UNITED STATES OF AMERICA

+ + + + +

DEPARTMENT OF HEALTH AND HUMAN SERVICES

+ + + + +

FOOD AND DRUG ADMINISTRATION

CENTER FOR BIOLOGICS EVALUATION AND RESEARCH

+ + + + +

ANTHRAX VACCINES: BRIDGING CORRELATES OF PROTECTION IN ANIMALS TO IMMUNOGENICITY IN HUMANS

+ + + + +

THURSDAY, NOVEMBER 8, 2007

+ + + + +

The workshop began at 8:30 a.m. in the Grand Ballroom of the Hilton Washington North, 620 Perry Parkway, Gaithersburg, Maryland

TABLE OF CONTENTS

		PAGE
I.	Introduction and Welcome	3
II.	Session 1: Background Information	8
III.	Session 2: Animal Models for General Use Prophylaxis	68
IV.	Session 3: Human Immunogenicity Data	
V.	Session 4: Panel Discussion on General Use Prophylaxis	
VI.	Adjourn	

P-R-O-C-E-E-D-I-N-G-S

(8:33 a.m.)

DR. LYNN: For those of you who haven't met me, my name is Freyja Lynn, and I'm with the Office of Biodefense Research Affairs at DMID at NIAID.

And before we start, I'd like to give you a few housekeeping matters. The first is one of the most important -- that's lunch. We are not providing lunch. However, the hotel and Starbucks have both been notified that they're going to have a lot of people to feed, and so we'll have lunch available in the hotel.

There are also some local restaurants fairly close by. I know we have only an hour for lunch, because we have a very tight schedule today. And there are maps and some of that information out front, if you haven't gotten it with your meeting packets.

The meeting is being recorded and will be transcribed, so when you speak please

state your name first and speak into a microphone. Transcripts will be available after the meeting.

I think there was one other thing, which I don't remember, Drusilla.

Okay. I think that's it for right now. Again -- oh, I remember what it was. We would appreciate it if you would allow the speakers to complete their talks. We have allowed time at the end of each talk for questions, so we'd appreciate you holding your questions until the end of each talk.

So let's go ahead and start. I'd like to introduce Dr. Karen Midthun, who is the Deputy Director at the Center for Biologics Research. Thank you very much.

DR. MIDTHUN: Thank you, Freyja. I'd like to welcome all of you to this workshop today that really has been a culmination of a lot of work by a lot of people, and I'd really like to acknowledge and thank the co-sponsors of this workshop which,

in addition to the Center for Biologics Evaluation and Research, also include the National Institutes of Allergy and Infectious Diseases, and also the Department of Health and Human Services Office of Biomedical Advanced Research and Development Authority.

The subject of today's workshop is how to bridge animal efficacy data to humans in support of developing new anthrax vaccines, and, of course, I think we all recognize that this is an important goal for public health preparedness.

In 2002, we held a workshop to discuss efficacy testing of new anthrax vaccines, and that workshop provided a lot of excellent direction for non-clinical and clinical studies that could potentially provide data to support efficacy of new anthrax vaccines.

And since that time, several studies have been conducted, and we now have a much better understanding --

PARTICIPANT: The sound is very bad. We cannot hear you back here.

DR. MIDTHUN: Is this not working? I'm so sorry. Let me speak directly into the microphone. My apologies.

I was just thanking those who have come today and welcoming them, and also thanking those who are sponsoring this workshop together with the Center for Biologics. That is the National Institutes of Allergy and Infectious Diseases, and also the Health and Human Services Office of Biomedical Advanced Research and Development Authority.

And the subject of today's workshop is on how to bridge animal efficacy data from animals to humans, and this of course is very important in support of developing new anthrax vaccines, and I think we all recognize this development of new vaccines is an important goal for public health preparedness.

In 2002, we held a workshop, and at that time got excellent directions on the

kinds of non-clinical and clinical studies that could be conducted that would help develop efficacy data in animals that could then potentially be bridged to humans. And I think since that time a lot of studies have been conducted, and now we have a much better understanding of the immune response that animals and humans have to anthrax vaccines, and also additional data in animals on efficacy.

And so I think today what we have the opportunity to do is to hear about those data and to further evaluate and get input on the approaches that have been taken on how to bridge data from animal studies to humans, and also figure out what data gaps there might be that would help to further assess this development of new vaccines.

And I guess I'd really like to take this opportunity to say that we really look forward to the scientific input from those who have come to this workshop today. We really

appreciate the input and also that people are so willing to share their data, because this is so important to really furthering the discussion and developing good approaches to this very important area.

And with that, I'd just like to say thank you. I'm really looking forward to hearing all of the discussions today. And with no further ado, I'll hand it back to Freyja Lynn.

DR. LYNN: Thank you, Karen.

Unfortunately, our first moderator, Julianne Clifford, we think is stuck in traffic, because we had -- there was an accident on the Beltway. So I'm going to moderate the first session, and our first speaker will be Dr. Drusilla Burns. Sorry, I can't see anything without my glasses.

So, Drusilla?

DR. BURNS: Thanks, Freyja.

What I want to do today is just set some background, so that everybody starts from

the same place. Now, I know a lot of you are very familiar with the Animal Rule, but what I wanted to do today is very quickly go over it, for those of you who may not be as familiar with it.

Then, I'd like to just summarize some of the very important points that came out of the 2002 Anthrax Vaccine Workshop that Dr. Midthun just told you about. And so let me start by describing the Animal Rule.

This regulation was first published in the Federal Register in 2002, and it's not called the Animal Rule. It has a much longer name. It's New Drug and Biological Drug Products/Evidence Needed to Demonstrate Effectiveness of New Drugs When Human Efficacy Studies are not Ethical or Feasible.

And there's four main criteria that must be fulfilled in order to use the Animal Rule. The first is that there is a reasonably well understood pathophysiological mechanism of the toxicity of the substance and its

prevention or substantial reduction by the product.

The second is the effect is demonstrated in more than one animal species expected to react with a response predicted for humans, unless the effect is demonstrated in a single animal species that represents a sufficiently well-characterized animal model for predicting the response in humans.

The third is the animal study endpoint is clearly related to the desired benefit in humans -- generally, the enhancement of survival or prevention of major morbidity.

And finally, the fourth criterion, which actually turns out to be the most difficult to fulfill, is that the data or information on the kinetics and pharmacodynamics of the product or other relevant data or information in animals or humans allows selection of an effective dose in humans.

So what does this mean for anthrax vaccines? It means that the vaccine dose must elicit an immune response in humans that is comparable to the immune response of animals protected by the vaccine. And it's really this fourth criterion that we're going to be spending the next day and a half discussing how to fulfill it.

Now, there's a number of potential misunderstandings about the Animal Rule. The rule does not apply if the product -- if product approval can be based on standards described elsewhere in FDA's regulations. The rule is not an accelerated or fast track approval.

And I think that it's important to know the rule is not a shortcut to approval, as I think many of the people in this audience are now -- now know. In fact, it may take longer. And human studies are still required under the Animal Rule. You need to have safety studies and immunogenicity studies for

anthrax vaccines.

The important thing to remember when -- as far as the Animal Rule is concerned is that the product is being developed for use in humans, not in animals. So the animal studies must be designed such that the data generated are relevant to humans.

This really means that the animal studies and the clinical studies need to be developed along a parallel track. That is, you have to have some human clinical data to know what the response in humans is likely to be, so that when you're developing your animal model you can keep that in mind and try and mimic the human response in the animals. Then, you can go back and do the larger clinical trials for the pivotal studies.

When designing the animal studies, the label indication is important -- that is, pre-exposure prophylaxis or post-exposure prophylaxis. And for pre-exposure prophylaxis during this meeting we're going to refer to it

as general use prophylaxis or GUP, so you'll be hearing that an awful lot.

When developing your animal model, you should consider route of exposure, appropriate challenge dose, need to have appropriate statistics, and the assays need to be measuring the appropriate parameters, and they should be validated for the pivotal studies.

Now, as you heard, in 2002 there was a workshop that was held, and at that time we were really just starting to develop the strategy for how to implement the Animal Rule in regards to anthrax vaccines.

And this workshop was very, very valuable at getting a lot of good scientific minds together to evaluate the data that were available at that time, and try to come to a consensus on some very important starting points that could be used to move forward, and I just want to summarize those today.

So the workshop had four sessions.

First session was review of pathogenic mechanisms, second was the review of animal models, third was possible strategies for the development of correlates or surrogates, and then we had a panel discussion, as we'll have two panel discussions in this workshop, and it's during those panel discussions where a lot of the ideas get kicked around and we can hear from not only our panelists but also people in the audience who might have some good thoughts and good ideas.

So in regards to the 2002 workshop, what were some of the consensus points that were reached? As far as the first criterion for the Animal Rule, that you have to understand, really, the pathogenesis of the organism and the host response and how the host is being protected.

The pathogenic mechanisms of B anthracis were reviewed and were thought to be reasonably well understood. And that is that the spores are inhaled, they then are

taken up by cells such as macrophages, the spores then germinate, and the vegetative cells escape from the macrophages and get into the bloodstream, and they secrete anthrax toxin. And it is believed that the -- it's the result of the toxin that you get the manifestations of the disease.

And the toxin is a tripartite toxin composed of protective antigen, and either -- and a lethal factor and edema factor. Protective antigen binds to Eukaryotic cells, oligomorizes, the LF or EF then binds to the PA, it's internalized, the PA when it hits the acidic environment of the endosome forms a pore, allowing entry of either LF or EF. And, again, it's believed that the disease symptoms are caused by the action of this toxin.

So the new generation anthrax vaccines are in general PA-based, with the idea that if you elicit toxin-neutralizing antibodies then that would abolish the effect of the toxin and prevent disease. So at the

2002 workshop it was really felt that there was a sound scientific basis for this.

One of the other things that came out of the 2002 workshop which was very, very important was the choice of the animal species to use in order to mimic the human response. And the animal data from a number of animal species was reviewed, and there was consensus that there were two animal species that would best mimic the human. And the gold standard was thought to be the non-human primate, and sort of the working model where you could get large numbers would be the rabbits.

There was also a discussion about the challenge and what should the challenge dose be, and the consensus was the appropriate challenge dose should be one that might be reasonably expected in an anthrax attack.

And then, finally, we come to the fourth criterion. And at the last workshop people just laid out possible strategies for how to -- what types of studies might help in

producing data that would be useful in fulfilling this criterion, and the consensus was that, really, probably both active and passive immunization studies in animals would provide valuable information that would help fulfill this criterion.

So what are we going to do in today's workshop? What we're going to do is review the overall strategy that has evolved since the 2002 workshop, review the data that have been generated since 2002, and then we'd like to obtain input from the panel members and you, the workshop participants, on how best to move forward.

Okay. Thanks so much, and I'll take any questions.

(Applause.)

DR. NASS: My name is Dr. Meryl Nass. I didn't attend the April 2002 meeting, but I have read the transcript, and I -- there must be something wrong with me, but I certainly don't recall that there was

consensus regarding acceptance of these two animal models as ideal for anthrax in humans, and I sent comments to FDA several years ago during a comment period when this current anthrax vaccine was relicensed, pointing out why these two animal models were not good.

So I just want to point out for the record that I don't believe there is consensus.

DR. LYNN: Okay. Thank you.

Anybody else?

(No response.)

Thank you, Drusilla.

Our next speaker will be Dr. Bob Kohberger.

DR. KOHBERGER: Okay. Thank you, Freyja.

I'm going to talk about some of the statistical considerations in correlates of protection, sort of to set the stage for the next day and a half, from at least a statistical point of view. And the outline of

my talk is, first of all, what is a correlate and what is a surrogate? I think there's some confusion on these terms as to what they mean, and I'd like to get some definitions down.

Second point is: how do we obtain correlates? And then, how do we obtain surrogates? And where do we stand today?

Well, what's a correlate and what's a surrogate? Now, this slide comes from Tom -- it's based on Tom Fleming's publication in Health Affairs. A Level 1 is true clinical efficacy, where we have a clinical endpoint -- survival, whatever your endpoint is.

The second level in the Fleming definition is called a validated surrogate. Vaccines will often refer to these as a surrogate. And this means that the variable, the immune response, explains all of the clinical benefit.

The third level is, in Tom's terms, the non-validated surrogate. It's reasonably likely to protect clinical -- predict clinical

efficacy, and in vaccines we can call those predictive correlates. And the key point here is there is no statistical validation of this, but people -- scientists, experts in the field -- feel that it's reasonably likely to predict benefit.

The fourth and lowest level is just a correlate, and here the immune response is related to the clinical endpoint. We'll call that a correlate. Now, since I made this slide, one of our panelists, Dr. Self, kindly published a paper last week that goes into more detail on this.

So the next slide is not in your package, but it takes what Steve and his colleagues have done and basically -- and, of course, Steve will have a chance to rebut this -- it seems to me that what they're doing is a Fleming Level 2, which is a validated surrogate, they break it down into three more refined levels, because after all when we say it explains all of the clinical benefit, what

does that really mean, and how do you do it?

Well, in this framework -- and there's the reference from JID, and it was just seven days ago -- starting with a Level 1 surrogate of protection -- and this is statistical -- that means that your immune response is predictive of vaccine efficacy within a defined setting, a defined population, a defined use. Usually it comes from a single large trial, and the typical analyses are the Prentice criteria for surrogates, and we're going to talk about those.

A Level 1 surrogate of protection principal -- same definition, it's within a defined setting, it's usually a single large trial. However, the analysis for causality -- and we're going to speak about this a little bit -- are using principal surrogates, which if you're familiar with this it's also known as the Rubin causal model, and Dr. Rubin is also on our panel. So questions about these

kinds of causal models, we have some good people to help you with.

The Level 2 surrogate of protection says that it's predictive of vaccine efficacy in different settings, different populations, different uses, and it comes mainly from multiple trials which means you need a meta-analysis where we would test this in different age groups, we would test it in immune-deficient subjects, and we'd find exactly the same relationship of an immune response to our clinical endpoint. That would be a Level 2.

So from what's -- Dr. Self has done, I think, what it really takes is this validated surrogate and helps us define what all of the clinical benefit really means. And we're going to talk about that a little bit. So we have surrogates, and we have correlates.

Why do we care? First of all, there's a scientific understanding of the process. When you're developing a vaccine, you can't do a Phase 2 trial for efficacy.

You're doing immunogenicity. So it helps in our vaccine development if we know how immunogenicity is related to the clinical endpoint.

If we're wrong on this, it's really the risk for the vaccine developer, because when it goes into the Phase 3 and clinical efficacy is tested the vaccine fails. An example of that is the recent Merck HIV vaccine. We use it to predict vaccine efficacy without an efficacy trial.

Very often efficacy trials are not feasible. In vaccines where you have combination vaccines we can't do efficacy trials. We can't have placebo groups. We need to use a surrogate or a correlate to predict efficacy and get products licensed. These are used for formulation changes, different products as I mentioned, combination vaccines.

If we're wrong, where is the risk? Well, the risk now is with public health,

because products are getting licensed and used. But that's why we care about it.

How do you get a correlate? First of all, we're going to relate an immune response. And it may not be one response, it may be multiple responses. It could be the same response over time, and I think you're going to hear some of that a little bit later.

It could be different responses, such as an acellular pertussis where we have four immune responses that we're measuring. But we want to relate that to some outcome of interest. Generally, it requires paired observations. You need the subject's immune response and the subject's clinical outcome. Some of the examples -- pneumococcal conjugate vaccines, and so on.

I emphasize paired responses in general because for pneumococcal conjugates, for invasive disease, when that product got licensed I believe there were only 20 cases. Just about all of them were in the placebo

group, and none in the vaccine, so the vaccine works.

The problem was we didn't know what the immune responses were for those 20 breakthrough cases. We didn't have paired responses. We did, however, in -- I say "we," because I used to work for Wyeth and was involved with that, so that's why I say "we." There were paired responses for otitis media and paired responses for colonization. So we are able to do correlates in those settings.

How do you choose the immune response? Well, IgG ELISA is often used. The reason is it's easy to use, you can do a lot of observations very quickly, rather inexpensively. So the developers would like to see IgG ELISA used.

There are also functional assays, and we're going to hear about that next. Typically, I think most people would prefer a functional assay because of its name. It measures function of the antibody as opposed

to the ELISA.

Second point -- when should you measure this immune response? You can measure it right after vaccination, or you can measure it prior to challenge. Now, this gets into duration of protection. If you measure right after vaccination -- and this example is from varicella, Merck's varicella vaccine, which is what they did -- and then follow up for two years, you can measure vaccine efficacy as would typically be done in a vaccine efficacy trial.

Sometimes you can't do that, and you measure prior to challenge whether it's an experimental challenge or whether it's like a household contact, and we'll talk a little bit.

How do you choose the event? Well, the type of the event is the clinical endpoint that you're interested in, and that you really want to use to predict for vaccine efficacy. Typically, it's infection, it's a clinical

disease state, it's death. And when it occurs, as I mentioned before, is it over time? Is it a longitudinal follow up? Is it over a two-year period, or is it right after challenge? "Right after" may just -- may be days. So you have to choose your event carefully.

And the consequences are, as I said, when you measure the duration, number one is a typical situation in clinical efficacy. For varicella, you measured it over two years, and that's what vaccine efficacy was. Same thing for pneumo conjugate. This couldn't be done for pertussis, because basically right after vaccination acellular pertussis vaccines had very high immune responses. Efficacy was in the 70 or 80 percent range.

Well, why is that, when the immune response is so high? Well, antibodies decay over time. So in order to get a correlate in pertussis what they needed to do was look at a

challenge type experiment, which was a household contact study where the subject -- the index case basically brought the organism into the household and subjects were there challenged. And luckily in Sweden they had multiple immune measurements on subjects and could come up with a pretty good estimate of what the immune response was just prior to exposure in the household.

So for pertussis you could do it, but it's important to remember what our inference is now for pertussis -- that it measures just prior to exposure. So the consequences of when your immune response is and when your event is are important, and you have to keep that in mind.

Well, how do you choose this relationship? There has been several different approaches. One is the step function, and this in a sense is very similar to protective levels, which we know for tetanus, diphtheria, and a whole host of

others. And basically step function says that below some level the risk for the clinical event is quite low and constant, above some level it's quite high and constant, and that's a step function.

It's a weak model, in that most of us would think that the probability, the risk of an event, is continuous as a function of an immune response. And the step function says -- steps up. It just changes very quickly. So to look at this continuously, logistic regression has been used, and the formula for the logistic model is there.

Since the probability event -- of an event is equal to that formula, it has been used with a single response. It has been used with multiple responses, as in acellular pertussis, as in responses measured over time. That X there can be more than one variable. It can be a whole host of variables.

In addition to logistic regression, time to event models have been used. Cox

proportional hazard models were used with varicella to look at the hazard ratios of the event occurring as a function of the varicella response.

Case control studies have also been used, and in one particular case -- Group B strep -- case control study where they were estimating protective levels. So there is a -- there's quite a few different ways of relating this immune response to the clinical endpoint. The two most popular or the two most that you see most of the time are this step function, which is just a protective level, and logistic regression.

