequency distribution. We do see that when they develop bullae. Some of these patients develop bullae, and you can see that from the distribution, because then you get a new peak corresponding to the bulla. So if air trapping was a big problem, that would change the shape so to speak of the distribution.

DR. TERRY: And if you do these at TLC, TLC can change as a function of body weight. So how do you standardize for that?

DR. DIRKSEN: That was what I tried to explain by the sponge model. You assume that the weight of the lung was not influenced by inspiratory level. So if you increase weight you may not take such a deep breath, and that means

that the inspiratory level changes from one examination to the next. But you can eliminate that source of error by adjusting to a fixed volume, which is usually predicted TLC.

DR. STOLLER: (Comments off mike.) I have a couple of questions pertinent to the questions we will be asked to respond to, primarily to Dr. Dirksen.

As we are asked to consider the impact of inhalation, Dr. Brantley pointed out appropriately that the inhalation of a protein, the volume of the protein is relatively small on any given inhalation. But the question that I have for you regards what we know about the serial inhalation episodes of protein to the lung and its impact on a CT even in normal individuals. Is there any such data about examining protein and its effect on lung density, not only with a single episode inhalation, but over a prolonged episode, as would be the case in a drug given by inhalation such as alpha-1 antitrypsin.

Two, to the question also framed in the first A1 that we are asked to comment on, when you look at exacerbation frequency and you had CT, the CT as I understand the trial were done at baseline six, 12, 24 and occasionally 30 months. I assume that one of those CT scans were done during exacerbations, but perhaps you have accumulated in your experience some data on CT density during an exacerbation. It would be relevant for us to know what the

impact of an exacerbation on CT density during the actual episode was.

So those are a few separate questions.

DR. DIRKSEN: The first question about accumulation of the drug in the lungs is a very relevant question. If you inhale 200 milligrams per day, then that would amount to 70 grams in a year. What the effect of treatment is, around one gram per liter per year. So if you have a lung of certain liters, where the effect of treatment will be to save seven grams of lung tissue per year, that is only ten percent of what you have inhaled.

So if ten percent is accumulating in the lungs, that would be the same as you obtain with your treatment. So this is a very relevant question. I do not know of any data that have looked at this accumulation, but there is data on radioactive drugs that have been inhaled that as far as I know show that they disappear, that they do not accumulate.

If you had such an accumulation, you would probably see the strongest effect during the first period, or you would see the same thing during the whole period. The two studies that we have show almost no effect during the first year. But I admit that they are small, so you cannot really prove that. But at least there is no indication of an accumulation.

The other question was about exacerbations. The

difficult thing about exacerbation is that they are very different. For example, if you have a severe pneumonia, this would be an exacerbation, and that obviously has a strong influence eon density, no doubt about that.

So the usual practice is that if the patient has any indication of exacerbation, we would postpone the CT scan for four to six weeks. But an exacerbation could also be for example that relatives of the patient go on holiday, and the patient gets stressed because he is afraid of being alone and gets a small dyspnea and come into the hospital with what we call an exacerbation. In that case, you would obviously not expect any change in lung density.

So I think exacerbation, we all know what it is. It is hard to define, and there is no doubt that it has -- it may have even a pronounced influence on lung density. I don't know if that answered your question about exacerbations. I'm not aware of any people who have studied lung density during exacerbations.

DR. STOLLER: Just to refine the question, so during individuals in criteria three, where they have copious secretions and worsening dyspnea, are you aware of any CT data during episodes like that, not the individual patient, because that is the exacerbation we would be --

DR. DIRKSEN: No, I don't know of any data. but it is very important when you do these studies that you do a

careful inspection of your CTs while you are still blinded about treatment. I have seen a few cases of pulmonary edema, for example, and also pneumonias that did not give symptoms. In that, CT scans must be excluded because they have a heavy influence eon the lung density. You can easily find them as outliers when you do the analysis. You can see that suddenly the density just goes way out. then when you look at the scan there is usually an obvious reason.

DR. KULKARNI: Of the patients who did not respond well or their lung functions went down, did they develop any kind of inhibitors to antithrombin? Or did they have mutations in their neutrophil elastin which makes them non-effective against the antithrombin?

The second question I had was, I was reading your American Graphics Society review about an 11-year-old children getting this. Do you think that this product given very early will prevent this from childhood onwards? I think the 11-year-old child did have lung findings on CT scan. Is there a response related to inhibitors?

DR. BRANTLEY: I am assuming instead of antithrombin you are talking about neutrophil assays in that case. There are a couple of different things. One is, like just about every other disease, alpha antitrypsin deficiency, treating it appropriately, is a moving target.

We have patients that live a normal life span that

have alpha antitrypsin deficiency. If they had never encountered a reason to recruit neutrophils into their lungs, they may never develop significant lung destruction. It is just not clinically apparent without a CT scan.

For those individuals that smoke, become a pediatrician, or many other things in which there will be an excuse for neutrophils to track, those patients can have a rapid decline in their lung function. There is both a lifetime decline as well as punctuated kinds where these patients will decline very quickly.

When we talk about somebody declining at 85 mls per year, that is the aggregate. What they actually do is, they get respiratory tract infections and they drop by 300 mls or more, and then they come back up and not quite up to baseline typically.

So in choosing doses and making decisions about which patients are going to go down, it is really hard to figure it out. As was mentioned, we are looking at longitudinal data and not being able to make a decision on that individual point as to what is going on with them. But clearly those individuals that have high neutrophil burdens and such because of exacerbations, because of heavy smoking, tend to go down much, much faster. Indeed, people that smoke, actively smoke and have alpha antitrypsin deficiency, can go down by about 120 mls per year, as compared to an

individual.

The other thing that is important as another trick about this is that patients' rate of decline as was mentioned is different at different stages of their life. For instance, people that are at what we call the fast burn stage, around 35 percent to about 65 percent, where they are declining quite quickly. On the upper end of the curve they go slower, on the lower end of the curve they go slower as well.

Those same issues also make it tough as far as which test to use as well. For instance, a six minute walk is not particular useful for an individual who has an FEV_1 of 65 percent predicted. They will be normal, normal, normal. But when you are 35 percent predictive, a six minute walk is a very valuable test, because that is an issue for those particular patients.

So when you design a clinical trial that has to encompass all the patients that you have, picking out the right tests to look at the whole frame of the individuals is tough. You have to think about it very carefully.

DR. TRUNKEY: I wanted a clarification from Dr. Chapman. You had a cartoon that showed the results of not having elastase. I think you said that if you give the protease you get rid of the elastase. But I also thought you said that it was a scavenger. It is not a scavenger of free

radicals, is it?

DR. CHAPMAN: I might take the microphone just to pass it on to Mark Brantley, but no, I don't think it is a free radical scavenger.

DR. BRANTLEY: It is.

DR. CHAPMAN: It is. Okay, I have been corrected. It is.

DR. TRUNKEY: Then the question is, when you have that one week curve that somebody showed and you only get a day and a half effect from the inhaled protease, it seems to me you are subjecting the patients to repeated stimulation of the white cells, is that correct?

DR. CHAPMAN: You are talking about the inhale form of therapy?

DR. TRUNKEY: Yes.

DR. CHAPMAN: That wasn't my presentation, I don't think.

DR. TRUNKEY: I know, but somebody showed a slide; over a week period you only got about 36 hours of normal level of the elastase, is that correct?

DR. SANDHAUS: (Comments off mike.)

DR. CHAPMAN: The serum levels from your presentation?

DR. SANDHAUS: (Comments off mike.)

DR. SIEGAL: Would you come to the microphone so

the transcriber can record our conversation?

DR. SANDHAUS: I think I am understanding the slide you are talking about. Ross Pierce showed a slide that showed the blood levels after a single weekly infusion of augmentation therapy in someone who had been receiving weekly infusions for at least four weeks, and showed the normal or super normal level in the initial days and then a rapid decline.

The inhalation therapy would be expected to be given daily or even twice a day. So that might obviate some of the issues that you are questioning.

DR. TRUNKEY: Thank you.

DR. HOLLINGER: Initially a comment. I don't know if there is a difference. The two studies that you commented on, the Dutch-Danish study and then the Swedish-Danish-UK study, I think the only other difference in that study was, one was given every month, 250 milligrams per kilogram was given every month, versus the 60 milligrams per kilogram. Where that might have had some differences is difficult, because it is another confounding variable.

A question is how one standardizes using -- when there are so many different CT scanners around and other things -- how one standardizes the high resolution CTs in a trial. What site do you look at? Do you look at the lower zone? And is it the better site to look at, or the middle?

How do you make that standardization? That is one of the questions I would like to ask, if anyone here could respond to that.

The second thing is just for education primarily. If smoking primarily causes an upper zone emphysema and so on, and alpha-1 antitrypsin causes a lower zone emphysema, and smoking makes alpha-1 antitrypsin disease much worse, the question I have is, do we know why patients with alpha-1 antitrypsin disease specifically have a lower zone emphysema as compared to say smokers in general? Do we know that information, or any reasons why pathophysiologically?

Anyway, the question has to do about standardization, but I would like to know about the other thing too, if anybody has an answer.

DR. STOLLER: I can speak to the latter. First of all, the notion that it is upper versus lower is like most clinical pearls a little bit oversimplified. In fact, the best data come from David Parr's group and Ralph Saki's group, who have done detailed CT imaging in these individuals.

The preponderance of folks have a lower lobed, about two-thirds, but a third have classic upper lobe disease. Again, it is biologically naive to say it is only lower lobe. It is usually a gradient with a preference toward the lower lobe, even in those two-thirds that have

that preference.

Now, having said that, that blurs the margins between classic upper lobe disease and lower lobe disease. It is not very well understood, even when there is a predilection for the lower lobe. It perhaps has something to do with lung perfusion, where lung blood flow is greater than one basis, therefore the neutrophil elastase, the number of polys in the lower lobes might theoretically be greater and therefore the neutrophil elastase burden might be greater in the lower lobe, exactly where there is a paucity of anti-elastase defense related to the deficiency of alpha-1 antitrypsin.

So that is a hand waving argument for which there is very little good clinical understanding. But the primary response to your question would be to say that like the classic blue bloater and pink puffer distinctions that we all grew up with, it actually doesn't work very well. If you depend on recognizing only individuals with lower lobe disease, you will miss most of the patients in your practice with alpha-1.

DR. SIEGAL: Could I come back to Tom Fleming's assertion that lung density might not have any real meaning? As a clinical immunologist, I confess to being very naive, and so I am asking this of the pulmonologists here. Isn't lung density central to the pathogenesis of emphysema? Isn't

that actually intuitively the most reasonable end point that one ought to be looking at?

DR. FLEMING: Just to clarify what I was saying, I surely am not saying it doesn't have any meaning, just as on that litany of examples that I gave where we did clinical studies looking at effects on biomarkers, in every one of those cases those biomarkers were chosen based on insight -- imperfect though it was, insight that this is part of the critical mechanism of the disease process.

So I am not quibbling on whether or not in this setting you would say that lung density isn't meaningful. I am asking, doe a patient approach their caregiver saying, hey, doc, I've got to do something about my lung density. I suspect not. They say, hey, doc, I want to be able to live longer. I want to be able to feel better. I want to be able to function better. The Temple definition of a clinical end point is a direct measure of that. The issue is, there are numerous examples where based on our partial insight, we do have a sense about meaningful mechanisms. But if you just change a mechanism that is not tangible to the patient and don't change how they can tangibly function, feel and survive, that is not clinical efficacy.

DR. SIEGAL: That is why I asked the question, really, of the pulmonologists. Is that in fact predictive of the way people eel ultimately? Does lung density over time

predict what is going to happen?

DR. FLEMING: I'm not even quibbling with that. What I am specifically asking -- this is critical, folks -- what I am asking is whether we know that a treatment induced effect on that measure is reliable evidence of a treatment induced effect on a measure that is directly representative of how a patient functions, feels and survives.

So I am completely agreeing that these are measures that have biological importance, or you may monitor them as you are monitoring patients in the clinic, and they are correlated with outcomes.

So biomarkers serve five purposes. It is important to realize they are distinct purposes. We use biomarkers to detect disease. We use it to assess prognosis. We use it to manage patients and adjust treatment strategies. All of those can be used by just being correlates. You don't have to be the causal mechanism.

But the two more complicated are enrichment, effect modification, identifying the patients that will truly benefit, and as a surrogate replacement end point such that it is sufficient to know I am achieving clinical benefit simply by showing an effect on that biomarker. That is where the science is far more complicated.

That is the only thing I am challenging, but that is critical to the answer to the first question we are going

to be asked.

DR. SIEGAL: So let's talk about it a little bit more. Can we hear from the pulmonary group?

DR. STOLLER: Well, as one of the talking pulmonologists in your midst, I will be happy to try to address that.

Dr. Fleming has appropriately addressed this earlier, but I would argue that if 100 years ago, before we popularized flowometry, we had the availability of CP and spirometry at the same time. It rules out in my mind that we would have picked CT as a more direct measure of the pathogenesis and pathophysiology of emphysema.