So the results -- what do you get out of these models? Well, as I said, with the step function you get protective levels, their cut-offs. And if you're going to take this approach, you need to look at the sensitivity and specificity at the particular level that you choose. And I always get this a little bit backwards, so I have to refer to

my notes on sensitivity and specificity, if you'll excuse me for a second.

Here we go. Sensitivity is the probability of being greater than the level given that you have the event. Specificity is the probability of being below the level, given that you don't have the event. In diagnostic testing these two are used most often to determine what the level -- what the cut point should be.

There is also something called a positive predictive value, negative predictive value. Positive predictive value is the probability of the event given that you're above the level. It's the reverse of sensitivity. And negative predictive value of probability, you don't have the event given that you're below the level. It's a little confusing. You'll see some examples later.

For the most part, positive predictive values and negatives are not used very often, because in epidemiology the

positive predictive value depends on the incidence rate of the event in the population, not just on the diagnostic test, whereas sensitivity and specificity, because it's conditional on getting the event, it's not dependent on that. So sensitivity and specificity are most often used.

So if you're going to deal with protective levels you need to look at sensitivity and specificity. As I said, some of the examples are in diphtheria, tetanus, polio, Hepatitis B, influenza, meningococcal. They have protective levels.

As I said, we can use continuous functions. These are logistic models, survival models, where the relationship of the immune response to the clinical event is continuous. And some of the examples where this has been done is varicella and pertussis, pneumococcal conjugate and colonization, and in otitis media.

So to summarize, the most simple

are a single response, antibody after vaccination, a single outcome, a disease state whether you have the disease or not, and a logistic model or a protective level. You can get more complex. You can get into time-varying, multi-variate immune responses. You can have time-varying longitudinal data series for the events that are happening, and the relationship model is going to depend on how you set this up.

And at the bottom one of my favorite quotations, which I think is attributed to George Box who is a statistician, is that all models are wrong, but some are more useful than others. So when we pick these models it's for their usefulness, knowing that they are not always completely correct.

So how do we obtain a surrogate? What we were talking about before are just correlates. All they do is correlate things. Well, we're looking for causality, and

correlation does not mean causality. And this quote is from Tom Fleming that "A correlate does not a surrogate make." Just because we found correlation, it doesn't mean that it's a surrogate.

In general, a surrogate explains all of the relationship, and let's talk about how we define this. Did I skip something? I guess not.

Just as a little thought experiment -- and what I mean by "all" -- suppose we have two groups, and they're randomized to vaccine and placebo. We vaccinate them, then we challenge them, and we measure the immune response prior to challenge.

And the results of this experiment are that in the vaccine group 80 percent survive, 10 percent survive in the placebo group, and the immune response in the survivors is 10, and the immune response in those that died was two.

Did the vaccine cause an increase

in survival? The answer is yes, because we've randomized these two groups. Randomization is what gets us to causality.

Did the immune response cause the increase in survival? Well, we have a significant difference in the mean response between those that survived and those that did not. We may have a significant logistic regression here where we can predict immune response and survival.

Is it causal? Did it cause it? Well, maybe. The immune response isn't randomized. It's a post-randomization event. The subjects that got these low responses may be somehow different from the ones that got high responses and that's why they survived.

So from this little thought experiment, the vaccine caused an increase in survival, but without some additional work we can't say that the immune response is a causative factor. So how do we do this? How do we get this causative factor? How do we

obtain a surrogate? Whether it's in Level 1, Level 2, but it's causal.

Four kinds of approaches. There are causal diagrams, and we're going to go into these. There's apprentice criteria, there's the principal surrogates or principal stratification. It's known now I guess as the Rubin causal model. And there's Tom Fleming's hierarchy.

Now, causal diagrams are diagrams that demonstrate the causal effects. And this experiment, which is from Judah Pearl's article, referenced here -- we have a fumigant, and we want to estimate its effect on crop yields. Well, the way the fumigant works is we have to worry about last year's worm population, the worms are eating our crops.

We have to worry about the predator -- the worm predator populations, and the worm populations before, after, and end of the season, the growing season. And this

structural equation model shows how all of these factors interact for the fumigant, the crop yields, and you can get a causality through something called structural equation models.

My opinion is that these kinds of causal diagrams are very good for looking at causal effects, for visualizing them. Using structural equation models are a little harder, and I have difficulty with them, and I -- I haven't seen too many applications in vaccines of these.

What is more popular and used more often are the Prentice criteria. And I mention these -- this would be a Level 1 surrogate of protection statistic in the recent publication. Four criteria -- references for these are down at the bottom.

The treatment impacts on the surrogate endpoints. The vaccine impacts the immune response. The treatment impacts the clinical endpoint. The vaccine increases

survival. Third one is that the surrogate is related to the clinical endpoint in a correlative sense. In other words, immune response is correlated to the clinical endpoint. And the fourth one is that the surrogate contains all of the information about the clinical endpoint. And if you meet these, you've met the Prentice criteria to get a surrogate.

Mathematically, and this is I think about the only little math slide I have in here, the first three are not hard to meet in particular, because they're tests of significance. Is vaccine related to survival? Is immune response related to survival? Is vaccine related to immune response? And that's just a significance test. That's pretty easy.

The last one, does it explain all of the response? That's a lot harder. It's an equivalence test. Basically, what you have to show is that when you have the immune

response in our model against the clinical endpoint, this term in here, which is treatment, and I just show it as one, that coefficient has to be zero. We can't prove anything is exactly equal to zero. It's an equivalence test.

So the thought in statistics is to look at proportion of treatment explained. In other words, if we explain 90, 95 percent, that's pretty close. By the way, this doesn't necessarily have to be just one factor -- vaccine, yes or no. We can look at things like, is the immune response consistent in the vaccine group and in the control group? Is it consistent across age groups? For the Prentice criteria, these would all have to be zero.

Now, I'm not going to talk about the Rubin causal models in the interest of time, and I think since Don is a panelist, if you want to get into that, you can. But I'd like to go into, where do we stand today?

Just to remind you, as we talk through these things, you're going to hear I think mostly about correlates, what's related to survival. We need to see -- we need to remember, three is a predictive correlate where the scientific body of knowledge says that, yes, it's reasonably likely to predict it.

Moving up is Level 2, which is our surrogate, and there can be three kinds of surrogates as in that JID paper, where Level 1 is in a specific application, Level 2 is general, and there's different ways of proving that.

So as we go through the next day and a half, there's a couple questions I'll leave on the table. Has a correlate been obtained? How do we move from a Level 4, which is the correlate, to the Level 3, which is a predictive correlate? What information is needed to show that it's reasonably likely to predict clinical benefit?

I think it may be unlikely that we can get the Level 2 surrogacy, that validated surrogate, but maybe we can. And, if so, what do we need to do?

So to summarize, these correlate models, we need to determine the most useful model for relating the immune response to the clinical outcome. We need to consider surrogates, most likely in the Prentice sense, but a realistic goal is to move from this Level 4, which is just correlation, up to Level 3 where we think that it's reasonably likely to predict clinical benefit.

And I haven't mentioned at all pan-species surrogacy. Is it acceptable to infer from rabbits and non-human primates into humans? And how do we do that?

So I will answer any questions that you have.

(Applause.)

MR. SUTER: Mark Suter is my name. I would like to go to the last point. How do

you actually do that? Maybe you compare -- is there a logical immune response between the rabbit and the human? The human has four IgG, the rabbit has one. The human has IgD, rabbit has none. Human has two IgA, the rabbit has 12.

DR. KOHBERGER: I'm a statistician, not an immunologist.

(Laughter.)

So I'm going to finesse this question, because I don't know. I mean, from a -- you know, from a statistical point of view, I mean, we think pretty -- I don't want to say simplistically, but we'd like to see efficacy trials in humans. But we can't do it. I mean, you know, it's impossible.

So what we need are the immunologists to come to some sort of an agreement that these immune responses that we obtain in non-human primates and rabbits are reasonably likely to predict efficacy in humans, in the face of all that you've said.

MR. SUTER: Thank you.

DR. KOHBERGER: Anything else?

(No response.)

Thank you.

DR. LYNN: I'll just introduce myself again. I'm Freyja Lynn, and what I want to talk about in the next 20 minutes or so is the status of the assay that you'll be hearing a lot about in the next day and a half.

This is the assay that is a likely candidate to be used as a correlate of protection, and I just want to sort of talk about the assay performance, so that in fact we can reassure ourselves that the assay is a good choice, just from an assay standpoint. So I'm not going to get into any of the correlative stuff. I just want to talk about the assay itself.

Before I go any further, I have to admit that I didn't generate any of the data, didn't do much of the data analysis either.

This assay was originally set up in the USAMRIID Laboratory, was transferred to the CDC. The CDC did a tremendous amount of work on standardizing the assay and tech transferring it, and I think some of the data I'll show will show you how successful they were in that effort.

I'll be showing you data from an interlab study. The participants are listed here. Tremendous amount of work from Battelle groups. And, finally, the data analysis was done by Precision Bioassay -- David Lansky is the head of that group -- and his people have done a tremendous job looking at these data for us.

So what do we think about when we think about an assay that's going to be used as a correlate of protection? The toxin neutralization assay that I'll be discussing today actually measures the ability of serum to neutralize the effect of lethal toxin on a cell substrate or a monolayer of cells. And I

kind of broke this into three areas that I tend to think of.

The first is the relevance. Obviously, that's the most important in certain regards, and I think we'll hear a lot in the next day and a half about the relevance of this assay. And just to touch on that briefly, we do know that antibodies to toxin are a mechanism of protection. You'll see -- and, again, I'm not going to present any data on the relevance issues right now.

TNA is attractive, as Bob said, because it measures the functional antibody rather than just all of the antibody that's generated, and you'll see data in the next day and a half that show that it correlates quite well with protection in rabbits and non-human primates.

An assay has to be applicable. You have to be able to use it, and it has to be appropriate to answer the questions that you're asking from an assay performance

standpoint. And I think what I'm going to try to do today is show you that, in fact, the assay is adequately sensitive, dilutionally linear and precise, and that it is a pan-specific assay, or a pan-species assay. And I think this is critical as we move forward.

The question we just got: how do you compare among species that have different antibody subclasses? Well, I think a functional assay that performs the same across species is a good first step.

Finally, the assay has got to be practical. You can't have an assay that takes you two weeks to run a single sample. And, again, I hope to convince you that we have good precision in this assay that allows for a high throughput, and that it actually is quite robust across multiple laboratories.

For those of you who are unfamiliar with what the assay does, essentially all you do is mix lethal toxin together with your serum sample. Antibodies in the serum will

neutralize the toxin. You add that to a cell monolayer, and if there's any free toxin left it will intoxicate the cells and kill them, and you measure the viability of the cells after the intoxication.

The data analysis, for those of you -- just a quick brief for those of you who are not assay geeks like me, just so you have a clue what I'm trying to tell you, you run a series of dilutions of each serum sample, and so you get a titration curve -- you get a titration curve that is just simply plotting the OD, which is the cell viability, against the dilution of the serum.

The data are reduced for each titration curve using a four-parameter logistic fit. The four parameters are the lower asymptote, the upper asymptote, the inflection point, or the ED_{50} as we call it, and the slope, and I'll be talking about those four parameters in a little bit -- a little bit later.

You'll see here that for very high titer samples we get very nice full curves. For lower titer samples we get what we call partial curves. Again, I'll be talking about that a little bit later. The readouts that we're using are the ED_{50}, which again is simply this inflection point, and something we're calling the NF_{50}, which is simply the ED_{50} of the sample divided by the ED_{50} of a reference run on the same plate. We've found that this actually normalizes data between assays and between labs, and I'll show you some data on that in a moment.

DMID has been sponsoring a variety of different studies that have been conducted in a variety of different locations. I'm not going to go into all of them. I just wanted to give you an idea of what we've been working on. What I'm going to talk about today are some of the validation that we've been doing for rabbit and non-human primate. We do also have a human validation underway, but I don't

have data on that right now.

And mostly I'm going to talk about the -- what we'll call the interlaboratory study, so let's get into that. The interlaboratory study was a tremendous effort from a lot of different people, and it involved seven different laboratories.

And what we did was we put together 108 samples that we sent out as a panel blinded to each of the seven different laboratories. We had a mix of rabbit, non-human primate, and human samples, and we asked the laboratories to provide us with two reportable values, so that we could look at precision as well as agreement.

We included in the panel -- we included in the panel low, medium, and high samples. We also did spiked samples where we took high samples, we spiked them into negative serum so we could look at dilutional linearity in each of the three species, and we asked each laboratory to run their own assay.

And I think it's important to note that six out of seven of the laboratories had participated in a common tech transfer sponsored by the CDC, and those data I think are quite interesting.

For analysis, we were interested in essentially three different areas. One was the pan-specific or pan-species performance. Can we really say that the assay performs the same for each of the three species and allow us to make direct comparisons among antibody levels among the three species? And we looked at titration curves, the individual titration curves, and the dilutional linearity for each species.

We also looked, then, at agreement among the laboratories. We have different laboratories doing different assays that ultimately will have to probably be compared, one laboratory doing clinical samples, another laboratory doing animal samples, and so we need to understand how the data from each of

these laboratories are related.

And, finally, we looked at some -- I'm going to show you some precision data, both from the interlaboratory study and from some of the assay validation work.

So right into the data. This is from the interlaboratory study, and we were looking at, again, the similarity of the species. So we're going to talk about the species comparability within the assay. What we have here is we had four different laboratories which provided us with the parameters from the four-parameter logistic curve fit.

We have the lower asymptote, the upper asymptote, and the four-parameter slope. And across the bottom you can see that for each laboratory we have the reference material, which is a human reference, the human samples, the non-human primate, and the rabbit samples. And all we're doing is comparing the parameters for the titrate --

for each of the -- or from all of the titration curves for each species.

And what you can see is that for the lower asymptote and the upper asymptote, for each laboratory each of the three species is essentially behaving exactly the same way. So this provides us evidence that the titration curves themselves among the species are very similar in the assay.

You'll note that within a lab you may see a little bit of difference from lab to lab, but within a lab they are very similar. In particular, I find it interesting that the slope looks very good.

This is a very similar analysis, except it combines all of the data from the four laboratories, and it looks at each of the three species with regards to the human reference material. So if we're going to use like a human reference in all species to incorporate the NF_{50}, which I spoke of earlier, then we need to show that the titration curves

are the same, so that we are legitimate in making that comparison.

And, again, the lower asymptote, the upper asymptote, and the slope, especially for the human and non-human primate, are very tight. The rabbit may be a little bit different here, but it's under 30 percent, and this is a cell-based assay.

I think it's interesting and probably predictable that the human and the non-human primate would be the most similar, and having a little bit of a different rabbit is not unexpected, and we don't feel that this is a huge, huge difference to raise any real concerns.

The other aspect we looked at was the dilutional linearity. Again, we took a sample and we created a series of spikes from that sample, and then measured the ED_{50}. So if you plot the spike versus the ED_{50}, you should get a straight line with a slope of one.

We had, you know, several samples

that we did dilutional linearity for, for each of the three species, and as you can see for the most part the broken line is the ideal, the solid line are the actual data, and for non-human primate and rabbit you can see that they are astonishingly dilutionally linear.

The human may be varying just a little bit. It turns out that this slope is about 1.16. Again, maybe a little steeper than the other two species, but, again, well within what we would expect for a bioassay.

So, essentially, this is just a conclusion stating that we think that essentially the species are performing the same within the assay.

The next thing we looked at were the laboratory-to-laboratory agreement in the interlab study. These are a modified Bland-Altman plot, where the -- each laboratory is compared to a consensus ED_{50} that was calculated by using all of the data from all of the laboratories.

And on the Y-axis, here is exact agreement one to one, so perfect agreement would be a straight line at the one to one. This would be a two-fold difference, a four-fold difference, and these are the 95 percent confidence intervals. So as you can see, when you lose the ED_{50} as a readout, we are seeing some systematic shifts like, for example, between Lab A and Lab C.

I think it's interesting to note Lab D was the one laboratory that did not participate in the tech transfer, and they are the ones that may be just a little more different. But if you start looking at the higher ED_{50} values, they also agree quite well. We lose agreement for the most part at the low end, and that's actually true generally across the board.

And, again, that would be expected. You tend to get your least precise, least accurate values at the lowest ends of the assay. And this is using all reportable

values, so we haven't taken into account a limit of detection or a limit of quantitation.

If you do the same analysis but you use the NF_{50}, again, that's the ratio of the ED_{50} of the sample to the ED_{50} of the reference, you can see how that starts to normalize the data, so that some of these shifts start to level out, everybody kind of comes closer together in terms of agreement. But overall we think the data show that the labs are very, very close.

This is the same kind of a plot. It's just each laboratory compared head to head to every other laboratory in the study. And I'm just including it; I'm not going to go through it. It essentially gives us the same message.

So essentially if you look at the data we had very good agreement among all seven laboratories, especially when you look at ED_{50}s well into the working range of the assay. And I'll show you some data on LOD and

LOQ in a minute. And the six laboratories that participated in this tech transfer had actually quite phenomenal agreement.

And, again, I remind you this is a bioassay. This is not an ELISA. This is a difficult assay to run, but it has been so well characterized and standardized that, in fact, it performs amazingly well.

When we looked at assay precision, again, this is from the interlaboratory study. If we take all of the data from all of the labs, and we say, okay, over everything, between all the species, between all the labs, how precise is this assay, and if you look at the ED_{50} -- and this is a percent relative standard deviation, which is essentially the same as a coefficient of variation, if you're used to seeing CVs -- the total variation is only 45 percent. That's seven laboratories and three species and a bioassay.

And I'll tell you that when we did the validations our criteria in an individual

laboratory was round about 50 percent. So even if we go to seven laboratories we're still under 50 percent.

Here is some nice data about the NF_{50}. If you go to using the NF_{50}, you can see that the laboratory variability goes from 29 percent to 13 percent, which in turn drops the total variability down to 35 percent. And I think at this point, for those who are interested, we are looking into moving forward with an NF_{50} readout, so that we can normalize data and hopefully make data more comparable as we move forward in the project.

That was all laboratories, all species. I just thought you might be interested in seeing just one species in a single laboratory. These are the rabbit validation data from Battelle Biomedical Research Center, who is performing validations on all of these laboratories. This is, again, just ED_{50}. The NF_{50} data are similar. And you can just see the various different components.

One thing to point out here is, as I pointed out in the data slide, we have full curves, partial reactive curves, and non-reactive curves. These are our pretty curves where we get upper and lower asymptotes and everything is really pretty. The partial reactives and the non-reactives rely on the software to do some extrapolation. And, again, these are the lower samples.

And as predicted, your CVs are lower. These are your CVs. The PRSD is lower for full than for your non-reactive. And, again, this is just a reflection of the fact that we're well within the working range of the assay, and you can see how tight -- this is just -- the plate is.

If you go down to the total variability, again, for full and partial reactive, we're running about 25 percent. We get to non-reactive, our lowest values, and our PRSD goes up as expected. But, again, even 37 percent at this low value is quite

good.