Notwithstanding your comments, Barchelli's work on the Bode index, body mass index, degree of obstruction, FEV_1, dyspnea and exercise on the six minute walk has been shown probably to have the best prognostic data, and embedded within it measures that would satisfy your criteria about functional status and symptoms.

If we had both measures available to us awhile ago, we would not be having this conversation today, I'm quite sure, because the pathophysiology and the pathogenesis of emphysema is of course loss of alveolar wall primarily, giving rise to bullae which in turn compromises the tethering function of the lung, which is manifested several pathophysiologic steps downstream as a decrement in air flow,

because the tethering below has compromised the springs that keep the airways open. When those springs are over stressed or lost, the elasticity of the lung changes, and one can exhale less gas less quickly.

So the FEV_1 is a pathophysiologic, several steps downstream, correlate for what we are trying to measure. So as its base, I believe both based on the data shown here and other data not included in these conversations, it is a far more direct measure of what we are looking at when we are looking at the impact of the drug on the disease processes alone.

There is one other comment that gets back to Dr. Fleming's excellent comments before. Given that the mortality as a hard end point, perhaps the only study large enough to have addressed mortality in alpha-1 as a function of augmentation therapy is the very same study, the NHLBI registry, that recruited 1,129 patients in 37 centers, followed over seven years. Most of the people in this room, as was I, were investigators in that study. There was a highly modeled observational cohort study, not randomized by location. There was a powerful survival benefit in that trial in favor of individuals receiving augmentation therapy, with a risk ratio of mortality highly statistically significant at .64 in the overall group, and in the stage two APS subset of .21. Arguably a 79 percent decrement in

mortality related to the receipt of this drug, modeled for everything else including socioeconomic status, be it data on education and income, which have been argued as perhaps the most important socioeconomic surrogates.

When we wrote the paper very carefully out, I plead guilty to that. I was one of the authors of the paper. The first paragraph of the discussion has all the disclaimers about the shortcomings of a non-randomized trial.

That said, this field is tortured by the mandate of multi-centered randomized trials which as we have heard are very difficult to do in the context. So we are often left with less than perfect associational studies around the very criteria you pointed out.

But the bottom line is, there is evidence, observational though it is, that augmentation therapy has concordant effects on FEV_1. In the only trial large enough to look at mortality, it sh owed a concordant effect on mortality in the very same groups in which there is a FEV_1 signal.

Then again fast forward to smaller randomized trials, there is data that high res CT is a more sensitive measure of the impact of those drugs which have been shown to be associated with decrement and decline of FEV_1 and a mortality benefit in the earlier trials in which CT was not done, for very good methodologic and practical reasons.

So that really is the landscape around which we have to make the associational decisions. We do not and probably will never have a randomized trial of thousands of patients that are receiving this drug, in which we have the opportunity to make the logical reasoning that you have quite articulately and beautifully laid out in terms of what the mandate for evidence is.

But we have been plagued by perfection is the enemy of the good, and the field has not moved along because of that somewhat staid perspective, which I would argue has been an impediment to actually making progress. We are often left with the very arguments that we are confronted with today, which is, are we looking at something that at its base is embedded within the pathogenesis of the disease. I would argue that high res CT is in fact that measurement.

DR. SIEGAL: Thank you, Dr. Stoller. That was very helpful. What about Dr. Terry? Do you have any comments along these lines? Or Dr. Choyke?

DR. TERRY: I would like to ask Dr. Stoller a question. I have been impressed with the alpha-1 patients I have taken care of, in the meticulous care by the nurses of these patients on a weekly or a bimonthly basis. I can't help but ask if part of the survival benefit in the study that you are describing wasn't a function of this meticulous observation by nurses with quick feedback to physicians, so

that they intervened more quickly with exacerbations, et cetera. That is why my biggest concern is that we will never answer this question without a randomized double blinded placebo control trial.

DR. STOLLER: Let me say that I am a great advocate and have published my opinion that randomized trials are sorely needed. But I am also grounded in the practical realities of the possibility of doing them. So I feel wholly schizophrenic in my response to those answers.

At least in the NHLBI registry, where we were looking at these differences, the frequency of patient visits is actually relatively comparable in groups. We were mindful of that. The number of times these patients saw their docs and nurses was relatively comparable in the compared groups. It would be hard to ascribe the differences observed there to the confounding variable of clinical attention in that particular study.

DR. CHOYKE: Can I return to Dr. Hollinger's question and address it to Drs. Brantley and Dirksen, about setting up a multi-center trial with CT, including the effect of multi-detector CTs, how do you make sure that all these different kinds of machines are calibrated in the same way so that the numbers are comparable, are any rules given about doing the same patient on the same scanner every time.

As far as the analysis is concerned, is it done in

a core facility or is it done locally? As you set up these multi-center trials you can introduce all kinds of errors. So what is your feelings about that?

DR. DIRKSEN: It is absolutely correct that the scanner type and all kinds of parameters are very important for your result. So the most important thing is that you do the scans on the same machine when you look at longitudinal changes, and preferably with the same radiographer, actually, to make sure that you have the same protocol in the machine. In my experience, various radiographers do this. They forget to take the right protocol and so on.

So this is a problem, but I don't know any other way to eliminate it than by just doing exactly the same each time. If you take another scanner there will definitely be some differences, although there should not. In principle it is standardized, but in the real world it is not.

DR. CHOYKE: Do you use phantoms, like a standard phantom that you would use across a study?

DR. DIRKSEN: Yes. There have been some studies of that. The problem is that to be useful, the measurements with the phantoms should be much better than the measurements that you are doing with a patient. That is a challenge. It is hard to do measurements on a phantom that are so good that you can use them to improve your data from the patient.

DR. SANDHAUS: The work that was done by Asker in

the studies that you have heard has gone into the improvement of our standardization of CT scanning in subsequent studies. The study that Ross mentioned that is being done in almost hundreds of sites around the world now to obtain a larger group, every site has their own phantom that is scanned each day that the patient is coming in for enrolled study evaluation. All of the data on both the phantoms and the patient studies from that day are sent to a central location.

As an investigator in that study, we have actually had to bring patients back within a week to repeat the CT when we get word virtually in real time from the central reading facility that there was a problem, whether it was, the tube was on the slow decline or something along those lines.

So those questions are important questions, but we have come up with ways to at least minimize the effect of the multi-center variation.

DR. BALLOW: Mr. Chairman, can I change the topic a little bit? We have talked an awful lot about the high resolution CT, but we haven't talked about biochemical and surrogate markers. It struck me when there was discussion about the different products, there was a very nice table about the different antigenic levels that are achieved in the various clinical trials, but there are holes when they talked about functional activity. This relates perhaps back to

finding the appropriate dose of these replacement therapies that it may be that at the antigenic level they would not be a good surrogate marker, that function activity may be much better. In my mind, there may be a better way of either measuring it in the lung as a potential biochemical surrogate marker. I don't know about the serum. Some things can happen between the local and the serum, particularly if you go by the inhalation route.

I think there should be some discussion about these biochemical surrogate markers. I would put in my push to measure functional activity certainly in the lung itself and perhaps in the plasma as well.

DR. BRANTLEY: I did all the assays for all of these studies. We have a very robust activity functional assay which has been used in every one of these studies. The dosing is done not on the antigenic amount, but is dosed on the functional amount of alpha-1 antitrypsin that is in the vial.

In all of these circumstances we measured both the antigenic amount and the functional amount. There is indeed as Ross mentioned a difference. The antigenic and the functional don't correlate 100 percent. But the reason for that is because when you take normal alpha antitrypsin and you inject it into somebody that is alpha-1 antitrypsin deficient, their alpha antitrypsin is half as functional.

So when you do the mixing experiment, where you take 100 percent functional and you take the native and mix it together, which is half as functional, you never quite achieve 100 percent.

DR. SIEGAL: Do you measure the amount of residual elastase activity?

DR. BRANTLEY: That is the methodology. It is an anti-neutrophil elastase capacity assay, and that is the function.

Again, just to remind you, alpha-1 antitrypsin does more than just act as an anti-protease. We have taken one biochemical feature of alpha-1 antitrypsin and used it as our assay to find function. I think it has since stand the test of time.

DR. CRYER: I'm not sure I heard the answer to Dr. Siegal's question. Can I get specifically drilled down on it? There must be data for this in emphysema. If you took a group of patients who got serial CT scans and they all have emphysema, and you take the group that got clinically worse over three years, are their CT scans uniformly worse in that three-year period of time?

Conversely, if you had a group of patients whose CT scans were no change over the three-year period, was their clinical respiratory function the same as it was three years ago? There must be data on that.

DR. STOLLER: Do you think there would be better data about that? Embedded within this issue is, in the specific context of alpha-1, the data exist about that. You have heard today. You have heard the trials. So I defer that to Asger, who can speak to the individual patients. But you have heard the world's body of experience in alpha-1 with regard to that specific issue today here.

DR. CRYER: But that is why I asked about emphysema as a whole, because there is lots of them around.

DR. DIRKSEN: It is hard to accept, but the measuring error is so large that you have to observe a single person for a very long time to get a precise estimate of what is happening in this patient. So it is anecdotal evidence.

For example, as I said, if you develop a bulla, then the patient may get worse and die. You can see that on CT, of course. But when you just say a patient gets worse, that is also a difficult thing. How do you define when the patient gets worse? If you define it in the sense that their questionnaire changes, for example, then there is proof that a change correlates to a change in lung density, but definitely not on an individual level.

But there is evidence on larger patient groups. But when you are talking about individuals, that is what I heard, you ask for single persons, then it is not a very precise measurement. So it is very difficult to correlate in

individual patients, unless you observe them for a very long time, of course. That would solve the problem. If you follow them for five years, for example, then you will be able to see within a patient that there is a correlation between the clinical situation and lung density by CT. But these data are very scarce, obviously.

DR. CRYER: Well, that is consistent. I think it addressed part of Dr. Fleming's issue. In other words, if you have to go a long time before you can see a big effect, a big change, but it is always the same direction, it is always a clinical deterioration and a worsening of the CT scan. Then the CT scan is more sensitive to small changes over time, then I think that is probably okay, as long as it is consistent. But you are saying we really don't have enough information to say that.

DR. DIRKSEN: No.

DR. FLEMING: Can I just mention that it can be even more complicated in the following sense. Here we are, we are intervening today. What we are trying to do is to achieve the more distal clinical benefit of reducing the tangible things that patients care about, mediated through an intended effect on a mechanism.

I'm not questioning that we have mechanisms here that we reasonably understand. My problem is, there are multiple mechanisms, some of which aren't captured by the

mechanisms that we are looking at, and there are off-target effects.

I'll give you two examples to show you that I can have the exact correlations that you were talking about, and still be way misled. If you have end stage renal disease it is very clear, the lower your hematocrit, the higher you are going to have a risk for death in MI. That is true in standard Epogen therapy. So we randomize patients to high dose Ipogen therapy to get even more normalization of hematocrit, and it worked. The higher dose Ipogen therapy gave a much better normalization of hematocrit. And within the low dose there was a clear relationship between hematocrit level and death, and within the high dose, a clear relationship of hematocrit level and death.

The trial was expected then to have a one-third reduction in death rate. It was stopped when there was a one-third increase in death rate. The reason is, you are tracking an intended positive mechanism, but the primary end points also get hit by off-target effects that are completely missed by those correlations. So even though you were normalizing hematocrit even better on high dose, you were introducing thrombotic events that were offsetting it.

You can also miss on-target effects. If you look at pertussis vaccine, big thing that we are looking at there are antibody responses. So when I was on the Bacterial and

Biological Products, there were two vaccines that we were looking at. One of them clearly had better antibody responses in the classic areas that we cared about, but it was far inferior in the overall assessment of protection against pertussis, because the other vaccine had better effects on other antibodies that weren't the ones that were targeted, but also contributed.

So you can track a mechanism and see that it is correlated with outcome, and it is not capturing the off-target effects and it is not capturing the on target effects that aren't part of that surrogate end point.

The bottom line is, we shouldn't overstate what we know. In 1998 this committee was asked to opine on the appropriateness of using these biochemical surrogate end points, and the committee said it was okay. That was based on what we knew 11 years ago. What I would like to avoid is, in 2020, 11 years from now, this committee comes back knowing a whole lot more about these biomarkers, stating that we were misled at this point.

It is an imperfect world. My biggest goal here is for us to acknowledge what we know and what we don't know. We know a fair amount, but there is no evidence here that is telling us that a treatment effect on these biomarkers is reliably telling us we have a treatment effect on the clinical end points. I think stating that we are isn't

helping today's caregivers, and it isn't helping the development of the scientific process.

Now, it doesn't mean that we can't move forward. It might be the judgment that where we are today is, we don't have validated surrogates that would be at the basis of using them as the sole end point. But we do have clues, and those could be the basis of developing interventions that in fact, either by proof of concept get further validated on clinical end points or maybe even by accelerated approval get marketed, but in the context of doing validation trials that will reliably answer the question on the true clinical end points long term.