So I think just in terms of assay precision, within a laboratory it's actually a lot tighter than we thought it was going to be at about 25 to 30 percent for the ED_{50}. And when we move to multiple laboratories, multiple species, we're still only at 45 percent. And, in fact, we can improve that if we go with the NF_{50} readout.

A little bit about assay sensitivity. Limit of detection, this is calculated looking at the probability of a non-zero ED_{50}. Essentially, if you measure a sample and you get a positive result, a value spit out at you, how confident are you that the next time you assay that sample it will be positive again?

And so this is where we know at about an ED_{50} of 25 percent -- or, I'm sorry, an ED_{50} of 25 we have about a 95 percent confidence that if we measure that again we'll get a positive value. So we know that any ED_{50}

above 25 is truly a positive measurable value. Anything below 25 we have a little bit of question as to whether it's truly negative or truly positive.

The limit of quantitation is the point at which you begin to improve your precision, you get more confidence in the precision of your data, the LOQ -- this is showing both rabbit and non-human primate, where the LOD is the same. This is where we show a little bit different between the rabbit and the non-human primate, the LOQ for the rabbit being 35, for the non-human primate being 45. This is based on the probability of a reactive curve.

Again, we know that the reactive curves are giving us our best data. So this is actually a pretty conservative estimate for the LOQ, and you can see that in fact it is quite good at 35 and 45. The other point is that the rabbit and non-human primate end up with the same LOD and very similar LOQs, which

again is evidence that this assay is measuring very similar things in the two species.

So, in summary, I think our data to date suggests that we don't have any really important differences among the species, and that this is in fact a pan-species assay. We've looked at the performance of the neutralization curves.

Again, one of the issues with the neutralization curve is if you are seeing very different mechanisms, if the avidities were very different among the species, if there was truly differences in character of the antibody, you'd start to see that reflected in the titration curves. And we're not seeing a big difference.

The species are performing the same with regard to dilutional linearity, precision, and the LOD, and the LOQ. I think the other thing is that the laboratories, especially when they cross-calibrated, are performing very similarly and reporting

results almost identically. And I think that bodes well for the use of this assay moving forward.

And with that, I'll take any questions.

(Applause.)

MS. VOLKMANN: Ariane Volkmann. You have shown -- I have two questions, actually. You have shown that the standard deviation is much smaller when you look at the rabbit assay as compared to the total where you have the three species. Is that because it was performed in one lab, or is that because there was a real difference between the species?

DR. LYNN: I think that's mostly because it was performed in one laboratory. I'd have to go back and pull up the slide again. I think it's mostly because it was performed in one laboratory. When we finish the non-human primate, we'll know the precision for the non-human primate versus the

precision for the rabbit, and right now it looks like those precision values will be very similar.

MS. VOLKMANN: So if you look at the ED_{50} and compare the species, they have the same values?

DR. LYNN: I don't understand what you mean by "the same values." Each animal has its own ED_{50} value, but when you --

MS. VOLKMANN: Okay.

DR. LYNN: -- repeatedly measure a sample, the variability that you get among your measurements is the same whether you're measuring a rabbit sample or a non-human primate sample.

MS. VOLKMANN: Okay. Yes, I'm asking because of the comparison.

DR. LYNN: Right. Right.

MS. VOLKMANN: And the second question is: does that functional assay correlate well with the ELISA? Because when you measure by ELISA, because it's easier, you

always measure the neutralizing antibodies as well. So if it correlates well, I don't quite see why you do have to do the functional assay, because you know that you have functional antibodies detected in your ELISA.

DR. LYNN: Exactly. And that's a very good point. And, in fact, these two assays do correlate quite well. The problem that's coming to light at this point as we get more data is that depending on when you measure the immune response, whether you measure post first vaccination or post second vaccination, early in the response or late responses, the correlation between those two assays changes.

And so you can't -- you can't universally say that the assays correlate all this time the same way. That correlation changes. And so to me that means that you're going to get a slightly different answer in different studies depending on which assay you use. And my bias is to go with the functional

assay where we can, because we think that is probably the more relevant measurement.

MS. VOLKMANN: Okay. Thanks.

DR. CHAWLA: I'm Anil Chawla from Panacea Biotec Limited. In slide 8, when you were showing the similarity of species with the laboratories, you have only used eight laboratories -- four laboratories, data from four laboratories. Why is that so?

DR. LYNN: That was -- that's simply convenience. Those four laboratories were actually the only ones that reported the four parameters directly, so it was very easy to extract for those four laboratories those data. We can go back and get those data, but for this analysis it wasn't worth going back to all seven laboratories. So it was purely convenience. We had the values for those four labs.

DR. CHAWLA: My second question is related to MTT dye. There are issues in using the MTT dyes. So there are better dyes,

water-soluble dyes, which are available now. Are you having any move to move toward that -- those kind of dyes?

DR. LYNN: There are some laboratories that are working on developing the assay further, and one of the aspects is using a different dye. And, in fact, one of the laboratories in this study does use a slightly different MTT method, and their data came out looking essentially exactly the same. So I think we could -- we could do that kind of improvement.

DR. FERRIERI: Pat Ferrieri. I wanted to be sure that I understood that the assay done and reflected in your data was based on the PIT publication in terms of the precise doses of recombinant PA and lethal factor. Is that the case or not?

DR. LYNN: I would have to go back and look at that, but the doses of toxin are very similar among all of the assays. In fact, most of the laboratories that are

running the assays are using the same material from List, and CDC is actually the one taking the assay from RID and did some more fine-tuning in terms of selecting the optimum doses, but, yes, they are essentially the same.

DR. FERRIERI: So it is the rPA.

DR. LYNN: Yes. Oh, absolutely, yes, it is rPA.

DR. BURNS: Drusilla Burns. I just want to get back to the question about can you use ELISA instead of the TNA. And I think one of the big problems in using ELISA is ELISA uses species-specific reagents in order to develop it. So I'm not sure that an ELISA titer from one species could be directly translated to the other. The beauty of the TNA is there is no species-specific reagents.

DR. LYNN: Yes, Conrad.

DR. QUINN: Conrad Quinn, CDC. I apologize, I'm losing my voice, so I'll be squeaking later this afternoon. To address

Dr. Ferrieri's question, the assay that we're talking about here was technology transferred from the CDC. It was based on publications by Steve Little and Art Friedlander.

The amount of toxin is titrated to give 95 percent cell death, so we're actually building a model around 100 percent survival and 95 percent cell lysis. So it's titrated to give those values.

MS. BELLE: I'm Archana Belle from Planet Biotechnology. With regards to species-specific, I had two questions, one is with regard to the ELISA and species specificity. We, of course, now allow us to do this without having a second detection agent, so we can bypass the issue of species-specific reagents.

A second question is this TNA does not look at the clearance mechanism of the toxin/anti-toxin complex. Have you any thoughts about differences in species? And are animal models still correlative to humans?

Or any thought about, how do we look at that as a big picture?

DR. LYNN: We've not done any looking at the clearance mechanisms, and I know that there are several -- there are several papers out there looking at the clearance and how the toxin is cleared at this point in time. But no, we haven't -- we haven't looked at that.

MR. KAMMANADIMINTI: Srinivas from Cangene. I would like to know what was the reference standard used for NF_{50} calculation.

DR. LYNN: AVR-801.

MR. KAMMANADIMINTI: 801.

DR. LYNN: Yes. That is -- that's the reference that was developed by the CDC. It is available through the BEI -- NIAID BEI program, and ultimately I think that's probably going to be our gold standard for our work.

DR. QUINN: Conrad Quinn, CDC, again. Regarding ELISA and TNA, or ELISA

versus TNA, I think in our perspective it is important to use both, because, yes, the TNA does measure neutralizing antibodies, but in relevance to clearance antibodies that bang protective antigen are still biologically active. Although they may not neutralize the toxin, they are still part of the clearance process and complex formation, so they should not be excluded from our analysis. So I would suggest that ELISA and TNA are both important.

DR. LYNN: Yes, I would agree.

Okay. We are, amazingly enough, running 15 minutes ahead. So let's go ahead and take our break, and we'll see you back here at the appointed time.

Thank you.

(Whereupon, the proceedings in the foregoing matter went off the record at 9:47 a.m. and went back on the record at 10:26 a.m.)

DR. HEWITT: Okay. We're going to Open Session Number 2 on Animal Models for

General Use Prophylaxis. Our first speaker from DMID is Ed Nuzum, and he's going to talk about the Rabbit Challenge Model: Interpretation and Implementation of the Animal Rule.

DR. NUZUM: Thanks, Judy. And good morning, everyone.

So in my talk today I'd like to just kind of reflect on some of our experiences and how those experiences impact our interpretation and implementation of the Animal Rule with regard to the rabbit anthrax aerosol challenge model.

PARTICIPANT: We can't hear you.

DR. NUZUM: Okay. Is that better? So I'm going to talk about our experiences with regard to implementation of the Animal Rule. As most of you probably know, this is a critical part of our rPA anthrax vaccine development program, so it has been something that we have given a lot of emphasis to.

So this fairly simple cartoon was

made by Scott Winram a few years ago, and it certainly oversimplifies the complexity of everything we're trying to do. But most of you are I'm sure familiar with non-clinical studies and clinical studies that are conducted in support of vaccine development.

There is a couple of pieces I think we tend to overlook or not get enough attention to early enough, the first one being the product itself, the countermeasure that you're working on. And it's important with regard to the animal model because once the natural history studies are done, you have to have a product to put into the model to continue development.

And there's a concept that we try to talk about of quality and maturity of both the model and the product. Early on in the product development stage we think it's fine to have product that -- well, you're not going to have product that's high quality product, and the model itself will also be immature.

But the idea would be that as the development path -- as you go down the development path, and the product matures, the model matures, such that when it's time for pivotal studies that are IND- or BLA-enabling studies, the model would be much better characterized.

The other piece is assays, and Freyja just gave a nice talk on what's involved with that and how complex that whole piece is. And it's absolutely essential, because that's the piece that ties all of the data from multiple species and multiple labs, and it's very complex. We're very fortunate to have Freyja help us run those trap lines.

Our approach is really very straightforward for the GUP indication. We do active immunization with increasing doses of vaccine. And as we evaluate the vaccine, dose-dependent immune response with regard to protection, then determine an immune response at the time of aerosol challenge, and then

concurrently we conduct clinical studies and then we evaluate the protective titers in animals with regard to the immune response in humans.

When we first started these studies, or this entire effort, we came up with several areas -- or folks -- those of us in the government, contractors, everyone came up with several areas, large focus areas that we thought would be needed to be looked at during the development cycle for the product.

And we probably -- well, we have the first few years concentrated on these first three bullets. We're currently moving our focus into passive immunization and time to protection, and ultimately we'll do high-dose challenge studies and duration of protection, so that -- we'll do those studies when we have a more final drug product that's consistency lot material, it's GLP product made at full scale.

I want to talk about several

assumptions, and on the surface they seem very straightforward. And those of us that work with this every day, they have become kind of second nature. But I want to emphasize that these assumptions have either came out of public workshops or considerable internal debate, and there is a lot that has been -- that has happened and discussed behind the scenes that go into these assumptions.

But they have been absolutely critical from the standpoint of starting the studies, and then also for advancing and making progress with the model as we get additional data and perform additional studies.

So the first assumption is that rabbit and non-human primate are relevant model species for prediction of protective behavior of anthrax vaccines in humans, that their protective correlates developed in non-human species will be protective -- predictive of protection for humans, and that the

clinical benefit provided by countermeasures and relevant models provides confidence that similar effect of countermeasure in humans will be predictive of clinical benefit in humans.

So that's kind of a mouthful. It seems rather circuitous, but the key piece here is the relevance of the animal models. And I think that's really the -- I want to emphasize that, because I think that's the basis of the Animal Rule. If the animal responses, immune responses, pharmacokinetic responses, the pathophysiology, are similar to what you see in man, then that enables you to make this extrapolation from animal efficacy to human immunogenicity.

Anti-PA antibody mediated neutralization of anthrax toxin is an acceptable correlate because it can be used to reasonably predict clinical benefit, and it is associated with the prevention of known pathological -- pathophysiological mechanisms.

Now, Drusilla touched on these same points, and I guess what I want to emphasize here is that the terminology, you know, regarding clinical benefit, pathophysiological mechanisms, comes straight from the Animal Rule. And what I want to emphasize, we do use the Animal Rule as our guiding principle in this model development, and we're actually implementing it. And I hope that this workshop will show that is in fact what we are doing.

Circulating antibodies to PA at the time of animal challenge are an appropriate predictor of protection. This is a point -- and Bob alluded to this in his talk -- there's -- this is our assumption. There's other ways to do this with regard to timing of when you measure the immune response and the event, and I think CDC will talk about this a little bit. But another option is to look at peak responses soon after vaccination, and then look at how that correlates with protection

when challenged at some point in the future.

I don't -- it's not that either is right or wrong. It's just -- I'm just pointing out there's different approaches, and this is the assumption we're using for our models.

Next, there is an antibody threshold above which protection is conventionally adequate, and the antibody threshold is the same in animals and humans. Again, conceptually, on the surface this sounds very simple, but implementation, doing the studies, getting the data to demonstrate this is not so straightforward.

The Animal Rule does not specifically require correlates or surrogates of protection. I mention -- I make this point, because depending on the complexity of the model, the endpoints you're looking at, it simply may not be feasible to obtain actual correlates. So I think that's something we need to keep in mind. But that said, the

correlates provide us a very powerful tool.

And if we can get a correlate, we need to do it, and that's the reason this discussion will -- this workshop that will focus a great deal -- well, it is the focus of the workshop. So it's just -- but it's a point that I think we need to be -- to keep in mind what the Animal Rule really says and what it doesn't say, and then keep that perspective.

So the assumption, then, is that -- here is that to the extent that correlates of protection are feasible, attainable, and facilitate implementation of the Animal Rule, every effort is made to develop meaningful correlates.

The Animal Rule requires that the effect is demonstrated in more than one animal species. However, it does not require that one non-human species is predictive of another, or that multiple non-human species are comparable. Again, this is something we

have debated, and it was -- this point was really kind of brought to my attention by Bob Kohberger in that establishment of the correlate will be much easier to do if multiple species are comparable, and if they are predictable. And we need to do it or make every effort to show that they are.

So, again, what the Animal Rule says, what it absolutely requires, and what we need to do or try to do maybe a little bit different, and the assumption here, then, is that the demonstration of comparability between non-human species is highly desirable and will be attempted but is not essential for product licensure.

The Animal Rule does not require 100 percent lethal animal models to the extent that human lethality is not 100 percent. Again, we full appreciate the value of models where all controls die, and we make every effort to develop models where that's the case. But just keep in mind it's not a

specific requirement, so, again, I'm trying to give the perspective here of what's desirable, what's nice to have, what's doable, and what you have to have.

So the assumption on this slide is the demonstration of fully lethal animal models will be attempted, but is not essential for product licensure.

The Animal Rule requires adequate and well controlled animal studies. It does not require validated animal models and systems. Again, this is -- this is kind of second nature to most of us working in this now, but it -- in the early stages this was a subject of considerable debate.

And what we've concluded and our approach is, that the -- and the assumption here is that components of an animal model system that can be validated will be validated, and models will be developed to generate data produced from adequate and well controlled animal studies.

So this is usually where the question will come up, "Well, what is good enough? When is it adequately developed, and how do I get there?" And unfortunately there is not a simple answer, and it's not going to be given up front. It will only come with data and numerous studies and numerous discussion -- or lots of discussion and analysis.

This slide kind of follows on to the same point. It's taken from a presentation given by CBER in January of 2006. And what I want to point out here, really, is that studies must be reproducible and predictive for infected negative controls.

I'm not going to talk about all these points here. They basically summarize the kind of things we would normally want to do in the conduct of good science.

I would mention that pivotal or definitive studies must be GLP, so they implication there is that studies prior to

pivotal studies do not need to be GLP. However, as the model matures, as the product matures, you want to get the increased quality and rigor in your studies, you will incorporate GLP studies probably well before pivotal or definitive studies.

However, for initial proof of concept studies you probably wouldn't want them to be GLP. In fact, I would argue they probably shouldn't be GLP from a resource perspective.

Now, as we've -- as we've gotten more data, we've done more studies, you know, other issues crop up, and so I've kind of lumped these under other considerations. You know, there was a question this morning on Ig subclasses, but antibody functionality may be affected by vaccine regimen and/or time since last vaccination.

Antibody from active immunization may be different than antibody that's passively transferred. Purified IgG may be

different than unpurified IgG that's passively transferred in plasma. There may be differences between similar vaccines. There may be differences between species. So this is a new area, a relatively new area that we've thought about and we're beginning to explore.

Correlate endpoint levels are generated from active and passive -- immunization studies may be different, and we'll see some data later in the workshop that makes this point.

Initial development of correlates for rPA vaccine will be for the GUP label indication, and -- but with an emphasis on time to protection. And this is probably a difference between HHS and DoD where DoD is probably more concerned about duration of immunity, HHS is more concerned about time to protection with regard to post-event scenarios.

I have several conclusions here.

Multiple studies are required in each species for each indication. BSL3 aerosol GLP studies are complex and costly, and, in fact, non-GLP studies are complex and costly in this environment. So there is -- the overall -- because of this complexity and cost the overall model development plan strategy needs to be well thought out.

From the standpoint of cost, staffing, facilities, animal utilization, especially non-human primates, it requires careful thought and planning -- a long-term plan to the extent that that's possible that the right studies are done in the right sequence and that we minimize redundancy.

And, again, the assumptions are important. Because of the cost of these studies, the assumptions we make help us move forward without having to address every possible question. Perfect solutions to specific issues are rare. I kind of modified Bob's quote when I did this, but perfect

solutions to specific issues are rare, but good planning, science, and data are essential to address them in the best manner possible.

And these are all kind of related. I mean, there's a theme here you can tell. But it's not -- neither is it feasible to appropriately address all possible questions. It's much easier to ask questions than to do the studies to answer them.

So when these questions come up, an I'm sure there will be many in this workshop, we have to ask ourselves constantly, "Do we really need that answer? If we have that answer, how will it help us?" What are the possible outcomes of the study that could be -- including the negative outcomes? And will the study really be worthwhile?"

And I guess, finally, when we're thinking about questions, we have to ask, "Is it attainable?" You know, it may be a very interesting question, it might be potentially very useful, but it simply may not be

attainable or not attainable in a practical and feasible manner.

And, finally, one of the main points that I like to make is that animal model and assay development consist of iterative process development studies and data-driven decisions that guide FDA, funding agency, and product sponsor decision-making.

The other statement that I guess I wanted to conclude with -- and it's not on here -- but if I can make one statement that the -- that a very high level, practical way to capture all of this, I would say that unless the animal model is very well developed it's unrealistic to expect that one study is going to address adequately any specific question or issue that has been raised.

So with that, I will conclude. I'm happy to take any questions.

Thank you.

(Applause.)

DR. CHAWLA: Hi. I'm Anil Chawla

from Panacea Biotec. What is the scientific basis of claiming that antibody threshold is same in animals and humans? Because of difference, you could drive the animal threshold to -- you can correlate really, but they cannot be same. What do you say about that?

DR. NUZUM: They cannot be the same in what regard?

DR. CHAWLA: The value. They can be same in terms of value?

DR. NUZUM: Well, if certain levels are protective in animals -- so I think your question is: how do you know what the level that's protective in animals is going to be protective in humans?

DR. CHAWLA: Exactly.

DR. NUZUM: Right. So that's where I think that the Animal Rule is -- we have to rely on the Animal Rule, and the requirements for animal models that are relevant and well characterized.

I mean, at some point it's a leap of faith, it's a prediction. But the data -- the efficacy data that you see in animals, if you get a similar endpoint, a similar threshold level, whatever -- and there's going to be more talk on this, and maybe this point will come out better, because we're going to have clinical and efficacy data. But at some point you make a leap of faith based on your animal efficacy data that those same endpoints in humans will be predictive of clinical benefit.

DR. CHAWLA: My second question is related to the multiple studies that are required in each species for each indication. Can you --

DR. NUZUM: Well, the one slide I listed with regimen studies -- I mean, you do your initial proof of concept, your regimen, the different studies for determining the correlate, time to protection. Those are what I'm referring to.

DR. CHAWLA: Can they be included in one study or -- because one study can have multiple arms or multiple outcomes?

DR. NUZUM: In our study designs, we try to get as much information as we can, you know, address more than one question if we can. The danger with that, of course, is that the studies can become too large, too complex, and that creates a level of risk in itself, so it's a balance.

I would say the short answer to your question is yes, but we do it with caution.

DR. FERRIERI: Ed, one of the leaps of faith for me is the assumption that the spore challenge in the animals is similar to what might be an anticipated exposure in real life. And I wonder if you might comment on that, because in the various papers I've scrutinized some of the spore doses have varied a lot in different experiments within some of the published papers, and I guess if I

were in the Metro, in the front somewhere where the ventilation is going to deliver, I don't know, tens of thousands of spores, a million, or if you're farthest away maybe none or 50 to 100 spores, could you reflect on that briefly, please?

DR. NUZUM: I'll try. Well, a couple of things. First of all, as we've -- we've done enough studies now where there is quite a range in spore challenge dose in our different studies. And we have done analyses on more than one study, and our general conclusion is that the effect, at least in animal models, is that the challenge dose does not correlate, does not impact the response.

With regard to what people might actually be exposed to, that's the reason for -- one reason for the high-dose challenge study we'll do down the road. And in our -- our understanding is that there is -- it hasn't been stated anywhere that the vaccine has to protect at 1,000 LD_{50}s or 2,000 LD_{50}s.

We're using a target of 200 to 400. We consider that reasonable and practical and it's feasible.

But we will do this high-dose challenge study just to have the information. What happens if there is high-dose exposure?

DR. FERRIERI: Thank you very much.

DR. NASS: I guess my question has to do with how you establish the LD_{50} when it is species -- anthrax strain-specific, species-specific, animal strain -- you know, there are many factors.

DR. NUZUM: Well, that's a good question, and, I mean, the thing you didn't mention is just -- well, or maybe what you were implying is there is a lot of biological variability associated with different species, with the challenge, what the assays measured -- used to measure the actual or calculated and held dose and all of that.

But basically it's done by challenging at different doses -- you know, a

challenge dose response, as any LD_{50} study would be done. And you look at the death at low doses going through the high doses.

DR. NASS: I guess what I'm saying is if you go back and look at different LD_{50}s for different species of animal and different strains of anthrax, you'll find very widely varying numbers. And although the number of 8- to 10,000 for a human has been batted around, there is really very little evidence for that. So how are you calculating your LD_{50}, and what's the reliability?

DR. NUZUM: Well, the calculation is no different than LD_{50} has ever done. I think the point we should make here is that the other aspect of what you're saying, they haven't been -- LD_{50}s haven't been done that many times. You can't do LD_{50} studies and NHPs over and over to get a lot of confidence that you have the right number.

I think rather than concentrate on LD_{50} value itself, we need to -- we need to

talk -- and this is another internal debate we've had, there has been discussion of doing away with the reference to LD_{50} period for some of the reasons that you state, and just give the challenge dose in terms of number of spores.

It's a -- I'm not sure what -- referring to this in terms of LD_{50} numbers adds that much value. It has historically been done. The main point is that these animals get lots of spores, and that's -- and it protects against them.

MR. SUTER: You said that a certain serological titer between the different species can be correlated to protection. Is there a correlation between the different titers on the LD_{50} between the different species? So say you have a rabbit of maybe five kilos. You calculate it to a human, and then you say, "This titer and the LD_{50} correlates what you know from human exposure to the bacteria."

DR. NUZUM: I'm not sure I understand your question. The ED_{50}s -- I'm sorry. I didn't understand.

MR. SUTER: If you normalize the serological titer, you should also be able to predict what this would mean in terms of overcoming a challenge. That is, if you have a titer of X in a rabbit and you say you have an LD_{50}, which this rabbit can support, can you then extrapolate what the dosage is you can tolerate in a human?

DR. NUZUM: The challenge dose?

MR. SUTER: Yes. I mean, you probably know in some of these bad cases how much spores were around, and we can probably extrapolate how many bacteria they had to fight against. So is there a certain correlation between --

DR. NUZUM: Well, I don't -- I don't think it's -- I think it's very difficult to make a direct extrapolation, because we don't -- we don't know the lethal

doses to that extent in humans. And, again, that's the reason for giving rabbits lots of spores, and at some point we will have more information on very high challenge doses. And that information I think would give us confidence that that information would extrapolate to protection in humans. I'm not sure I've answered it.

MS. VOLKMANN: I have a comment to the first question, which is that you're assuming that the threshold of titer for protection is the same in all species. And when we look at most titers, for instance, in small pox using a well characterized vaccine such as Dryvax, we always get much higher titers in mice than we would get in monkeys, and yet those titers, although they are different, they are protective.

So what do you think about rather than comparing titers directly between species using a well characterized vaccine as available for anthrax as well as a comparison

and always run that comparison for all assays for all species when you have I guess a better comparison, don't you think?

DR. NUZUM: Do you mean the same -- a same vaccine as the control --

MS. VOLKMANN: Yes, like a gold standard or well characterized or licensed vaccine as a comparison in all your assays.

DR. NUZUM: In many of our studies we do include BioThrax.

MS. VOLKMANN: Because then you don't have to rely on the same titer in a mouse or rabbit or a human.

DR. NUZUM: Well, it's another reference point, yes. And we do include BioThrax in many of our studies.

DR. NASS: But the obvious problem -- Meryl Nass -- doing that with BioThrax is that the different lots of BioThrax contain different amounts of PA and other proteins and have not been individually characterized. So there really is no gold standard --

DR. NUZUM: Well, we're --

DR. NASS: -- using BioThrax.

DR. NUZUM: Right. Well, we're aware of that, but it is licensed and it provides a reference. And, certainly, for the toxin neutralization assays, ELISA capitalizes -- it's focused on PA. It is in consideration.

DR. HEWITT: Okay. Thank you.

I would like to remind all the questioners to please identify yourselves when you're posing your question.

And our next talk is going to be by John Bigger from Battelle, and he's going to present his data on the rabbit model.

DR. BIGGER: Thank you. I guess the platform can be raised, can't it? Yes, for the altitudely-challenged.

Good morning. Can you hear me in the back if I just leave the microphone right here? Are we okay? Okay, great.

(Laughter.)

Barely? Okay.

I'd like to thank the organizers for allowing me the opportunity to come here today and share this data. We received the task to provide small animal models for bacillus anthracis vaccine testing using rPA vaccine candidates. And then, having established that model, we were then tasked to test two rPA vaccines that contained alhydrogel adjuvant for efficacy within this rabbit aerosol challenge model, and then evaluate the immune response to the vaccines.

The test articles themselves were two rPA vaccines in alhydrogel. They were provided by two separate commercial companies. The route of vaccination was intramuscular. We used New Zealand white rabbits, both male and female, balanced set, and then we challenged them with a target dose of 200 LD_{50}, which comes out to be about 2.1 times 10^7 colony-forming units. That was our target inhaled aerosol challenge dose.

And then, for this study -- the studies that I'm going to present, the endpoints were limited strictly to antibody titer and to survival.

So here is our study design. We had six groups of rabbits, ten rabbits per group, and we vaccinated them with diluted doses of rPA, depending upon the cohort, at week 0 and week 4. We then collected serum from these animals every other week for 10 weeks for ELISA and TNAs, and then at week 10 we then provided an aerosol challenge.

Logistically, we could not challenge 60 animals in a single day, so we spread the challenge out over three days. So the animals were randomized both by cohort and by challenge order and by challenge day. We then monitored the animals for 14 days, and at the end we did the whole thing over again for the second vaccine.

So let's look at the immune response. We're looking at vaccine A and

vaccine B for ELISA or TNA, and what we see is that after the first vaccination at week 0 we did get a dose-dependent immune response, and here at week 4 they received a second vaccination which then boosted the immune response in all vaccinated cohorts.

The unvaccinated controls show their ELISA data right here. The ELISA, the antibody titer, peaked at week 6, and then began to decline out here at week 10, which of course is our challenge date. Importantly, the immune response between both vaccines looked very similar, and again, importantly, down here in the week 10 ELISA or TNA titers we also see a dose-dependent immune response by TNA.

I'd point out that the two -- I do not have the same data on the TNA. Over here I'm representing the TNA ED_{50}. Over here I'm representing the TNA NF_{50}, which as you saw earlier the NF_{50} is normalized to a control serum. But as you can see, again, we get a

peak at week 6 and a dose-dependent response at week 10.

So in deference to some questions that were asked earlier, here is a correlation experiment or analysis looking at ELISA versus TNA results, and what we see in each of these slides is -- in the darker black is a week 6 correlation between ELISA and TNA, and then in the lighter line, lighter colored line we have the week 10 correlation.

Now, keeping in mind the fact that we had -- we did not use as many animal cohorts in the week 6 analysis we still feel confident that the week 6 correlation is a little lower than the week 10 correlation in both vaccines -- again, demonstrating another comment that was made earlier that we do have some evidence that there is a change in the relationship between the ELISA titers and the TNA titers as time progresses in this model.

So a moment to discuss our aerosol challenge. We start with a well characterized

challenge material spore lot -- bacillus anthrax, Ames strain -- characterized for 10 separate areas, including purity, genotype, and phenotype characteristics, virulence, and aerosol performance.

Having done that, then, we use a muzzle-only exposure chamber where the animal is loaded into a real-time plethysmography chamber. Plethysmography, if you're not familiar, is a way to measure the real-time inhalation volume and inhalation rate, and then by comparing that with the predetermined spray factors of the aerosol system we're able to estimate as the challenge is going on how many spores the animal is inhaling.

When we reach the targeted challenge dose, we turn the aerosol challenge off, and then bring the animal out. During the challenge, the aerosol chamber itself is sampled by glass impingement. And then, taking the impinged sample and enumerating the spores by spread plate, we can then back

calculate the actual number of spores that the animal actually inhales during the challenge.

So this is the -- a little analysis of our aerosol challenge across both of the experiments. Here we have three days of challenge for vaccine A, three days of challenge for vaccine B, and the take-home point here is that the mean challenge across all six days of challenge across both experiments was very close, very tight within the day and very tight across all six days.

Over here we can see the first experiment graphically. Each one of these numbers represents a rabbit, and then each number on the Y-axis represents the challenge dose that that animal received. So here is challenge day 1, challenge day 2, and challenge day 3 of the rabbits in order. And the important point here is there is no real pattern to how the animals -- to the challenge that the animals receive.

While we believe this is a very

tight and very reproducible pattern, we do recognize that there is a range of doses -- for instance, on day 1 -- from about two times 10^7 to four and a half times 10^7. So there is a range of challenge doses that the animals receive, so we did ask the question: did the challenge dose affect survival?

And this was statistically analyzed by our Stats Department, and without going into it the short answer is no, challenge dose did not play a role in survivorship. As long as the animals received a challenge, then the amount of challenge did not affect the endpoint.

Well, as you note here -- okay, you don't note here, that's okay -- we had ELISA titers across a full spectrum. And as it turns out, above 30 micrograms per ml ELISA all of the animals survived challenge. Similarly, if the animal had an ELISA titer below seven, they succumbed to challenge.

So we actually have a range where

animals both lived and died, and so we asked the question, okay, if you had a 30 microgram per ml ELISA titer, but you were given a lower challenge dose, did that affect your ability to survive? Maybe you could survive that challenge, where if you had received a high challenge dose you would have succumbed. And, again, the short answer is no, that did not play a role.

Okay. So let's look at the survival data. Here we have, again, vaccine A and vaccine B, 60 animals per experiment, 10 animals per group. Here we have the dark line showing you the survivorship of the unvaccinated or the mock vaccinated controls. The control animals began dying at day 2, and we had 100 percent fatality in both experiments by day 5.

Contra-wise, the animals that received the highest vaccine groups either had in experiment B no fatalities or in experiment A we had one fatality out at day 11. The

vaccinations in between showed a very nice dose-dependent survivorship.

And just as importantly, while we received fatalities in the low vaccine groups that statistically were not different than the controls, we did see a dose-dependent time to death change showing that while we -- at the endpoint we still had a statistically insignificant survivorship, we did have a statistically significant protection offered by time to death.

Okay. So we then took a look at the immune response of both of these vaccines, and we compared their ELISA titers and the TNA titers to survivorship in these experiments, and we found that in both experiments there was a strong correlation between immune response and survivorship. And then, comparing those statistically we saw no difference between the two experiments, so the data were combined and then the immune titers were compared to survival and provided in

these logistic regression plots.

So each plot represents 100 animals that were vaccinated with rPA. The control animals are not represented here. So we have animals that had higher immune responses on the right-hand side on each of these scales. Here is your ELISA down at the bottom, TNA ED_{50} here, and TNA NF_{50} here.

Animals that received a lower -- had a lower immune response to vaccination are here, and then by comparing these to the actual survival data we were then able to show a probability of survival as shown by the dark black line with the 95 percent confidence intervals shown in the dotted gray lines on either side. And then, the actual survivorship and immune response data are binned and pointed out in these black dots here.

So, importantly, each of these plots show a statistic correlation between the immune response and survivorship to the P

equals lots of 0001, .0001 level. All of the plots show a very similar curve, and because we're able to do this with 100 animals the 95 percent confidence intervals are very tight in each of these curves. Especially in the mid-range of the curve, the 95 percent confidence intervals are very tight, whereas you get up toward the asymptotes and you've got a little more room here, a little more uncertainty.

So if we then take these data and put them in tablature format, we can provide an estimate of probability of survival at any given ELISA or TNA measurement. Looking at the 75 percent level of -- or probability of survival, we find 25 micrograms per ml ELISA, a TNA ED_{50} of 131, or a TNA NF_{50} of .12. So if you had a TNA ED_{50} titer of 131, and then the animal was challenged in our system, the animal would have a 75 percent probability of surviving.

Again, down in the 95 percent level, we have an ELISA titer of 71 micrograms

per ml, or a TNA ED_{50} of 951, TNA NF_{50} of .35.

In our hands, however, during this experiment the animals that had above an ELISA titer of 29 all survived. So we had 100 percent survival in this experiment at an ELISA titer of 29, as compared to the NF_{50} of 72 basically.

Now, the confidence interval, the 95 percent confidence intervals are shown in the parentheses, and my statisticians assure me that if we had an infinite number of animals that this value would come up, and we would indeed see only 95 percent survival given that budgetary non-restraint.

So at this time, hopefully I've convinced you that the rPA vaccines used here did provide a dose-dependent immune response. Our ELISA titers and TNA titers correlated highly, and we showed a change in correlation over time. Our aerosol challenge model is well characterized and reproducible. The survival of the animals was not challenge dose

dependent, but was vaccine dose dependent, and the survival of the animals correlated to pre-challenge -- week 10 pre-challenge serological titers.

I'd like to acknowledge Dr. Roy Barnewall who is our aerobiologist and supervised all of the aerosol challenges. Our lead statistician on this was Dr. Greg Stark, who is here with us today, if we have any questions on the statistical analysis.

Many of the Battelle staff that were -- or maybe not many, but some of the Battelle staff that are involved in the spore growth and analyses and the TNA and ELISA experiments are here with us today. Again, if you have any technical questions that I can't address, they are here, and I'd like to point out their extremely hard work in this endeavor, and also the animal studies group.

Thank you very much. At this point, I'll open it up for questions.

(Applause.)

DR. NASS: Meryl Nass. You didn't specify what type of rabbit, but I'm assuming these are all genetically identical.

DR. BIGGER: The rabbits were New Zealand white rabbits, and they are an outbred population. So they are not genetically identical. They are an outbred population.

DR. NASS: Did you try this experiment with any more genetically diverse group of rabbits?

DR. BIGGER: The short answer is no, and while I'm limited and could open up the audience on my knowledge of the model, I believe that New Zealand white rabbits are preferred. Does anybody use any other species of rabbit?

(No response.)

Going once, sold. Sorry. No, we haven't.

Come on down.

(Laughter.)

DR. CHAWLA: Anil Chawla from

Panacea Biotec. What are the scientific bases of the schedules of zero to four weeks and then challenge at tenth week?

DR. BIGGER: What is the --

DR. CHAWLA: What is the scientific basis of choosing that schedule?

DR. BIGGER: The scientific basis of choosing zero weeks and four weeks for the vaccination --

DR. CHAWLA: And then challenge at tenth week when you have the peak titer at sixth week.

DR. BIGGER: And then -- okay. And there was lots of discussion when these experiments were being set up as to whether or not we should use zero weeks and two weeks, zero weeks and three weeks, zero weeks and four weeks. Obviously, you know, what you'd like -- prefer to do in a vaccination is to vaccinate them and allow them to rest for a period of time.

And it was discussed in depth and

four weeks was arrived to as a consensus. Again, the challenge date of ten weeks was discussed, whether or not we wanted to do it earlier, whether or not we wanted to do it later, and really that was just -- we wanted to do it early enough so that it was not a long-term wait. Okay? So that it was not a six month or year or two year wait to -- we wanted to address the near-term efficacy of the vaccines, not the long-term efficacy.

DR. CHAWLA: When you say that X microgram of rPA was used, do you really check it at the time of vaccination, it was X microgram or it was a predetermined value which was tested maybe at the time, which was manufactured two or three months back?

DR. BIGGER: Yes, that's an outstanding question. And the question was: did we do a dose confirmation on the rPA vaccines at the time of vaccination to ensure that we were really giving them the dose that we had anticipated? And the answer is is that

at the time -- and I don't believe this has changed -- the rPA mixture with alhydrogel made it impossible for us to do a back dose titration on the rPA. It binds irrevocably to the alhydrogel, and that made it impossible for us to assay.

So I don't know if the technology -- and there's an experiment now that can make that happen, but at the time that was not possible. So we had to rely on the manufacturers to assay the amount of rPA, and then let us know what that was, so that we could dilute appropriately.

DR. CHAWLA: Thank you.

DR. NABORS: Hi. Gary Nabors from Emergent Biosolutions. My question is, for the TNA NF_{50} assay, or that endpoint from the study, was that the same standard that was discussed before the AVR-801, or was that a rabbit standard that was used in the assay?