So my main goal first and foremost is to not overstate what we know scientifically.

DR. GOLDING: Can I make a point, Mr. Chairman? Dr. Fleming is highlighting the possible off-target effects. In my understanding, what we are talking about here is a replacement therapy or a replacement of a protein that the body normally has, and we understand its function. Plus, there is a large amount of clinical data. We are using this over a long period of time with repeated therapy.

Unless I am mistaken, there aren't any severe off-target side effects that I have understood happened in other trials where drugs were used for arrhythmias or EPO was used for renal failure. So I think it is a different paradigm

here where we are talking about replacement therapy. I think that does make a difference, because we are talking about not only replacement therapy, but replacement therapy where we have a lot of experience of the product used for these patients.

The other area which I think has already been studied by many of the pulmonologists, we have been talking about a biomarker, and we are talking about CT, high resolution CT. The question is whether this really is a simple biomarker, or whether it really is looking at the disease itself.

You are actually defining the emphysema. The only other way to define the emphysema is at postmortem when you fix the lungs and you look at it. You are looking at this in a live person, and I would argue that this is higher than most biomarkers, and maybe shouldn't be called a biomarker at all.

DR. FLEMING: But what we are not acknowledging with that is, it took us ten million plus patients treated erythropoietin stimulating agents to recognize these off-target effects. We don't have ten million patients.

I'm not blaming us that we don't know all the off-target effects, but to argue that we know what the off-target effects are of these interventions because they have been used is certainly inconsistent with many other settings that

have led to across FDA a greatly enhanced recognition that off-target effects are inadequately understood, even in settings where we have the ability to do larger scale randomized trials.

It only was when we had 50,000 patients in randomized trials with the COX 2 inhibitors and ten million doses given with erythropoietin stimulating agent that we are really understanding hat these off-target effects are.

A non-randomized trial is effective at recognizing relative risks of ten. So if you are inducing intussusception with a rotovirus at a tenfold increase or PML with Cassavri at 1,000-fold increase, the kind of data that you are talking about gives us that insight. But it doesn't give us insight about odds ratios of two, in terms of doubling risks of things that we are not assessing in randomized trials. It could be highly influential on the ultimate benefit to risk.

DR. STOLLER: I'll just share with you, in terms of off-target effects, I hear two potential things. One is the risk of the intervention and the other is the mechanistic implications that go beyond the current paradigm as we understand it.

Let me just speak to the risk, to do some of these calculations in my head. The experience with regard to risk of the drug, and this has been looked at both in the registry

as well as in the European experience, in this country it is probably hard to know precisely, but probably in the 5,000 recipients of augmentation therapy with one drug or another.

The first drug available approved by the agency in 1987 was Prolastin. This is obviously an accumulated frequency, but let's say there are 2,000 patients -- I'm guessing here -- that have been receiving once weekly drug at a prescribed interval since 1987. I don't know how many doses that is, you can figure it out in your head. In that context, there is pretty good surveillance about adverse events. In terms of cataclysmic outcome events, they have not been observed.

Now, that may not belie the subtle issues that you brought up, but I am reasonably confident that the administration of this drug is not producing those adverse catastrophic, acute apparent events.

DR. FLEMING: That is a half a million people a year, so were the cardiologists not very observant? I would defend the cardiologists to say, a relative increase, even by a factor of three, is not transparent in the absence of a randomized trial.

DR. STOLLER: Fair enough. One would have to grind these data. I offer the numbers in the hope of allowing that kind of calculation of what signal would be apparent in the context of those data.

DR. HOLLINGER: I know one of the other things that we are supposed to discuss is dosage, potential dosing of these things. I think it is important to point out that it looks like at least that even in patients who are treated, there is a deterioration in most of the trials that are looked at. It is not as much as those that have not been treated, but it could suggest, it doesn't have to suggest, but it could suggest either it is effective but not that effective, or that a larger dose may be beneficial.

So therefore, I think that any trial that goes on with new drugs at least, new medications, should look at different doses.

DR. SIEGAL: Do we all lose lung function over time? So the question is comparative rates.

DR. ZIMREN: I had a question. As part of the briefing materials, there was a paper about the use of inhaled insulin and what looked like a significant decrease in the initial part of the therapy. There is a parallel until the end, where the decrement in FEV_1 appeared to resolve after therapy.

I'm not a pulmonologist, and it was hard for me to put this into perspective. Is this something that is meaningful? Is there maybe more known about this? And is there any reason to think that something more worrisome might happen with this mechanism if it is real -- I don't know if

it is -- in patients with alpha-1 antitrypsin deficiency?

DR. STOLLER: Let me speak to that. I actually happen to sit on the endocrine committee for the agency that reviewed inhaled insulin. I for the record voted against that drug on the basis of the consideration you have brought up.

Having said that, two comments are germane. I will just remind the committee that we are talking about the existing paradigm of serum levels, which again are a consideration for intravenously administered drug, but are not a consideration for inhaled drug. It is clinically impossible to measure serum levels as an outcome measure for an inhaled drug, for the reasons that Dr. Brantley and others pointed out, because the lung is a very good membrane that prevents the passage across the epithelial barrier into the interstitium and into the epithelial space into the bloodstream, which is where you would be measuring it in inhaled drug.

So to insist on the existing metric of serum level as an outcome measure in a clinical trial of inhaled drug is to say that there is no trial because it is an impossible measure to make in that context. So one is left with other metrics.

Having said that, the question you raise is an important one. I don't know these data. It would be quite

important to know in the context of the inhaled insulin trials about the impact of an inhaled protein on spirometry as a simple function of the inhalation of a drug in itself into the lung, which theoretically could have an adverse effect, and having nothing to do with the efficacy or lack of efficacy of the drug on the basic pathogenetic process, which is the development of emphysema, because of course the FEV_1 is an accurate measure of air flow. It aggregates effects in the central airways, in the tethering function of the distal airways.

It is a gestalt, if you will, a pulmonary gestalt measurement of air flow, which could clearly be affected either favorably or potentially unfavorably by the inhalation of a protein. So one would have to know those data as a prelude to a good understanding. But again, serum level, which is the conventional paradigm for the three existing -- four existing drugs, is not a possibility for an evaluation of an inhaled drug. It is not possible to measure the drug by conventional criteria by this new route of administration.

It has been said in the briefing documents. I just needed to reiterate that point for the consideration of the group.