DR. BIGGER: I believe -- and Chris can give me the nod here -- that was a rabbit

standard. And I would like to point out that we've had, you know, several years of assay development and increase in fidelity and changes in platform as we continue to refine that assay since these data were conducted.

DR. NABORS: So just as a quick followup, do you think that these data are translatable to immunogenicity data in any way that you would see in humans, or was this more of a sort of model development effort?

DR. BIGGER: I think this was a model development effort, and as we continue in TNAs, as they continue to refine, we're going to get closer and closer to be able to make that comparison. Still, I think if we were to go back and rerun these samples today, the data are sound. And whether or not they are comparable to humans is part of what this workshop I guess is all about. I'm not going to comment on that.

Sir?

DR. GOTSCHLICH: Emil Gotschlich,

Rockefeller University. I must have missed this, but what was the intended dose to be given to these rabbits of antigen?

DR. BIGGER: Sir, I'm not at liberty to discuss the intended -- you're talking about the vaccine dose?

DR. GOTSCHLICH: Yes.

DR. BIGGER: I'm not at liberty to discuss the vaccine dose given to these animals. It's proprietary information for the companies that provided the vaccine. So, I mean, the -- our intent here was to focus more on the immune response that was generated and how that immune response correlated to survivorship.

DR. GOTSCHLICH: I find that an amazing answer.

MS. PASETTI: Marcela Pasetti, University of Maryland. It's a beautiful study. I know you are concentrated on TNA, but did you perhaps free cells or did you do some cell-based assays -- cytokines or

antibody-secreting cells or B memory? I know there are limited reagents for rabbits, but --

DR. BIGGER: So, yes, the -- you know, the ability to do that kind of work has evolved so much in the past couple of years since these studies were done. But no, at the time we did not try to do any of that work. We have since then brought in -- I'm sorry? Some of my staff members were hinting to me maybe?

But since then, we've brought online B-cell memory assays and we can detect various Ig levels in different species. And I'm not sure if we can do that in rabbits with the IgA, but we can certainly do it in non-human primates at least.

MS. PASETTI: Thank you.

MR. SUTER: Maybe I've done that experiment. Would it be possible to transfer serum from an immunized rabbit into a naive rabbit, get the same titer, and then challenge it? And do you see the same protection?

DR. BIGGER: Later this morning --

MR. SUTER: Okay.

DR. BIGGER: -- Mark Perry is going to present some data on rabbit passive transfer studies, and I think following that -- and the discussions that we'll have with the panel -- the discussion is going to be very exciting, comparing the results of this active vaccination experiment to his passive transfer experiment.

Thank you. Thank you very much.

(Applause.)

DR. HEWITT: We are going to move on. Our next speaker is Louise Pitt from USAMRIID, and she is going to talk to us today about her rabbit model of active immunization.

DR. PITT: Well, good morning. I'm going to present a series of experiments that were carried out at USAMRIID over the last maybe eight to ten years. I'm looking at an in vitro correlate of immunity in the rabbit model.

These studies were initiated, as I said, probably 10 years ago. At that time, there was a scientific opinion that antibodies actually did not correlate with protection. That assumption was based on a series of non-human primate vaccine efficacy studies, which had small numbers of animals, and on varied guinea pig efficacy studies looking at a variety of different vaccines.

But at that time nobody had actually designed a study to look at the question specifically. So at USAMRIID we decided to approach that. And in terms of where we stood at the time we knew that the New Zealand white rabbit was probably an appropriate model.

We had done the disease and pathology comparison with non-human primates and humans, and somewhat understood the differences and comparisons between the human and non-human primates and the rabbit. We also had done vaccine efficacy studies in the

rabbit with both rPA and the licensed anthrax vaccine, and knew that it was predictive of efficacy in the non-human primate.

We also came from the standpoint that protective antigen combined with an adjuvant provided complete protection, as I said, and that we did have this quantitative ELISA and the toxin-neutralizing antibody assay that could be used for correlates.

So the approach we took in the first study was to take the licensed anthrax vaccine and dilute it down. We knew that the human dose gave full protection, two doses of the human dose, and so dilute it down in order to start getting survival and non-survivors so we could then compare the responses and see if there was a correlation.

The study design was two doses at zero and four weeks. We bled the animals and looked at the titers at week 6, and then immediately prior to challenge. We chose to challenge the animals at week 10 to match a

lot of the vaccine efficacy data that we already had, and have found that that was appropriate time -- six weeks after the second dose -- to look at efficacy.

There were two studies performed with the licensed vaccine looking at two different lots, and their challenge doses were an average of 133 LD_{50}s or 84 LD_{50}.

So this is a summary of the initial study. Here, as you can see, we did one in four dilutions in groups of animals. The undiluted, as expected, gave the 100 percent efficacy. And as you can see, as you go down, we get excellent efficacy from 100 to 90 percent when you get down to one in 64 dilution. And then, at the one in 256 in this lot, we started to lose animals.

If you look at the week 6 quantitative ELISA, we got a very nice titration if you look at it as a group, both at the six week and at the 10 week. And, indeed, in the TNA ED_{50}s, the titer gave a nice

gradation.

In the second study, we increased the numbers in the middle groups in an attempt to increase the non-survivors so as it would improve our statistics. And, again, you can see excellent efficacy at the first two dilutions, and then starting to drop off with zero survival at the one in 256.

The same pattern in terms of the quantitative ELISA quantities, both at the six weeks and at the 10 weeks, and, again, a similar gradation in the toxin-neutralizing antibody levels.

This is just a graph of the actual individual animals with the live in the closed diamonds and the dead in the open to show you exactly the titers of each individual animal. As you can see in the top group, you clearly have a group that is solidly protected. In the lower groups, although they do have some levels of antibody, they are clearly not protected, and then in between you have some

that are and some that aren't. And this was true for both lots of vaccine that we did experiments.

So this is the concentration at the time of challenge; the previous slide was at the peak, the six weeks. And this shows you a very similar pattern in terms of the solid protected and the solidly not protected. That gives us the gradation in between. The TNA level at six week again followed the similar pattern of the groups.

So in terms of predictions of survival, both the six-week peak and the 10 week ELISAs were significant predictors of survival, and as was the toxin-neutralizing antibody assays.

So in the next series of studies -- and these were led by Steve Little from USAMRIID -- the next logical step was to look at the rPA vaccine and say, "Did this hold true, and was the pattern similar?" Again, a similar study design. In the initial rPA

studies we looked at a one-dose vaccination to see if this -- if there would be any correlation following one dose and a challenge.

The doses of rPA that were chosen varied between .08 micrograms and 100 micrograms of rPA. They were combined with .5 milligrams of aluminum, so the amount of aluminum in each dose remained constant. This is different from when we diluted the anthrax licensed vaccine, because we diluted that in PBS.

So the amount of aluminum changed in the initial experiment, but the ratio of antigen to aluminum maintained. In this experiment, the aluminum remained constant and the rPA was titrated.

The animals were bled at week 2, and then at time of challenge, and, again, we looked at the ELISA and the toxin-neutralizing antibody levels, and the animals were challenged at week 4 with approximately 200

LD_{50}s.

So this is the table that summarizes. Here we have the dose of the rPA going from 100 down to .08. This column just shows you the number of experiments. As you know, we can't do hundreds of rabbits at any given time, and so it had to be split up into several experiments.

And here we have the survival column where at 100 micrograms of rPA, one dose, four weeks later, 93 percent are protected; 65 with 25 micrograms; 43 with five, 16 percent with one, 10 percent with .2, and then zero at .08. So we got a very nice titration in survival that follows the dose of the rPA.

When look at the ELISAs, we got a similar pattern of gradation at both week 2 and week 4, and the toxin-neutralizing antibodies followed a similar titration pattern.

This is a graph of the actual live,

dead animals. This shows you there is quite considerable overlap after the one dose of rPA in terms between 100, 25, and five micrograms, and then the response starts to drop off. And this is the quantitative ELISA at four weeks, which is the time of challenge.

This is looking at the TNA response at two weeks, again showing a similar pattern. Clearly, down here, a group that are solidly not protected, but after one dose you have groups where clearly there is a more mixed group of survival and of non-survival.

But looking at it in terms of significant predictors of survival, the PA ELISA was significant at week 4, and indeed it was also significant at week 1. Looking at the TNA, it was a significant predictor at both week 2 and week 4.

So moving on to the next series of experiments, we then looked at what would happen with two doses of rPA. This slide actually has some mistakes on it, and I'll go

through it. The doses were varying from .08 to 10 micrograms of rPA. Again, it was combined with aluminum at the same concentration, so each -- each injection had the same amounts of aluminum and the rPA was titrated.

This sera -- they were bled pretty much weekly, but we concentrated really on looking at the six-week and then the ten-week, the prior to challenge. The aerosol was done at week 10, not week 4, and they were given over 200 LD_{50}s.

So this is the summary table of the survival with the doses going from 10, 1.2, and .08, two doses zero and four weeks, 100 percent survival with two doses of 10 micrograms rPA. No difference seen in survival with the one microgram or the .2 microgram, and then a drop-off at two doses of .08.

In looking at the ELISA at week 6, again, you can see a nice gradation, but right

here you can see there is somewhat difference in terms of the ELISA, but no difference in terms of survival.

At week 10, similar gradation. And when you look at the TNA, again, a nice titration in terms of the assay. This then shows you the individual animals and their levels of ELISA at week 10 just prior to challenge, and you can see here solidly protected group here. These are the controls down here, and your 1.2 and .08.

The toxin-neutralizing titers at week 8 showing a very similar pattern where you've got the solidly protected and then the mixed in between.

So in terms of significant predictors of survival, the PA ELISA at week 10 was indeed a significant predictor, and the TNA at week 8.

So the last study that I will present approached in a little different fashion. Instead of looking at a short-term

challenge, this was looking at a six-month and a 12-month challenge. In this study, the animals received two doses of 50 micrograms of rPA combined with the same amount of aluminum. Blood was drawn at various times through the experiment, and one group was challenged at six months and another group challenged at 12 months.

So this is the efficacy at six months where 74 percent of the animals -- 20 out of 27 -- survived the challenge. You can see this is the weeks 4, 6, 8, 13, and 26 levels between survivors and non-survivors. Week 26, in terms of the ELISA, there was a significant difference between the survivors and the non-survivors, and indeed that was a significant predictor of survival.

In terms of the TNA assay, there was a significant difference at week -- between survivors and non-survivors at weeks 8, 13, and 26, and the week 13 was shown to be a significant predictor of survival.

Looking at the 12-month challenge, at this time we got nine out of 24 survivors. The response -- there was a significant difference between survivors and non-survivors at 26, 39, and 52 weeks, and the 26-week turned out to be the significant predictor of survival.

Looking at the TNA assay, there was a significant difference between survivors and non-survivors weeks 6, 8, 13, 26, 39, and 52. But the week 39 in the TNA was the most significant predictor of survival.

So in summary, looking at these series of experiments, we showed that these two assays -- both the quantitative ELISA and the TNA -- are useful assays to serve for correlates in estimating the immunological status of rabbits. We found that the antibodies to PA are a serological correlate of vaccino-genized immunity in this model. And this could provide a basis of an in vitro test to serve as a correlate.

This data has all been published, and the references are available. And I would like to acknowledge all of the people at USAMRIID who have participated in these studies over the years, particularly Steve Little who did so much work not only in the animal studies but on the assays and the development of the assays.

And as you know, these studies take a large number of people, and I would like to acknowledge them.

Thank you.

(Applause.)

DR. FERRIERI: A quick question, Dr. Pitt. I gather that the rPA is not absorbed to the aluminum hydroxide. And my question is: is there -- what is your opinion of what it would do, its behavior immunologically, if it had been absorbed?

DR. PITT: It was absorbed. It was just --

DR. FERRIERI: It was.

DR. PITT: -- not a formulation. It was absorbed within 24 hours, and then given to the animals.

DR. FERRIERI: Okay. Thank you.

MS. WILLIAMSON: Louise, I just wanted to ask about the duration of immunity. In terms of anti-PA ELISA, that seems to be fairly standard in these longer-term studies, that week 26 was the critical time point. But the dynamics for the TNA titer seemed to vary much more.

Can you explain why, or any theories why that might be? Because one would like to -- the function antibody titer, I would have expected the function antibody titer to follow the ELISA titer. So, you know, you're developing antibodies, and within that you're developing a functional antibody. You might expect that to be a slower process, but it doesn't seem to be from this data necessarily.

DR. PITT: No, I agree, but that's

-- I honestly don't have an explanation at this time.

MS. WILLIAMSON: Thank you.

DR. CHAWLA: Anil Chawla from Panacea Biotec. The question is in clarification of the first question she asked. You have used different amount of rPA starting from 0.8 microgram to 100 microgram on same amount of aluminum that is 0.5 milligram. Did you carry out absorption studies that how much was absorbed? Was it 100 percent absorbed in all cases?

DR. PITT: I don't know that.

DR. CHAWLA: Thank you.

DR. NASS: Did you perform a functional assay of the PA to find out whether it was biologically active?

DR. PITT: In terms of using the TNA to show that PA is active, yes.

DR. CHAWLA: Again, a clarification on the question she asked. When she said that rPA which was used, was it biologically

active, I mean, was in microphage license assays done? Or not the TNA, because for checking the biological activity of PA you need to carry out the microphage license assay.

DR. PITT: Yes, that was performed.

DR. CHAWLA: It was. Okay.

DR. HEWITT: Thank you, Louise.

We'll move on to our next talk by David Madigan, and he is going to tell us about his non-human primate analysis.

DR. MADIGAN: Thank you. Good morning. Indeed, I'm going to tell you about the non-human primate anthrax vaccine study run by CDC. And just at the outset, I'm very grateful to Brian Plikaytis and Conrad Quinn, who are here from the CDC, and they are going to answer all of the questions.

I have the wrong slides. Can we take a five-minute break? These are the wrong slides that are loaded on the laptop. Can we take a five-minute break, please?

(Whereupon, the proceedings in the foregoing matter went off the record at 11:40 a.m. and went back on the record at 11:43 a.m.)

DR. MADIGAN: Okay. Can we resume?

So I apologize. The version of the slides that I'm now showing you are updated versions of what's in the handout. And if there are some extra slides and some corrections. If you would like a copy of these slides, feel free to e-mail me, and I'll send you this updated copy of the slides.

Okay. Okay. So this particular study was run by the CDC. The goal of the study was to find immunologic markers that endorsed the human clinical trial endpoint, and confirms human vaccine protection, and identifies when protection is achieved, and also it quantifies how long protection lasts.

And this study was heavily scrutinized by an IOM Committee several years ago. Several members of the Committee are

here in the audience. Subsequent to that Committee, the Statistical Advisory Committee prepared a statistical analysis plan for this study, and there are a couple members of that Committee here also. And then, it fell to me to actually implement the statistical plan.

Basically, I'm primarily just going to show you some of the data from the study, and very briefly I'll describe some of the statistical methods that we implemented.

So this study was a lot like some of the other studies we've just been hearing about. It was in non-human primates, and we -- there were different doses of the vaccine used, the human dose of 1/5, 1/10, 1/20, 1/40, and as well as saving controls. And the human -- the proposed human vaccination schedule was 0 weeks, 4 weeks, and 26 weeks IM.

And so our goal was to build a comprehensive -- or the study did build a comprehensive immunological profile. The animals were challenged at different times,

some at 12 months, 32 months, and 42 months, and the statistical goal was to build a model to -- for predicting survival using these assorted -- large number of measurements of the state of the immune system gathered throughout the study. And the longer term goal is to apply this relationship to the human clinical study.

So a little bit more specifically, the question was: are measurable aspects of the state of the immune system predictive of survival? The answer to that is yes, as I'll show you. And the basic statistical problem we had here is that we had -- we had literally hundreds of different assay time points, different assays measured at different time points, but there are fewer than 100 animals in the study.

I'll describe a descriptive analysis, and then briefly I'll talk about some of the fancier statistical things that we also explored.

So here are some basic statistics about the study. So there were 12 different groups of animals. The first three groups, there were -- the dose of the vaccine were the human dose, which is one and one, and then a 1/5 and a 1/10 dose. And there were approximately 12 groups, 12 animals in each of these groups, and with one or two controls in each of the groups.

These animals were challenged at 228 weeks. The next group of animals at different doses, one in 20, one in 10, one in 40, were challenged at 52 weeks, and then these group of animals were challenged at 124 weeks, and there's various doses in there from the human dose down to one in 20. So there's a total of 114 vaccinated animals, and 23 controls in the study.

Here's a broad-brush summary. This is the death rate by dose, so of the 20 animals that received the human dose two of those animals died, and so on. Zero animals

and one in five, and nine of 29, 31 percent died at the one in 10, one in 20, one in 40, and 70 percent of the control animals died, which means there were 23 control animals, and seven of the control animals actually survived the challenge. The overall death rate is 32 percent.

So IgG was the strongest predictor of survival, and very similar to what we've seen in some of the other presentations this morning. Here is a plot of IgG at week 8 on the left-hand side and week 30 on the right-hand side, and as a function of dose going from the control animals receiving zero up to the human dose. And so there's a strong dose-response in terms of IgG, and it will be the same -- at any week I show you, the plot would look very similar.

This is a plot of, again, week 8, week 30. IgG is on the vertical axis, comparing the animals that died and the animals that survived. And quite clearly, the

animals that died had lower IgG responses than the animals that survived. And in some sense, that's the primary conclusion.

And the plots -- it looks similar at various -- no matter which week you look at, the plots will be broadly similar.

Here's another way of looking at the same thing. At the top here are the animals that received the human dose, and this is a trace of their IgG levels over time. These are the different -- so in here, and in all of the plots I'm going to show you -- it doesn't come out very clear in the handouts -- I'm using black to denote the animals that survived and red to denote the animals that died.

So there were two animals. This is not perhaps so clear here, but there are two animals here who received the full human dose who -- but died, and they're at the -- they have -- they're at the low end of the IgG response. And these are the control animals,

and as you can see these animals did not mount -- as you would expect, did not show an immune response, although they did not all die upon challenge.

I'm now going to show you a series of pictures that showed -- that go through the groups, and basically showing you the data. And, in particular, I'm looking at IgG on the log scale. In general, I'm showing these things on the log scale.

And this is group -- these are the group 1 animals, and full human dose. Two of them died. And as you can see now more clearly in this picture, the two animals that died had a lower IgG response generally than the other animals. These are post-challenge measurements, which are in some sense not interesting from a predictive point of view.

So that's group 1, human dose. That's group 2, the one in five dose. And group 3, the one in 10 dose. And you can see kind of the -- as I play this move there is --

you know, this is the dose response showing again. There's a -- you know, clearly, the animals that received the human dose mounted a higher level of immune response.

No animals died in the one in five, and then in the one in 10 several animals died. I think it's three animals died in that group.

And I'll quickly flip through the other groups. And all of these pictures are on the same vertical scale, so you can get some sense of what the level of the IgG measurements are, but the horizontal scale is changing because these different groups are -- were challenged at different times.