DR. PIERCE: To review the data with Exubera, the inhaled insulin on FEV_1, in the paper by Jay Schuyler et al. published in Diabetes Care in 2007, Volume 30, page 579, it

stated that over a two-year period of time, the difference between the inhaled insulin group and the subcutaneous insulin group was 17 milliliters worse for the inhaled insulin group. That was over the two-year period.

Between months three and 24, the difference was 11 millimeters. But that difference between months there and 24 wasn't significant, whereas if you subtract those two you get about six millimeter difference in the first three months. That was actually significant.

DR. ZIMREN: It may be numerically significant, but it doesn't sound like --

DR. PIERCE: Right, you are talking about small differences, but you want to put the 17 milliliters -- compare that to the 27-milliliter difference that was seen for example in the registry study.

DR. SIEGAL: Is there any other discussion? Dr. Fleming, anyone?

DR. FLEMING: Not until we get to the clarification of the question.

DR. EPSTEIN: Without trying to anticipate voting or individual comments, let me just ask Dr. Fleming, would you accept HR CT as a likely valid surrogate at this stage of knowledge? Because then all we are really debating is whether we can use HR CT is a surrogate end point for conventional approval, or a likely valid surrogate for an

accelerated approval.

When you think about this field, everyone understands that in the long run you want hard clinical end points. Therefore, phase IV studies are essentially inevitable, whatever you do up front. So if we are going to have longer term monitoring and outcome reporting regardless, then the only thing that we are really debating is whether to give full approvals because we think we have a valid surrogate or conditional approvals because we think we have a likely valid surrogate.

I think that perspective might enable us to reach some kind of consensus on how to use the HR CT. I would just draw attention to the fact that if we decide we can't use the HR CT, what we are left with are the historic clinical end points, FEV_1, exercise tolerance, et cetera.

Those studies are likely to take upwards of five years, because as has been pointed out, they are downstream consequences of the deterioration of the lung. Its functional ability tends to be preserved beyond the point where you can see anatomic deterioration.

So what does that mean? That means we won't have new products for a longer time if we wait for the hard clinical end points versus, we accept HR CT, which the EXACTLE study showed us had earlier sensitivity.

That didn't prove that it was valid. I understand

that point very clearly. But if we think it is likely valid, then we can get to new product approvals in two to three years, at least on the basis of accelerated approval, but it would require a scientific assessment that it is at least a likely valid surrogate.

DR. FLEMING: You anticipated exactly what, when our Chair just said, are there any more questions, I said not until we get to the questions.

I was hoping that you would accept what you just proposed. That is, that question one could be softened to simply asking whether we think it is reasonably likely to predict clinical benefit. My own sense is, it might predict clinical benefit. We do not have the data at this point to view this measure, this high resolution CT of lung density, as a validated surrogate. But it could readily be judged by the committee to be reasonably likely to predict clinical benefit.

The criteria that we typically looked at for that are consistent with what you could know at this point. First of all, is it the committee's judgment that when you look at high resolution CT of lung density, you really are capturing a critical, if not the essential, pathway through which the disease process leads to the risk of clinical end point. That is point number one, do you have the essence of the disease pathways captured.

Secondly, is it based on current understanding our judgment that off-target effects are really implausible at a level that would meaningfully offset the intended benefit through this on-target effect.

Third, is the magnitude of the effect on the biomarker that we are asking for here substantial? The goal of clinical research isn't to get statistical significance; you have heard that from the statisticians. It is to get statistically reliable evidence of meaningful effects. So if you are asking for an ability to approve an agent, to get regulatory approval under accelerated approval using high resolution CT of lung density and you state that to be a sufficiently substantial effect, that it makes it very plausible that it will translate to clinical benefit, and it is capturing a principal mechanism through which the disease process influences outcome. The arguments that were given, off-target effects, while they can't be ruled out, seem to be unlikely relative to the magnitude of the effect, then that is the basis to judge that this would be a measure that could be used for an accelerated approval.

Now, I wouldn't quite de-emphasize the distinction between that and full approval. To me it is a very big difference, because you are still recognizing with accelerated approval that you need to do a postmarketing validation trial to ultimately reliably answer what the

effect is on clinical end points.

But my sense is, yes, I think the answer to one would be very different if it were stated as reason to likely predict rather than reliably.

Agenda Item: Questions for the Committee

DR. SIEGAL: Are there other comments? Thank you, Dr. Fleming. In that case, let's proceed to the questions.

DR. PIERCE: Actually, in my reading of the question, I don't see the term reasonably reliable in the actual question. Question number one as we have it for VRBPAC is first of all just the statement, introductory statement, that CBER has identified serial high resolution computerized tomography, HR CT lung density measurements, as an appropriate clinically meaningful end point to assess the efficacy of augmentation with IV alpha-1 PI products on emphysema disease progression. That is just a statement.

Then the question, does the committee agree that the rate of change of lung density as measured by serial HR CT could potentially be used as a primary end point in pivotal studies of efficacy in alpha-1 PI augmentation for inhalation therapy. So there is nothing in the question directly that addresses whether we would intend to use an accelerated approval mechanism or standard approval mechanism.

DR. FLEMING: There should be, because in any other

advisory committee setting, those are profoundly important distinctions. So in my view, the way this is asked, it is at least implicit that we are asking whether you could give full approval based on this. In fact, I may even say it is explicit by the wording in the preceding sentence that says, you have judged this as a clinically meaningful end point.

So if you instead were stating, does the committee agree that this rate of change could be used as a biomarker reasonably likely to predict clinical benefit that could therefore be used as a primary end point in an accelerated approval, my own view is, the answer to that one is very different from the way it is currently worded.

DR. PIERCE: We certainly have discussed the accelerated approval mechanism. Earlier it was misstated that some of the newer products were approved via an accelerated approval mechanism. That wasn't the case, but we certainly have discussed using an accelerated approval mechanism for any future intravenous alpha-1 PI augmentation therapy products.

DR. SIEGAL: The word potentially is there, and that is not so very different from reasonably likely to be.

DR. PIERCE: No, it says potentially used as a primary end point. That doesn't in any way preclude using this as a primary end point for a full approval.

In my experience on other advisory committees, we

are asked to be more explicit in advising the agency regarding full approval end points versus accelerated approval end points. They are very different.

DR. EPSTEIN: I would like to suggest some rewording of question one that I hope will clarify this.

First of all, the first sentence of question one is really introductory. It is to remind you that were have already accepted this in phase IV trials of intravenous therapy, rightly or wrongly so. Again, as was stated clearly by Dr. Pierce, those studies do involve concomitant collection of clinical end point measures.