So it's a one in 10 group, one in 20 group, that's a one in 40 group, where a lot of the animals, perhaps surprisingly, that did not mount much of an immune response did survive challenge. That's a human dose group, group 10, a one in five group, one in 10 group, and here are some animals -- there were

no deaths in this group, and there are several deaths in this group, and only one death in this one in 20 group.

And so this is a summary of the pictures that I just showed you, and this is the human dose, one in five, one in 10, one in 20, one in 40, showing IgG. And there is basically more red as you go -- you know, as the dose goes down, there is more red meaning more animals died in general.

This is the same picture for -- showing the same picture for ED_{50}, and there's a strong correlation between the IgG measurements and ED_{50}. Very similar pictures. Now I'm going to show you the same picture for some of the other assays. There were many assays measured throughout these -- the duration of this study.

That's what the picture looks like for IFN, and a lot of the pictures are going to look like this. So, you know, basically certainly it's the human eye. There is no

predictive power here. There is this -- this particular assay does not appear to be discriminating between the animals that lived and died, so that's IFNELI. That's SI, the stimulation index, so it's very similar. You know, it looks like there is absolutely no predictive power in this assay, in discriminating between the animals that lived and died.

This is IL4, IFN -- we are missing some of the later measurements. But, again, you know, the -- to the naked eye there certainly does not seem to be anything going on here.

This is -- we looked at a number of ratios. This is a ratio of ED_{50} to IgG. It does not appear to be particularly useful, and so on. I could show you lots of pictures like this. So, basically, the TNA measurements and the IgG measurements show strong discrimination -- discriminatory power between predicting -- for predicting survival, and

basically none of the other assays seem to show much predictive power as what happened.

Okay. So that's descriptively what the data show. We tried out a variety of analyses to see if there was some predictive power in there that was not apparent to the naked eye. The primary tool we used was just -- was logistic regression, so I'm going to describe this briefly. And it's small but technical.

So logistic regression -- I assume most people are familiar with -- is a binary regression model, so it's a model when you're doing a regression with a binary response -- in our case survived or died. And it models the log odds of this binary response as a linear function of the predictor variables. In our case, the predictor variables are the assay measurements at each of the time points, and there are 100-plus of these measurements.

And the standard way of fitting a logistic regression model, if you open up SAS

and you press the logistic regression button, it will be doing maximum likelihood estimation. So it estimates the regression parameters that maximize the likelihood of the data, and in many applications that's a perfectly reasonable thing to do.

In our context, we have a problem, which is we have about the same number of assay measurements as we have animals. So if you like, the number of parameters is roughly of the same order of magnitude as the number of observations. So you run into problems with many statistical -- classical statistical techniques in that situation.

If you do logistic regression, it will tend to overfit, and, in fact, it's not defined if the number of measurements exceeds the number of animals, and that is exactly the situation that we are in. So a standard way of dealing with this problem is to do some sort of feature selection. So instead of using all of the candidate predictors, select

out a small number of them and then fit a logistic regression model using those.

I'm not describing that here. We did pursue that type of analysis, and indeed IgG and TNA turned out to be statistically highly significant, and basically nothing else is. An alternative approach that we explored in some detail is to use shrinkage methods in this context.

So the basic idea here is instead of doing maximum likelihood estimation, use a shrunken version of a maximum likelihood estimate, so you sort of hedge your bet somewhere between estimating the regression coefficients at zero and the full maximum likelihood estimates.

This has become a very popular way to deal with overfitting and to fit standard logistic regression, other kinds of regression models, in context where you have more parameters than you have observations. And it has become fairly routine in many settings.

It's a very simple idea. The idea is you do maximum likelihood estimation, but you put a constraint on the regression coefficients. You don't allow the regression coefficients to get very large.

There are two -- well, there are an infinite number of varieties of it, but there are two varieties of this that are in common use. One is Ridge regression and the other is called Lasso regression. In Ridge regression, you do maximum likelihood estimation plus a constraint on the squares of the regression coefficients. And with Lasso logistic regression you put a constraint on the absolute values of the regression coefficients.

This is the net effect of these two types of shrinkage. So if you do -- if you put a constraint on the squares of the regression coefficients, basically depending on that number, you get to choose that number, depending on where you choose that number, you

either get the maximum likelihood estimates if you choose the number to be large, or you get zero if you choose S to be zero.

And in general you would choose a value of S somewhere in the middle of the range, and you get estimates of the parameters that are somewhere between the maximum likelihood estimates and zero. This procedure has some very attractive theoretical properties, and you can use it in situations where you have more parameters than you have observations.

Generally, you would choose S by cross-validation, and that's exactly what we did in our context. So that's the picture for Ridge regression. This is the picture for Lasso regression, and they are very different.

And here you get the same type of shrinkage. These are the maximum likelihood estimates. This is zero over here. But now when you start shrinking you shrink the regression coefficients toward zero, and at

some point they actually hit zero.

So for instance here if you chose S to be here, you would get shrunken estimates for this parameter and for this parameter -- these coefficients -- and all of the other coefficients would be shrunk all the way to zero. So it simultaneously selects variables and gives you shrinkage estimates of the parameters, and it's very attractive.

So this is the primary method we employed. This was in the statistical plan that was developed and the primary analysis that we carried out. And so here -- and doing this Lasso logistic regression, L1 logistic regression, these are the particular variables that it selected, and it's not terribly surprising.

So IgG at week 38 was the largest predictor, in terms of the coefficient, and coefficients are negative here meaning the higher this is, the less likely you are to die. That's the way they're phrased. So IgG

comes out on top at week 38, ED_{50} at week 30, and then you get into other stuff. So the stimulation index at week 8 is the third variable into the model, and so on. ED_{50} appears here and here, and then some of the other assays make an appearance.

But note, these are the standard errors associated with these coefficients, so in this -- you know, I wouldn't get too excited about this in the sense that, you know, the standard error associated with that coefficient is very large. So even though it is picking the stimulation index, it is still not a very strong predictor of survival, and IgG -- also not significant in this particular analysis, but IgG comes out as the most important predictor.

For technical reasons, we did a number of different versions of this. This is a different one where we did -- there are a number of missing assay values, and we used different kinds of imputation schemes to

impute some of these missing values. With this particular imputation only three measurements were selected as predictors of survival -- an IgG measurement, an ED_{50} measurement, and, again, the stimulation index at week 8.

This I'll go through very quickly. We explored an alternative kind of analysis based on decision trees. And decision trees, if you're not familiar with them, is a very simple type of regression model where you just recursively partition the predictor space.

This is the decision tree that was selected from the data using groups 1 through 3, and basically you -- I'm sure you can't see that from where you're sitting, but the first split corresponds to kind of the most important predictor in this model in this analysis, and TNA at week 38 is the most important split according to this analysis.

But it's sort of Tweedledum/ Tweedledee whether you use an IGG measurement

or you use a TNA measurement, and you get much the same quality of split. And then, further down the tree it splits on other measurements.

There's something sort of biologically very unsatisfactory with all of these analyses, which is particularly -- let me go back to this one. You know, what I'm doing here is I'm saying the IgG measurement at a particular moment in time is the single most important predictor.

Well, you know, the fact that it's week 34, it could just as well have been week 38 or week 20 or whatever. If you do standard variable selection on this kind of a product, it will take specific measurements at specific moments in time.

We explored -- we developed some other techniques, which instead used the entire trajectory of an assay as a predictor. And I'm not going to go into the details, but we developed a technique called functional decision trees, and there's a paper or two

describing this, which does this. It's a decision tree analysis, but it splits on the entire trajectory of the assay. And surprise, surprise, what it splits on at the top is IgG.

So basically animals that have an IgG trajectory that's more like this one are more likely to -- the preponderance of them survive, and animals that have an IgG trajectory that's more like this one, the preponderance of those animals died. And then, there are further splits down the trees, which may or may not be of interest, but IgG comes out as by far and away the most important split.

Okay. So to conclude, overall, there is significant survival at 53 months with animal -- with vaccinated animals as shown here, so the human animals, 80 percent of them survived -- the animals that received the human dose. The animals that received the 1/5 dose, 100 percent of them survived. And by the way you get down to 1/10, there are

more -- there's larger numbers of deaths occurring.

IgG and TNA are highly correlated with survival. They are also highly correlated with each other, which I'm showing you that. But from a predictive point of view, they're more or less interchangeable. The other assays that are measured are at best weakly correlated with survival.

And just a note -- TNA levels above 250 ED_{50} give greater than 90 percent survival, which is very close to the number we just saw for the rabbit study.

Thank you.

(Applause.)

DR. NASS: I'm sorry if I missed this. What dose did the monkeys receive? Did they get .5 ccs?

DR. MADIGAN: Conrad can answer that question.

DR. QUINN: In terms of volume -- Conrad Quinn, CDC. In terms of volume, the

animals got .5 mls of the vaccine dose at different dilutions to accommodate different antigen loads, so that we could modulate the immune response.

DR. NASS: Okay. So that the animals that got the full dose received a non-weight-based dose identical to the human.

DR. QUINN: That's correct.

DR. NASS: And I assume these were Rhesus Macaques? And what did they weigh?

DR. QUINN: These were Rhesus Macaques of Chinese origin, and the minimum entry weight into the study at time of initiation was 2.6 kilos.

DR. NASS: And the maximum weight?

DR. QUINN: There was no maximum at study start, but obviously over 53 months they put on some weight.

DR. NASS: Okay. So you've got a six-pound monkey getting a full human dose, and you're saying that you're doing a dose reduction study.

DR. QUINN: I'm not sure I understood the question.

DR. NASS: You're reducing from six doses over 18 months to three doses, but each dose is approximately 20 or 30 times by weight the human dose. So I can't -- don't see how this can really correlate to giving you data that suggest that a dose reduction in humans is viable.

DR. QUINN: Shall I take that one?

DR. MADIGAN: Sure.

DR. QUINN: This is a correlate of protection study, and the objective, as David pointed out, is to determine what components, either singularly or in combination with the immune response -- the immune response profile correlate with protection in Rhesus Macaques. So this is a Rhesus Macaque study.

DR. CHAWLA: Anil Chawla from Panacea Biotec. Groups 1, 2, and 3, did they receive the same batch of AVA? Was it the same batch of AVA which was diluted --

DR. QUINN: Yes.

DR. CHAWLA: -- different concentration?

DR. QUINN: Conrad Quinn, CDC. Yes, all of these animals received the same batch, same lot of AVA.

DR. CHAWLA: Okay. Thank you.

MR. SUTER: You looked at several variables. Do you also look at IgA?

DR. MADIGAN: No, that was not one of the ones that was -- that was measured.

MR. SUTER: It's interesting, because it's an aerosol study, and obviously IgA is the predominant antibody in this location. Maybe you looked at the wrong immunoglobin isotype.

DR. QUINN: Conrad Quinn, CDC.

(Laughter.)

DR. MADIGAN: See, I told you he'd answer all the questions.

DR. QUINN: I should stay up here. To address the question, this is a response

to the vaccination, and although it is possible that intramuscular vaccine with AVA does elicit an immune response at the mucosal surface, it is a very different immunological presentation to generate an IgA response.

And other studies in animals, and data, I believe, in humans indicate that to get a good IgA response you need to give a specific intranasal or inhalation vaccination using specific adjuvants. It's a different immunological compartment.

MS. SABOURIN: Carol Sabourin, Battelle. I think it's important to point out in your cytokine analysis that the peripheral blood mononuclear cells, there were two different stimulation times where the cytokine levels were evaluated by mRNA levels in ELISA. So there was a difference between the groups for the cytokine stimulation time with rPA.

DR. MADIGAN: Okay.

DR. HEWITT: Okay. Thank you very much.

Since we're a little bit ahead of schedule, I think we're going to move on to the next talk, so that we can finish our discussion of the animal models, and then we'll reconvene at the time this afternoon, 1:50, and pick up the human presentations then.

So Mark Perry from Battelle is going to tell us about his passive transfer models.

MR. PERRY: I'd like to start off by thanking the organizers for letting me present our study. And also, for the person who asked about passive transfer, it's a great lead-in to my presentation.

My name is Mark Perry, and I work at Battelle's Biomedical Research Center, and I will be presenting data from two passive transfer studies in the rabbit model. The objective of these studies was to assess the protective efficacy of human anti-PA IgG administered to the rabbit model via IP route,

and see how that protected against a bacillus anthracis aerosol challenge.

In the first study, we assessed the foreign protein tolerance and the pharmacokinetics, and 14 days into that -- after IP dosing we also challenged the high-dose group with an aerosol challenge.

In the second study, we refined some of our dosing levels, and actually added some plasma, straight plasma, and assessed the protection efficacy 24 hours after IP dosing. And at the end of this presentation I'll go over the correlates of protection.

So in study 1, like I already mentioned, we assessed the foreign protein tolerance in kinetics and 14-day protection efficacy from human anti-PA IgG. Now, this is purified from pooled human plasma. It's to isolate the IgG, and then we administered it into the rabbit via the IP route.

Our animal model was specific pathogen-free New Zealand white rabbits

weighing between 2.3 and 2.5 kilograms. They were dosed on day zero, and aerosol -- only the high-dose group and a control group were challenged on day 14, and for 28 days we collected blood samples for anti-PA ELISA and TNA analysis. We collected clinical observations, and for those animals that received an aerosol challenge we took some blood samples for bacteremia evaluation.

Our test material, like I have already stated, was human material, human plasma, which we had purified the IgG out of that. And we also -- for control material we got some naive plasma and which we analyzed to be negative for TNA and anti-PA ELISA.

And as you can see from our -- this data here, we have the sample material pretty much normalized at the same total protein load, but the normal was zero. Actually, that's probably below the detection of -- below the detection limit, not actually zero.

In our challenge model, here is the

dose group we had. We had three groups that had progressively increasing dose levels, and this is a dose of milligrams of anti-PA IgG per kilogram of rabbit weight. Group 4 was just the normal IgG, and that animal group got comparable volume -- no, actually a comparable total protein dose as those animals in Group 3.

And Group 5 was basically our control, which we used for the aerosol challenge. You can see the quite extensive blood sample collection time. We wanted to make sure we had sufficient data to do a pharmacokinetic event analysis. Only Groups 3, 4, and 5 got the aerosol challenge on day 14, and those animals also had bacteremias after the aerosol challenge.

Clinical observations. We did not note any significant adverse clinical observations for the animals that received these dosings. And for those animals that received the aerosol challenge we saw the

typical clinical signs that -- of an anthrax progression until they succumbed.

The ELISA and TNA results, which are presented here, showed a very nice dose response. With the blue lines you can see on each side, here's the high-dose rabbits and their TNA -- anti-PA ELISA response. And here are the TNA high-dose groups over here. And we had a very dose -- nice dose response for all of the three dose groups.

Now, you'll see a break in the data here. As I stated already, the high-dose group and the controls were the only ones who received an aerosol challenge, and that was at day 14. As I get into the next slide, all those dosed animals did die. But you'll also see an interesting data point on both these points right here. One animal did show an increased titer in ELISA and TNA, just before succumbing to an infection.

It's important to note that we did some analysis on our ELISA data, and we found

that the human conjugate did cross-react with the rabbit anti-PA IgG. So it was -- it's -- our inference from this is that the rabbit started to build its own -- develop its own immune response just before it succumbed.

Pharmacokinetic analysis of ELISA and TNA data -- we did get maximum concentration between one and two days, and that was consistent between the ELISA and the TNA results. And our half-lives again were consistent for both sets of data with the half-life somewhere between two and three days.

Our C_{max} levels showed a nice dose response for all of the three different dose groups. That was consistent for the ELISA and the TNA data.

Analysis of our pharmacokinetic data -- this is just another way of showing what you saw with your eye on this other table. We did have some good dose proportionality with our results.

None of the animals that received the aerosol challenge that were dosed with the anti-PA IgG from the human product survived, and their deaths was very similar to our two control groups, the ones that just had normal -- naive IgG and also the controls that were not treated at all.

With that foundation set, we moved on to the second study, and in this study we wanted to assess the protection efficacy of the same AVA IgG that we had in the first study, but this time we actually added the straight AVA pooled plasma to see if that could provide protection without us having to do additional processing of the material.

They both -- they were dosed in the IP route, just like the first study, and the new part to this is that we did the aerosol challenge 24 hours after they received their IP dosing. Same rabbit model. Again, the dosing -- the IP dosing was on day zero, but the aerosol challenge on day one.

We kept our bleeds for the TNA and ELISA consistent from the last study. And bacteremia was for all of the animals in this study, because they were all aerosol challenge, and clinical obs were also similar.

So, as I stated, we added another material -- the AVA plasma. This is the straight material, and, as you would expect, because we did not purify it, the total protein load of that material was quite a bit higher. So to ensure that we had a proper control we used our normal pooled plasma, and had that as our plasma control.

Both all of our naive, our normal plasma as we're calling it here, were negative for the anti-PA ELISA and the TNA. These should be below detection limits.

Another -- several other important things to note is that our plasmas had comparable total protein concentrations, and our purified IgG also had comparable concentration. So the matrix almost doubled.

If you can look from one table to the next, you'll notice that our actual dosing levels did increase slightly. We wanted to hedge our study to make sure that we had some protection there, and to demonstrate some efficacy.

The IgG animals had comparable anti-PA dosing levels as the plasma dose to animals. But as you would expect, because the plasma had other protein components, they received much greater protein load during the dose.

Aerosol challenge was 200 LD_{50}s inhalation, 24 to 36 hours after dosing, and these bleed time points are very comparable to the previous study.

Clinical observations -- the rabbits that received the plasma did show some adverse clinical signs within the first four hours after dosing. Obviously, we were giving them quite a protein load from a human, so there was probably some foreign protein

intolerance issues there.

It would have been nice if we had some clinical camera hematology of this, and we have actually looked at followup studies. But for this one we did not have that data. But the animals that received the purified IgG, they did not show any adverse clinical signs. They were actually up bouncing around right after dosing.

And all of the animals that did succumb to anthrax infection had the clinical signs you'd expect from a normal anthrax infection.

Here's our ELISA results for all of the groups. The one on the left, this is all of the animals that were dosed with the IgG. The animals on the right had the straight plasma. If you compare this to the graph which I showed you previously, you'll see that the first few days of this plot are almost identical -- nice dose response, peak concentration around the first day, and they

just have similar elimination phase.

It's very interesting that as the -- as after the aerosol challenge, in those animals that succumbed to -- well, all of these succumbed very quickly. In the high-dose group, we ultimately had 75 percent of those rabbits surviving. The middle dose group, I think it was 25 -- next graph will show you that. And all of the low-dose group ones died.

This -- because of the cross-reaction of our ELISA, I believe this is the rabbit developing its own immune response. It's provided just enough protection from what -- the passive transfer of the human anti-PA to allow it to generate its own immune response.

The plasma -- only the high-dose group has shown any protection at all, and that was 50 percent of those animals that survived. All of the other animals have succumbed fairly quickly.

Similar -- the TNA data showed fairly similar results. Dose response early on for both the IgG-dosed animals and the plasma-dosed animals -- well, the one thing I wanted to mention, which I thought was very interesting, and that is that the plasma-dosed animals had a more rounded time to get peak concentration, whereas the purified IgG was -- seemed to be more quickly uptake -- taken up. And that was shown in both the TNA and the ELISA results.