So the real question lies in the second sentence of the first paragraph. The proposed revision is as follows: Does the committee agree that the rate of change of lung density as measured by serial HR CT is a reasonably likely surrogate that can be used as a primary end point in pivotal studies of efficacy of alpha-1 PI augmentation for inhalation therapy. So we have changed it to, ask whether we think it is a reasonably likely surrogate. I think the committee can vote that.

DR. FLEMING: Yes. I would have liked it if you would also have said explicitly for accelerated approval. But I think it is now implicit for that because you said reasonably likely; I agree.

DR. EPSTEIN: That is implicit. So again, we need

to record this for the record, but does the committee agree that the rate of change of lung density as measured by serial HR CT is a reasonably likely surrogate that can be used as a primary end point in pivotal studies of efficacy of alpha-1 PI augmentation for inhalation therapy.

DR. FREAS: We are now ready to vote on the modified question that Dr. Epstein just read. For the people that have just joined us in the afternoon, there was three ways you can vote. Yes, you agree, no, you do not agree, or you can abstain. I would like to get a show of hands in that order. Please keep your hand up until I call out your name, until we have a name with every vote.

All those who are voting yes, they agree with the statement, please raise your hand. There are 17 voting members at the table right now, and this is a unanimous decision. I will not call the names. So unanimous yes vote.

Moving on to question 1A now.

DR. PIERCE: Question 1A is, before embarking on pivotal studies, should sponsors first establish to what extent CT density measurements are confounded by, one, inhalation therapy and two, pulmonary exacerbations of COPD?

DR. SIEGAL: Do we discuss this or call a vote? Is there any discussion on this point?

DR. FLEMING: Does this question need a vote, or is it sufficient for us to discuss the answer?

DR. SIEGAL: Jay, do you want to vote, or is it sufficient to discuss it?

DR. EPSTEIN: We would prefer a vote, yes. There has been a discussion. We talked about it.

DR. SIEGAL: Is there any further discussion before we vote? Anybody want to say a few words? In that case, let's proceed.

DR. FREAS: I am going to call for yes votes on question 1A. Those voting yes please raise your hand at this time. Again, it is a unanimous vote for 17 people at the table. No need to call names. Do you need question 1B?

DR. PIERCE: Question 1B again pertaining to the inhalation therapy. Does the committee recommend any additional information regarding HR CT lung density measurements be obtained prior to sponsors initiating pivotal studies of efficacy of alpha-1 and proteinase inhibitor augmentation for inhalation therapy? So that is in addition to the information that was in 1A that you just voted on.

DR. SIEGAL: Any discussion on this point?

DR. STOLLER: I would simply say that this has perhaps not been the subject of much discussion earlier today. Perhaps tipping my hand to my own vote, for example, some of the measures in my own mind that would be relevant in this regard would be an analysis of existing data about the slope of high res CT in normals, to the extent at which it

exists, were advantaged with spirometry. We know the rate of change of normal spirometry and therefore can distinguish between the effect of an intervention and a natural history of the disease.

I'm not aware of it, but I suspect that there are some data already existing perhaps that would allow clarification of that, and if there aren't, subject to the ethical issues of doing CT scans on normal individuals, an attempt should be made to understand that more clearly, as well as just to be mindful that patients with chronic obstructive pulmonary disease often have a number of other comorbidities that in theory could affect lung density.

For example, the prevalence of heart disease among these individuals is very high, and the cause of death is often cardiac rather than pulmonary. So the assessment of heart failure, of pneumonia and some of these other inter-current events, manifestations of these other comorbidities that these patients have would be important to better understand in order to tease out drug effect versus the effect of other comorbidities.

So it has not been the subject of huge discussion at the panel today, but those would be my views on that.

DR. SIEGAL: Anybody else?

DR. ZIMREN: Just a question. So are you suggesting that trials of normals should be instituted, or

that data available should be analyzed? To start a trial now looking at multi-year
degradation of normal lung would postpone things by quite a bit.

DR. STOLLER: In fairness, it is my ignorance as to whether those data exist currently. I suspect there are some data, and it probably could be mined without mounting a large prospective multi-year trial. I suspect that they exist. I have not seen them published. They are not embedded with any of the trials that we see, as one might think they were if they were robust. But it is a question that should be asked, I believe.

DR. PIERCE: Dr. Dirksen I believe did point to a rate of decline in normals that he said should be fairly negligible over the particular 30 or so month time period of one of the trials, in one of the slides that he was showing.

DR. CRYER: It might be important to know the data in smokers, for instance.

DR. HOLLINGER: Yes, and I think we have already discussed the potential possibility of looking at different methodologies, the CT scanners and things like this. I think that was already discussed today, but that clearly needs to be looked at.

A very simple thing. Some of these patients may have sclerosis, so since there can be some hepatopulmonary

problems with these patients, that should be assessed down the line. There is not going to be a large number, but it should be assessed.

DR. SIEGAL: With that in mind, shall we proceed?

DR. FREAS: Again, the options are yes, no, or abstain. May I have a show of hands of those voting yes for additional data? That is a unanimous yes vote.

DR. EPSTEIN: Can I suggest that asking for additional information is an essay question? It is not helpful. I don't think we need a yes-no vote on that question, because we have asked for what additional information would be helpful. I think we heard a useful discussion.

DR. PIERCE: Question number two. Does the committee recommend that FDA reconsider the use of biochemical surrogate end points, serum ELF, antigenic and functional alpha-1 PI levels to provide substantial evidence of efficacy prelicensure of new IV therapy products in favor of more clinically meaningful end points, i.e., HR CT lung density, FEV_1, pulmonary exacerbations or mortality?

So this question is really asking, do we want to move beyond in the prelicensure phase the biochemical surrogates and look at these examples of more clinically meaningful end points in the prelicensure phase, as opposed to using for example an accelerated approval mechanism,

continuing to license new IV therapy products on the basis of the biochemical end points, and attempt only in the phase IV postmarketing to verify the efficacy of the products using the more clinically meaningful end points, as the examples listed.

DR. SIEGAL: Discussion on this? Do we need a vote on this, Jay?

DR. EPSTEIN: Yes, in this case I think we are in need of a vote, because this is somewhat fundamental, whether we are in effect steering companies away from biochemical markers and telling them to use either HR CT or in combination with clinical end points.

This is the reflection of the discussion that we had about uncertainty about levels. We can't just take a serum level or trough level and know its predictive value. We don't know the threshold for an effective therapy. We have had a discussion about ability to use or not use ELF measurements.

So we are basically asking the committee whether we should reverse course and focus on non-biochemical markers. Again, we are talking about the context of IV therapies here.