Protection efficacy -- as I stated, only the group that -- the high-dose groups had any protection efficacy at all with those animals getting 28 milligrams per kilogram of human anti-PA IgG, showing 75 percent protection. The high plasma dose group showed 50 percent protection, and then the IgG group that got 14 milligrams per kilogram showed 25 percent. All of the other animals succumbed in fairly quick order.

We took the data from the second

study and did some corollary protection analysis. Separate logistic regression models were fitted to the TNA and ELISA data and to survival at each of the time points.

And this was only done for the animals that were dosed, so the control groups were not included in this model. And as you would expect, because we had a lot of deaths as the study progressed, fewer animals were included as a model as the time points went out.

A statistically significant slope provides evidence of the correlation between the titer value and the probability of survival. This is shown very nicely in these graphs right here.

This is the correlation plots that was provided by our statistics group, and this is only for the day of challenge. The solid line is our estimate on -- based on the model, and you can see our confidence intervals are here in the dotted lines.

As you get higher up to the higher level of survival, you can see that the confidence interval really spreads out. That's to be expected, because we didn't have 100 percent survival in any of our animal groups. The highest was 75 percent, if you recall.

But we do have a significant correlation between the antibody levels and survival.

This same model was done for every one of the time points. And, again, if you recall, these are the time points. We did bleeds post-IP dosing, and we found significant correlation at multiple time points. For ELISA, we found them the day of challenge, which we just went over, and all of the way out to day 4.

Similar correlations -- significance to the .05 level was seen for the TNA for day 1 all the way out to day 7.

You probably saw a graph similar to

this, if you recall, during Dr. Bigger's presentation where we tried to give you probability survival for different ELISA and TNA levels. If you recall, his levels were significantly lower.

The possible reasoning for this is because there is no cell-mediated immune response here. This is just transferred anti-PA, so the rabbit itself didn't have anything else to help fight off the infection. But for our study these are the highest levels we had in which an animal died.

It's important to note that we didn't have any titers at this level, because we had a lot more deaths -- ours were lower, and we never did have 100 percent survival.

In summary, human anti-PA IgG can be passively transferred to a rabbit, and we are able to get very nice kinetic results from our animal model here. And from our clinical observations, the purified material was much better tolerated than the straight plasma.

Protection efficacy against bacillus anthracis was shown when we had a challenge 24 hours after for some of our higher dose groups. It was not protective at the 14-day post-IP dosing for the 20 milligrams per kilogram dose level.

Significant correlation between anti-PA titers and also TNA data and survival were found. For TNA, it was for days 1 through 7, and for the ELISA we had them for days 1 through 4.

I had a lot of support on that, and here's the people that helped me. Do you have any questions?

(Applause.)

DR. CHAWLA: Anil Chawla from Panacea Biotec. Because the antibodies from humans will act as an antigen in rabbit, so they will be cleared more quickly than if you had used rabbit IgG or rabbit plasma, so I think that study could be having an audit of protection there.

If human antibodies are used in humans, that will give more -- or antibodies from rabbit, if they are used in rabbit they will give more -- better picture of protection.

You can answer to that or --

MR. PERRY: I believe you're right. I think the rabbit -- would probably be tolerated better in a rabbit.

DR. HEWLETT: Erik Hewlett, University of Virginia. You didn't show any of the bacteremia data. Were there any relationships between dose of material, level of bacteremia, and did the animals that survived have -- did any of those animals have bacteremia that was measurable?

MR. PERRY: We've done a lot of bacteremia evaluation at our facility, and early deaths -- sometimes the animals succumbed so quickly that the blood doesn't always become positive in bacteremia -- we have positive bacteremia results.

And trying to keep in the confines of what I was supposed to, I couldn't present all of the data. That correlation has not been evaluated at this time. I can't answer that at this time, sorry.

MR. WINBERRY: Larry Winberry with Biologics Consulting Group. Again, it's very similar in context -- in terms of the timing of the bacteremia, when do you normally see bacteremia arise in these challenged animals?

MR. PERRY: There's several people from Battelle here who could probably echo this. But I do recall seeing positive bacteremias as early as two days after challenge, but I think as you go on and infection starts to manifest, you -- the rates of bacteremia does increase.

I can look into that in more depth, and then answer --

MR. WINBERRY: The reason for the question is that you're pre-treating with the prophylaxis, whereas in most instances with a

hyper immunoglobulin it would be post-exposure. So that the timing of your peak relative to the challenge and the bacteremia, since this is spore-based, they are -- there's going to be a time base for germination.

MR. PERRY: Correct.

MR. WINBERRY: You're working against the functionality of the prophylaxis in terms of the timing.

MR. PERRY: Yes.

MR. WINBERRY: And I'm just wondering if, because you went to 14 to 28 days post-prophylaxis, correct?

MR. PERRY: Yes.

MR. WINBERRY: And then, challenged, you're looking at your pharmacokinetics. You're not maximizing the opportunity for the antibody to be protective.

MR. PERRY: Oh, yes. In Study 1, we did the aerosol challenge at day 14. Study 2, we challenged them 24 hours after IP dosing.

MR. WINBERRY: Okay.

MR. PERRY: The intent there was to hit them with an aerosol challenge when we were around the peak concentration of the human --

MR. WINBERRY: That's why I'm asking in terms of the timing of the post-challenge germination and bacteremia, because we're looking for toxin neutralization. So it's going to take some time to elaborate. I'm just wondering how -- what that timeframe might be. Maybe we can catch up later.

MR. PERRY: Yes. Sorry about that. I haven't got to that chunk of data yet.

MS. WILLIAMSON: Hello. Diane Williamson from DSL, Porton Down. Very nice demonstration of passive transfer of human plasma or IgG works in the rabbit. But I was just wondering whether in view of the rabbit response starting at about 14 days, whether the passive -- the duration of the passive transfer study should be terminated about that

time, so that we're looking at only the response of the human transferred IgG or plasma, and not having the assay confused with the rabbit active response kicking in.

MR. PERRY: I'm sorry. Can you ask that question again?

MS. WILLIAMSON: Okay. So your data seems to suggest that the rabbits are actually developing their own active immune response to PA from about day 14 onwards.

MR. PERRY: Yes.

MS. WILLIAMSON: So I'm just wondering whether the assay should be capped at around that time, so that what we're looking at is clearly the response just to the human IgG in the rabbit and not confused with this incoming rabbit anti-PA response.

MR. PERRY: Yes. It would be nice if we had a human -- an analytical approach that was this specific just to the humans, and a couple of my co-workers had suggested some other analyses or some other ways we can

probably make a more specific assay. But at this point, we don't have that luxury.

MS. WILLIAMSON: Thank you.

DR. LYONS: Hi. Rick Lyons, University of New Mexico. Conrad may know this as well, but I'm just wondering, if you look at -- you sort of extrapolate back from the TNA that you found that was protective, has there been -- has anybody looked at the humanized monoclonals to see if that correlates on a microgram basis or microgram per ml, or do you know the information there?

MR. PERRY: No, I don't. Conrad hasn't come up and defended me once yet.

(Laughter.)

MR. SUTER: I think this is an exciting model. Have you actually tried now to transfer rabbit antibodies into the naive rabbits and see how much either ELISA titer or DNA titer you actually use -- need to protect the animals?

MR. PERRY: No, I haven't. I don't

think anyone at our facility has. Maybe Louise may have tried it at her facility. I'm not sure. Okay. Yes, she has. We haven't at our place.

MR. SUTER: The other question is: have you actually tried to add purified human antibodies into just normal human plasma and so the same experiments? And so to figure out what the plasma effect is in terms of protection.

MR. PERRY: That specific study we haven't done. But we did -- as you saw, had naive human plasma that was negative, and there was no protection. And we did straight human plasma also that had positive anti-PA titers, and that did provide some protection at the highest dose level.

Am I getting your question wrong?

MR. SUTER: Yes. I mean, if you have it in the same soup, you probably can compare directly to antibodies in plasma.

MR. PERRY: We haven't tried that.

MR. SUTER: The last question, have you -- or are you trying to distinguish between the neutralized and neutralizing effect and the effect of an antibody being XE-mediated or performing XE-mediated uptake? That is, if you transfer FAB fragments, you should also get a protection because it's neutralizing, as compared to the whole antibody.

MR. PERRY: Maybe we'll do that in the future. That wasn't part of the objective of this study, but that might be an interesting way to go.

DR. NASS: Did you choose the intraperitoneal route because that was the only one the rabbits could tolerate? Because it seems you are already minimizing the positive effect of the plasma or the hyperimmune plasma.

MR. PERRY: Well, that's an excellent question. Why do we pick the IP route? There are multiple ways to give this

material. We were kind of confined that we only had so much concentration of the anti-PA IgG in our human material. And how do we give sufficient volume to the animal to afford -- to give it high enough titer so it provides some of the protection?

We had considered the IV route, and there's a lot of difficulty in that. And actually, I've seen some past data where the IV route sometimes had led to quicker deaths. The IP -- our route, our IP route, had shown kind of a dampening effect, if you will, so able to get into the circulatory system without shocking the animal too much.

Some other people may have more take on that, because I know that -- I think CDC has tried both routes and had some interesting results also.

MS. VOLKMANN: Ariane Volkmann. It seems to be quite difficult to find a value predicting protection in terms of TNA and ELISA, because when you compare those values

with your colleague, Dr. Bigger, who has reported about the active immunity when you have a factor of 10 or more difference. Could you comment on that?

MR. PERRY: Yes. The term that I was told to use --

(Laughter.)

-- that we had no cellular components adding to the protection here. This immunology is not my strong field. There are some other people in the area here who I'm sure can field that much better than I, but that was the read on -- the best response that we have at this time.

MS. VOLKMANN: Yes. I think you are right, and that's exactly the reason why it is difficult to find such a value. So I think we probably won't be able to compare passive and active immunization and find exactly a value where you can then protect this animal or that person will be protected or not.

MR. PERRY: Okay.

MR. TINO: Yes, Bill Tino, LigoCyte Pharmaceuticals. With your purified antibodies, are you purifying just total IgG, or in some instances where you're going the next step to specifically purify anti-PA specific antibodies, can you briefly describe your methods for purifying those fractions?

MR. PERRY: It was just complete IgG. We did not try to isolate just the anti-PA IgG.

DR. NUZUM: I just thought I would maybe have a few comments. There's a couple of questions on rabbit-to-rabbit transfer.

PARTICIPANT: Speak up.

DR. NUZUM: There were a couple of questions on rabbit-to-rabbit transfer. The important point to remember here -- and I didn't say it in my talk, and if Carol Oestre is here she'll start laughing, but I often say the Animal Rule is not about animals, it's about humans and people. So you have to know

-- we're trying to relate animal stuff to humans. So there is -- yes, rabbit-to-rabbit transfer might work better, but the question is: does human antibody protect animals in a challenge efficacy model?

So that's -- and we've done a lot of other studies that we're not showing here, but the purpose of this workshop is correlates of protection. So we're -- that's the focus.

And then, I wanted to comment on the question on the four-week regimen and 10-week challenge. Remember that the purpose of this study was not a regimen study. It wasn't to identify the optimum regimen. The purpose was to see if we could repeat what Louise had done and show a correlate of protection in rabbits.

At different facility, different staff, at different point in time, different rabbits, different challenge material, could it be repeated? Could we get our own model and a correlate of protection in our own

system for our own purposes?

So the purpose of this study was to get a correlate of protection, not to look at regimen. And it gets to my points of my talk was -- there's lot of questions, but you have to focus your question on the study you want to do, because you can't answer everything in one study. When you try to do that, you run into trouble.

The 10-week challenge was done so that we didn't want a long-term duration study, but we thought that would allow for maturation of antibody that would be consistent with antibody that would be persistent. So the reason we didn't do a challenge at peak titer was that would only have been two weeks after, and the antibody wouldn't have been matured.

So, again, I'm trying to bring the perspective and focus and keep it -- you know, we have to focus studies on specific questions and try to get -- do the best study to get the

right answer.

DR. HEWITT: Okay. I want to thank all of the speakers for some very good presentations this morning on the animal models. And Freyja is going to make an announcement.

(Laughter.)

DR. LYNN: Sorry. I just want to let everybody know that there is a little booth -- not booth, a sort of bar thing right outside the door where the hotel has arranged lunch for us for $8 apiece. You can simply go out, pay there, and then go over and pick up your lunch at the cafeteria, or -- the cafeteria, the restaurant, whatever the food service is here.

The other thing is you -- there was a handout available that has some local restaurants that are more or less within walking distance, if you'd prefer to do that. That's kind of what we have available for lunch today.

And then, we'll reconvene -- do you have anything else, Judy?

DR. HEWITT: No. We'll reconvene at 1:50.

DR. LYNN: Yes.

DR. HEWITT: And pick up with the human session.

DR. LYNN: Right. Thank you all.

(Whereupon, at 12:38 p.m., the proceedings in the foregoing matter recessed for lunch.)

DR. LYNN: All right. If we could start to take our seats, please, and we'll get started again. Once again, I'm Freyja Lynn, and welcome back. I hope everybody had a successful lunch.

The next session we'll be covering, immunogenicity data that we have in humans. We don't have a whole lot of data at this point in time, but we wanted to share what we do have to sort of give a context in terms of having heard about the animal work, and the

animal immunogenicity, what we're starting to learn about the human immunogenicity for both AVA and rPA-based vaccines. So our first speaker will be Ed Nuzum, who will present a University of Maryland study that was sponsored by DMID.

DR. NUZUM: Okay. Thanks, Freyja, and good afternoon. It seems like it wasn't too long ago I was just up here. That may be bad news for all of you.

So this will be fairly quick. And, really, this now, in terms of RPA history, at least at DMID, is fairly ancient history, because this study was actually started, the planning for it was started before I arrived at DMID. Lydia Falk and Carmen Mayer were a couple of key players in the initial planning for this study. And as the slide shows, it was conducted between July of 2003 and February 2004. It wasn't the University of Maryland, it was one of the BTEU sites, and it was sponsored by DMID.

Now this really, as I say, it's going to be quick. It's mainly just to make sure everyone is aware this study was done, and kind of give a high level -- the high level results.

It was a Phase I dose-escalating study to assess the safety, tolerability, and immunogenicity of Recombinant Protective Antigen Anthrax Vaccine administered in two IM doses. And, as you'll see, it also included BioThrax.

The rPA used was the one originated and developed at USAMRIID, and it was made under CGMP by SARC Frederick, at the BDP facility there at NCI.

The general study design is covered on this slide. It was a Phase I study. It was one of the first rPA vaccines produced in the U.S., and that maybe is one of the most significant things about this study. It was one of our first indications that rPA is immunogenic in humans; eighty healthy adults.

And another key point here is the rPA was not pre-adsorbed, it was mixed at the bedside. We gave two doses of rPA only unadjuvanted, and four doses with alhydrogel, and there were two arms with BioThrax IM, and BioThrax SQ, and so the BioThrax SQ dose, this is the first three doses of the license regime. And we used two doses of BioThrax IM as a control or reference for the rPA studies, ranging from 5-75 micrograms rPA.

This slide shows the TNA titers for the two high doses of rPA, 50 micrograms in red, 75 micrograms in black here. That's with alhydrogel, and the blue line shows BioThrax. That's the license regime, SQ, zero, fourteen, and twenty-eight days.

As you can see, the titers, the peak titers are essentially the same. The biggest difference is the boost you get here from the two-week vaccination with AVA.

These RCDCs were included in the publication. This was recently published in

Human Vaccines just this fall, and these RCDCs show the AVA curves up here, rPA with alhydrogel, and rPA without alhydrogel. And I think it's certainly interesting to note the difference that the adjuvant makes in the study.

So on to the summary already. And really, again, this is just a very quick preview, so you're aware that the study was done. All doses of adjuvanted rPA were well-tolerated and immunogenic. The unadjuvanted rPA was poorly immunogenic, and the Anti-PA and toxin neutralization antibody responses following rPA adjuvanted with alhydrogel were similar to responses following BioThrax, either SQ or IM.

So that's really all I wanted to cover on that. I'm happy to take any questions.

(Applause.)

DR. BURNS: Drusilla Burns. Ed, I don't know if a lot of people could see your

TNA slide, but did you want to comment on the titers that were achieved, just so that -- I didn't hear you say what some of those peaks were.

DR. NUZUM: Let me see if I -- I already closed it up. I thought I was done.

So as far as the peak levels you're talking about? Yes. If I remember right, they're in the thousand, the TNA levels, peak levels are in the thousand range. You want to go back, right? So the TNA levels here, peak levels are in the 500-1,000 range, and that's consistent with I think what CDC has found, and with the AVRP data. And I think it's -- to me, the interesting thing here, the rPA and BioThrax are similar.

DR. FERRIERI: For day 42, can you point out the dates for the blue dotted lines? We can't see them.

DR. NUZUM: Oh, okay. So here's the blue for BioThrax, and the red is -- I mean, they're essentially overlaid on each

other. I mean, I can hardly tell. The red is right here, and the black is underneath both of them.

DR. NASS: Two issues. The first was, this was a study also to identify safety and tolerability, but you didn't mention that.

DR. NUZUM: Right. We're not covering the safety aspects, again, because the purpose of the workshop is the correlates. And what we want to do -- and this morning's session concentrated on efficacy in animals. Now we're going to talk about immunogenicity in humans, and then there will be discussion about how we tied the two together. So it's not that we didn't do -- we don't have that data, it's just that's not a purpose for this workshop.

DR. NASS: There's a recent paper which I can't recall too many details of, and I'll bet dozens of people in the room know this paper well, which showed that although the titers, when you gave PA without an

adjuvant, the titers were much lower, that actually survival rates were similar, such that the adjuvant may be artificially raising titers without actually producing any increase in the immune response.

DR. NUZUM: So this would be survival rate in an animal study. I'm not familiar with that.

DR. NASS: I think it was out of Ft. Dietrich. I know there's a few people here.

DR. NUZUM: I'm sorry. I don't know.

DR. CHAWLA: This slide does not show the dose related to five microgram, 35 microgram, does it?

DR. NUZUM: Fifty and seventy-five micrograms. The two high doses.

DR. CHAWLA: What about five and twenty-five, because you have used five, twenty-five also.

DR. NUZUM: Right. I just didn't

put it on the -- we just didn't put it on this slide.

DR. CHAWLA: Okay. How did they react actually?

DR. NUZUM: Well, they were a little lower. I mean, there was a general dose response.

DR. CHAWLA: So you could see a dose response curve between5, 25, 50, and 75?

DR. NUZUM: Yes.

DR. CHAWLA: And that data is available in the research paper?

DR. NUZUM: It's in the paper.

DR. CHAWLA: Thank you.

DR. NUZUM: Yes. And, again, the point here is, what we want to know, and it's probably why Drusilla asked the question, what we want to know are what titers are possible in humans, so that when we look at titers in animals that are protective, can we -- are they at the same level? And so, again, it comes to back to knowing what happens in

people, so that we can show similar response in animal efficacy modeling.