DR. KULKARNI: Can I just make a comment? Maybe we can insert the word, reconsider the use of current biochemical surrogate end points. What if somebody develops a fantastic test in the future which can measure everything?

DR. EPSTEIN: Sure.

DR. SIEGAL: Is that agreeable?

DR. PIERCE: Our intent is that we are referring to the same biochemical end points that have been used before as a basis for licensure of the products to date.

DR. ZIMREN: Are you asking us what we think should happen tomorrow, or what we should be moving to eventually? We had a long discussion about how we need to accumulate more information, for example, about the HR CT. You are not asking us whether we think now the whole paradigm should be changed, is that right?

DR. PIERCE: First of all, this question, as we are intending it, is only going to apply to brand new IV products. Any IV products which are currently under development or they are active INDs, this would not apply to. We would abide by the negotiations and discussions that we have had with those companies to continue to rely on the biochemical surrogates.

But a brand new company who comes to us, yes, even tomorrow, starting their development program from scratch in terms of the human phase, the question to the committee is, for such a new, new product, does the committee recommend to move away from solely relying on the traditional biochemical end points toward more clinically meaningful end points such as the four mentioned here, including but not necessarily

limiting to HR CT lung density.

DR. ZIMREN: But you are still using language like move away toward. That implies more of a gradualism than abandon and start.

DR. PIERCE: I'm sorry. Usually initially in a development program we have a tolerability dose ranging study. There may or may not be a phase II study looking at maybe some more novel and sophisticated type of biochemical surrogates in addition to the traditional ones, and then a pivotal trial. So all of that takes a bit of time during which there are opportunities to get additional information in the natural history of changes in CT measurements in normals, if it is judged that the data in that regard are absolutely necessary and the current data are sufficient in that regard. There ought to be at least some current data on that.

DR. BALLOW: Might not they complement one another? Might not these biochemical markers complement the other surrogate markers that we talked about?

DR. GOLDING: We are not talking about taking all the other biochemical markers and throwing them out the window. What we are saying is, the HR CT should be the main -- if we are going to call it a surrogate, should be the main surrogate for new IV products. The IV levels will still be a secondary end point, You are not going to throw them out.

It is still something that we want to know about.

We also want to know about other things such as clinical end points where there is a whole list of things that we have said several times. So those things would be secondary end points. The primary end point would be HR CT. That is the point, what would happen tomorrow if a new company came along with an IV product.

DR. STOLLER: Just for my clarity, I am reading the text to say that tomorrow instead of a tough serum level as the primary outcome measure, that we are asked to consider any of the four that are mentioned here, high res CT, FEV_1, exacerbations or mortality, as a preferable primary end point. Any one of those compared to a trough serum level, which has been the existing metric primary outcome.

DR. PIERCE: That is correct. The sponsor would be free to choose any of those four. If there is a fifth that they could adequately justify, we could entertain that as well. But the idea would be, move it in a direction that would be more clinically meaningful.

DR. STOLLER: It deviates a little bit from the question which is, are we endorsing high res CT tomorrow, given what we recognize as the need to do more studies. This allows a smorgasbord of metrics other than the existing serum level, as I read the question. I am just reading it for confirmation of that understanding.

DR. CRYER: It seems to me that you ought to change it to exactly what we already talked about for inhalational therapy. That is basically what we discussed. Isn't that what you want to know? As I recall, the language we used, a reasonable chance that it is a surrogate and potential use, rather than -- this is pretty strict language.

DR. PIERCE: When you say it, are you referring to HR CT?

DR. CRYER: Yes.

DR. PIERCE: Dr. Epstein, is that agreeable?

DR. EPSTEIN: I think the subtlety here, not anticipated before today's discussion, is that we have grouped HR CT with other markers that no one debates is clinically meaningful. So maybe we need to subset this. But what we are really asking the committee is whether we should move away from serum levels.

DR. PIERCE: Serum levels as sole primary end points?

DR. EPSTEIN: Right. So the choice here would be a known clinical end point such as FEV_1, pulmonary exacerbation, mortality, or a likely valid surrogate end point, HR CT, either in preference to serum levels. I think that is the understanding on which one would vote.

DR. SIEGAL: Everybody in agreement?

DR. FREAS: We are voting now on the agreement that

we understand for question number two. All those voting yes? Unanimous yes votes.

DR. PIERCE: Question number three. Does the committee recommend any other alternatives as primary end points for alpha-1 PI product premarketing clinical trials, A, for inhalation therapy, B, for new submissions of IV therapy?

DR. SIEGAL: Any discussion on this?

DR. FLEMING: Is this a discussion question as opposed to a vote?

DR. SIEGAL: An essay question.

DR. FLEMING: So it is an essay question. I am being somewhat repetitive to what is already in two, but what you have listed in two are some of the key things to consider. I would agree, Jay, with exactly what you stated, that there are options as to whether someone would choose to use any of these, FEV_1, pulmonary exacerbations, lung transplantation, mortality, HR CT, lung density. I am happy with all of those. I might throw out six minute walk as something to consider, but I agree with the agency that there would need to be proper validation of that measure, and standardization and assessment of what a difference would be that would be clinically meaningful.

As a comment was made, it would probably be sensitive in a subpopulation that would have a lower baseline

level of FEV_1. But at least I put that on the board as something that could be an option if done properly in the right subpopulation.

DR. SIEGAL: Any other comments? Do you want to vote on question number four?

DR. PIERCE: No, it is an essay question. Only discussion questions. I'm sorry, three is. Are we ready to go to four?

DR. SIEGAL: Yes.

DR. PIERCE: Number four. Does the committee recommend studies of intravenous alpha-1 proteinase inhibitor augmentation therapy include higher doses than previously approved, assuming adequate safety?

I would remind the committee that we do have one ongoing postmarketing study of an IV augmentation therapy product that does not include more than one dose, a placebo control trial. The additional phase IV studies that the other sponsors had agreed to conduct, the protocols for those are still under development. So there would be an opportunity there. If for example the committee decides to recommend exploring higher doses, that would be one opportunity where that could be looked at.

DR. SIEGAL: Does anyone wish to discuss this?

DR. FREAS: Do you want to vote on this? Again, there are three options, yes, no or abstain. May I have a

show of hands, the yes votes for question number four? Again that it a unanimous vote. All 17 voted yes.

DR. SIEGAL: I believe that concludes our day. Unless there are objections, we are dismissed.

(Whereupon, the meeting was recessed at 6:10 p.m., to reconvene Tuesday, July 7, 2009 at 9:00 a.m.)