Anything else? Thank you.

DR. LYNN: Our next speaker will be Conrad Quinn from CDC. And, Conrad, I'm assuming that the one in the folder was your new one?

DR. QUINN: Thank you, Freyja. Thank you, organizers, for inviting us to present at today's meeting. I'd like to begin with a couple of statements. First is, I hope my voice lasts. I'm suffering from laryngitis today. Second point is, we have prepared these data for peer review publication, so there are a few changes to the slides I'm going to show, compared to those that are in your packet. There are a few additional slides, and a few changes in terminology, but the data are the same.

So I'm going to tell you about a current analysis of a dose reduction and rate change study in humans, Human Clinical Trial,

Phase IV, of the Licensed Anthrax Vaccine in the United States, AVA, Anthrax Vaccine Adsorbed.

The background to the study is that AVA is currently the only licensed Anthrax Vaccine in the United States. It's also the only licensed aluminum adjuvant vaccine that's given subcutaneously. The immunization regime is currently giving .5 ML doses subcutaneously at weeks zero, two, and four, following up at month six, twelve, and eighteen, and then annual boosters. The principal immunogen of the AVA is Anthrax Toxin Protective Antigen, rPA.

The background to the CDC study is that data supporting the license regime, the way it's given, the number of doses, and the rate of administration are quite limited. They are based on animal studies, and a single fetal evaluation done in the 1960s.

Some safety concerns were raised following the immunizations for the Department

of Defense in the late 1990s, and a pilot study also conducted by USAMRIID and the Department of Defense, Phil Pittman, et al, in 2002, demonstrated in a smaller-scale study that a reduced schedule, and a change to the intramuscular rate of administration elicited similar antibody responses to the licensed regime, and elicited fewer injection site adverse events, or AEs.

Based on these studies, and the concern with the existing vaccine, in 1998, the U.S. Congress mandated CDC in Atlanta to undertake and implement and Anthrax Vaccine Research Program in cooperation and collaboration with the National Institutes of Health, and the Department of Health, and the FDA.

The AVRP, as we refer to it, is Phase IV Clinical Trial, post-manufacture. It's randomized, double blinded, and placebo controlled, and the objective is to assess the immunogenicity and the reactogenicity of

alternate schedules, and the different rate of administration of AVA.

The comparative analysis that we undertook are the change of rate to intramuscular. In the first instance, dropping the dose at week two, and subsequently, and data that we have not yet unblinded and analyzed, is a reduction of the booster regime.

This study is undertaken at five clinical sites across the United States, the University of Alabama at Birmingham; Walter Reed Army Institute for Research in Maryland; the Baylor College of Medicine in Texas; Mayo Clinic and Foundation in Minnesota; and the Emory University School of Medicine in Atlanta, Georgia.

The enrollment and exclusion criteria are quite extensive, and are available at the ClinicalTrials.gov website. And I've reduced them down to a few bullet points here for the sake of brevity.

Inclusion criteria were that these were healthy adults between 18-61 years old with no history of Bacillus anthracis infection, and no history of Anthrax vaccination. Exclusion criteria are also extensive, but I've reduced those to three bullet points; specific allergies, immuno suppression or history of immuno suppression, pregnancy or planning to become pregnant during the course of the study.

The evaluation criteria were based on serology and the clinical reactionicity. In terms of serology, this is a non-inferiority study, and we are looking at the geometric mean concentrations of Anti-Protective Antigen IGG at week eight and month seven. And we're also looking at the injection and systemic adverse events. For example, but not exclusive to penal injection, POI, warmth, tenderness, itchy, erythema, induration, edema, and nodules at site of injection. Adverse events were treated as

dichotomous endpoints.

For the current analysis that I'm going to present today, it focused exclusively in the first 1005 subjects, and the completion of their month six vaccination, and the month seven endpoint.

Injection sites and systemic adverse events fall into two categories, solicited adverse events, and severe adverse events. The selected AEs were predefined based on existing studies, and the pilot study conducted by USAMRIID, and these are grouped into three classifications, mild, moderate, and severe.

Severe adverse events fall into the five categories of death or life-threatening resulting in hospitalization, causing disability or incapacity, congenital abnormalities, or any medical intervention deemed related to or deemed necessary.

Reactogenicity reporting was based on scheduled and clinic examinations, pre-

vaccination, 15 to 60 minutes post injection, and one to three days post injection. We also had a 28 day follow-up for injections three and four. Enrollees and participants also had self-reporting diaries. They were given the opportunity for unsolicited reports, and there were also telephone follow-ups for all participants.

In terms of immunogenicity, which is the focus of today's workshop with the correlates protection, the primary endpoint were the NTPA, IgG, Geometric Mean Concentrations, Geometric Mean Titers, and a proportion of vaccinees with a four-fold rise in titer compared to the baseline or pre-vaccination values. It's a non-inferiority study, and the non-inferiority criteria are listed here.

The upper bound of the 95 percent confidence intervals for the ratio of the four SQ group, which is the licensed regime, to the test group's geometric mean concentrations and

geometric mean titers were to be less than 1.5. And the analogous upper bound for the differences in proportions of the four-fold responses was to be less than 0.1.

This table shows the schedule of injections. In the top line, we have the licensed regime, which we refer to as 8-SQ, where we have the zero, two, four week injections, month six, twelve, eighteen, thirty, and forty-two, and this is the termination of the study. We actually take another blood point after the 42-month injection.

Between this and the intramuscular right, we have a gradation of schedules where AVA vaccinations are sequentially replaced by saline placebo. We also have saline placebo groups that received the saline either intramuscularly or SQ.

This is the target regime, four doses intramuscularly, where we have zero, four, twenty-eight, and then the booster at 42

months. Four doses are dropped.

For the purpose of the interim analysis, we're focusing on injections up to month six and the immune response to that injection. Therefore, the immune response at month seven.

For the purpose of the analysis that allowed us to compress or reduce these three groups to one. And for the purposes of the presentation, we've renamed these as the 4-SQs, the Licensed Regime, 4-IM is the license schedule with the intramuscular right of administration, 3IM is the two-week dose dropped, and then we have the placebo controls.

The demographics of enrollment at this point in the study are we have 1,563 enrollees, and we selected the first 1,005 for this interim analysis. Mean study group size is 168, with a range of 165-170. The proportion of male and female participants are similar, 505 and 500, respectively. Mean age

is 38.4 years, median 39, and the distribution of groups of each participants across the different age groups here were quite similar, but a little lower in the 50-61 years category. Race distribution, 76 percent white, 19 percent black, and 5 percent other, and ethnicity, 95 percent non-Hispanic, 5 percent Hispanic.

In terms of the seven month analysis, these are, according to protocol, unimputed data. And what we see is that we have a very low rate of non-responders in all groups for SQ being license regime, just over 1 percent of the participants did not respond. And in the placebo group, high percentage of responders, as one would expect, less than 1 percent did not respond, or had a measurable response in the absence of vaccination.

These data are good, but they improve again at seven months, where we have no non-responders in either of the two four dose groups, .5 percent non-responders in the

3IM, and approaching 100 percent non-responders in the placebo group, as one would anticipate.

Focusing on the immunogenicity, the first thing I would like to point out is that the relationship between the concentration of antibody and the titer, the dilution of titer measured for those responses have a very high positive correlation, or a value of .99. And for the purposes of the rest of the presentation, we'll be focusing on geometric mean concentrations, or concentrations values only, and we will not be referring to the titers because they're so positively correlated.

Immunogenicity data are as follows; at week eight, at the two important time points, week and month seven, at week eight, in terms of the primary endpoints of the study, the 4-IM group was non-inferior to the 4-SQ, the license regime schedule for all three primary endpoints, geometric mean

concentration, geometric mean titer, and proportion of four-fold responders. The 3-IM group, those that did not get the vaccination at week two, was non-inferior for the proportion of participants with a four-fold rise, but there were differences in the magnitude of that response. At month seven, however, when all participants had either three or four doses, all primary endpoints were non-inferior, and I will show you the data now.

This is the serology antibody curves. We have weeks across the bottom, we have antibody concentrations in micrograms per mil on the Y axis. The vertical dotted lines are the injection points, which are vaccine or placebo at zero, two, four, and then at 26 weeks. The measurement time points are at zero, four, eight, 26, and 30. So what we see, first of all, is that we have at the first time point differences in magnitudes between the 4-SQ license regime, and the 4-IM,

rate change of 106 versus 89.7. These are not statistically significantly different, and these are -- they're not statistically significantly different.

However, in the 3-IM group, we do have a statistically significant difference. It's about 50 percent of the magnitude of response. However, by month seven, which is the response to this vaccination here, we see that all groups are providing essentially the same magnitude of response, which is statistically significantly not different, and they are non-inferior. So our interpretation of these data is that between here and here, the priming of the immune system is equivalent in all of the regimes tested.

This data show the proportion of four-fold responders. And, again, at week eight you can see that by inspection the non-inferiority is evident by the superimposition of these time points, and these magnitudes of response here, and at month seven it's the

same.

These are the reverse cumulative distribution data with the same groups. The grids here are for reference points only, and what these data tell us is that these are the point estimates for the proportion of responders that get to these levels in the different regimes. In green we have the 3-IM, in red we have the 4-IM, and in black we have the 4-SQ, the license regime. These are according to protocol unimputed data set.

Visually, using the ITT imputed data set, they are very similar visually. These are point estimate curves, and the importance of the ITT data set is that they give us a better representation of the standard errors of the data. This will become much more important at the end of the study, as we expect to have more missing data at that time point. At this point in the study, what they -- they indicate that the prevalence of missing data is most likely random, and it is

low. If you look at the month seven data, we see that the three RCD curves for the three different groups are essentially superimposable, and this is also reflected in the ITT imputed data set.

Continuing with immunogenicity, these are new data, and this is the first correlation that -- opportunity that we've had to correlate the toxin neutralization activity levels of the first 1,005 participants with their IgG levels. We see that there's very strong positive correlation between the magnitude of neutralizing power of the antibody response, and the magnitude of the IgG response.

In our study, we're doing a 30 percent subset in the TNA assay, the objective being to demonstrate or evaluate that antibody responses by the different regimes have similar neutralizing capabilities.

Other interesting facets that have emerged from the study to this point is that

there's a trend of reduced immune responses with age. These are not statistically significantly different, but it is an interesting trend. This is the eighth month data for age groups less than 30, 30-39, 40-49, and greater than 50. This is the 3-IM group, so we can see that the difference in magnitude of response at week eight is also reflected in the -- the trend is reflected in the age groups.

At month seven, we still see the same trend, with decreasing response with the -- again, they're much tighter, and there's no statistically significance between the groups.

If we look at gender-related differences in immunogenicity, you see that at week eight there is a significant difference between male and female responses, the female being 112 geometric concentration versus 73 geometric mean concentration, significantly different in both the 4-IM and in the 3-IM groups, not in the licensed regime

subcutaneous, and we see that by month seven that these differences have disappeared. So gender-related differences at week eight and month seven at the end of the regime, of the schedule.

I realize this is Carl's protection data, but I'm so tempted to go through the reactogenicity data, at least briefly. To start with, the punch line, or the bottom line, intramuscular administration was associated with significantly fewer and less severe injection site adverse events. In the first seven months of the study, no serious AES were reported assessed as causally related to the study agent by our data safety and monitoring work. And up to the seven month analysis, we have 221 reports of adverse events in 179 participants. These events related to only five persons deemed causally related to the investigational agent, and the adverse events fall into these five categories here.

The intramuscular rate, overall, by group was associated with significantly reduced frequency, significantly reduced severity and duration for local adverse events. If we break these out by gender and by time point, we do see some subtle differences. For example here, generalized arm pain, not pain on injection, was not different within males at this time point, comparing 4-IM to 4-SQ. Obviously, it was not worse, but it was not better compared to the other end points. And, similarly, arm motion limitation of these two may be related, not significantly different between the male proportion of participants in this study at this time point. But, again, it was not worse, either.

Similarly, bruising. Although, overall a study cohort intramuscular had significantly reduced frequency and severity of these adverse events, again within the male participants, the 4-IM versus 4-SQ was not

statistically significantly different, no worse, no better, but significantly better in the study group as a whole, and certainly in the female participants.

Moving on to serious adverse events, fatigue, muscle ache, and headache, again, had a gender-related difference. You can see that in female versus males, there was not a significant -- they were significantly different, but in 4-IM versus 4-SQ they were not for the male components. Similarly, for muscle ache, and headache.

So, to conclude, at this intermediate stage in the study, we still have more work to do. We anticipate there is -- we have just enrolled, actually taken the last blood sample from our last participant. We anticipate about 18 months more of laboratory work to close the study and evaluate the effect of dropping booster doses. But at this point in the study, looking at the first 1,0005 participants to month seven of the

enrollment, we can conclude that the 4-SQ, and the 4-IM, and the 3-IM regimes provide equivalent immunological priming as assessed by the response to the immune -- the immune response at month seven to the injection at week 26, and that intramuscular administration significantly reduces the occurrence of injection site adverse events.

Not only is the frequency reduced, but these are less severe adverse events. And to this point in the study, for the first seven months, there were no serious adverse events reported that were assessed as casually related to the Anthrax vaccine adsorbed.

I'd like to finish with acknowledgments of the participants. This is a cast of thousands represented here by the prominent players, shall we say, our five clinical study sites, Baylor College of Medicine, Emory University, Mayo Clinic, University of Alabama, and Walter Reed Army Institute, the branch at CDC that organized,

and sponsored, and managed the study, and continues to do so, our Biostatistics and Information Management Office headed by Brian Plikaytis, who is here in the audience today, for doing statistical analyses, and the various lab groups, and supporting agency who have contributed significantly to the course of this study.

Thank you. I'd be happy to take questions, if my voice holds out.

(Applause.)

DR. NASS: I have a few questions, but I'll start out with what were the criteria that allowed you to determine that only seven of the reactions were caused by vaccination?

DR. QUINN: I don't actually know the answer to that. That was assessed by our Data Safety and Monitoring Board, who are the only members of the team that can unblind the data, so I'm not -- these data are still blinded to me, so I do not know the answer to that question.

DR. NASS: How many reports were filed with the vaccine adverse event reporting system?

DR. QUINN: I believe all of these were filed -- with VAERS?

DR. NASS: Yes.

DR. QUINN: I believe all of these were filed.

DR. NASS: All being 179?

DR. QUINN: To the best of my knowledge, yes.

DR. NASS: Okay. Now there are other data sets that show that women have two to three times the rate of many reactions as men, and that includes studies that have been published by the Army, such as the Tripler Study, as well as the Anthrax Vaccine Expert Committee, which analyzed the VAERS reports, and showed that those that had symptom complexes that looked something like Gulf War Syndrome with headache, fatigue, and pain, also had two to three times as many women

reporting those symptom complexes as man for Anthrax vaccine. And now you're showing data that demonstrates that women are, again, having approximately twice as many systemic reactions for those types of reactions as men, and yet you're claiming that most of those are completely unrelated to vaccination. And I find it hard to understand how all these data sets show that women have a much higher rate, but somebody has determined it's not related to vaccination. Help me out.

DR. QUINN: And the question was?

DR. NASS: Help me understand how you could possibly have concluded, or how anyone could possibly conclude that this female prevalence, which has been discussed widely in the literature, and in Congressional hearings, does not represent reactions that are causal, causally related to vaccination?

DR. QUINN: Well, the objective of this study was to determine what the effect of changing from subcutaneous to intramuscular.

DR. NASS: That's not true. This is a Congress -- this is part of a Congressionally mandated research program that the CDC is carrying out, and one big part of that program is to assess reactogenicity.

I also have a question as to what time period -- you suggest that you had a 28-day time period after vaccinations for checking reactions. Was there any longer term follow-up?

DR. QUINN: I believe there was longer term follow-up, for the duration of the study. Brian, did you want to comment on that?

MR. PLIKAYTIS: The participants are given a diary so they have a 28-day period to self-report adverse events, but any SAE, any severe adverse event, no matter what time is experienced, is reportable, so there's no time limit on that.

MR. BLAKE: I'm Milan Blake from CBER. At the eight week, you say that these

-- all of the regimes that you show were similar, or suggest that they were similar. But when you look, start looking at the fall-off from that eight week to twenty-six, you see quite a difference between the fall-off of each one of them. Do you want to comment on that?

DR. QUINN: Sure. This particular plot is of the four-fold responders, so these are frequencies, and there are no data points between here and here, so the rates are driven by the two points. It would be nicer to have data points between here and here to show is this a real rate, and they are, therefore, comparable, but we don't have those time points, so there's not a lot we can say beyond this a two point line.

MR. SUTER: In this study, immuno deficient individuals and children were excluded. Do you think they could be included in a further study?

DR. QUINN: There are many studies

we could continue to do with Anthrax vaccines, be they the existing vaccine, or new vaccines coming out.

MR. SUTER: No, I mean the existing vaccines.

DR. QUINN: Those are studies that could be done, yes. I'm not aware that they're planned.

MR. SUTER: Okay.

DR. QUINN: Okay? Thank you.

DR. LYNN: Our next speaker will be Dr. Matthew Duchars from AVECIA.

DR. DUCHARS: Okay. Good afternoon. I'd like to start by thanking the organizers for giving us the opportunity to present some of our data today.

So what I wanted to go through today, slightly different approach to the one that Conrad has just been through. I'm going to really link clinical data with some of our non-clinical data, and start to consider how we can start to draw a correlate between what

we're seeing in the non-clinical experiments with the clinical trials.

So I shall start by going through an overview, some of our clinical data, and where we've got to with that. We have completed one Phase I trial, and two Phase II trials now, so the Phase I trial was a fairly typical safety study. It was a dose escalation study that was conducted in the U.S., and it evaluated four different dose levels of rPA vaccine starting at the 5 microgram level, and moving up to 100 micrograms of rPA.

Two dose schedules were evaluated, so we looked at dosing on days zero-twenty-one, and zero-twenty-eight, and we also included an AVA control cohort, as well, which is on a zero-twenty-eight schedule. There were 16 subjects in each of the cohorts that were examined.

The results of that Phase I study, and I don't want to spend a long time on that,

because I want to move on to the Phase II study, but the results showed that the vaccine was safe, and it was well-tolerated. There were no significant differences between the schedules we could observe. There were one or two issues around that we had some -- a high percentage of baseline responders, which I think was probably due to the fact that this is our first foray into clinical trials in the Anthrax arena, and our exclusion criteria may not have been quite as tight as we would have liked. And there was a fairly wide range of data that came out of that, as well. However, we did see a dose response across the 5-50 microgram range. And, in fact, the titers that were observed after the two doses were very similar to the titers that were seen in the Phase II study, as well.

So moving on to the Phase II study and design. So having looked at the results from the Phase I study, we concluded that although the vaccine was safe and well-

tolerated, we didn't feel that we had produced a saturated response, so we wanted to include a third dose to see if we could further improve on the titers that were being obtained, so we moved to a three-dose priming schedule. And we looked at two dose levels in that study, and we actually ran two separate trials to examine this, to really look at the dose level, and to look at the schedules, there was a dose finding, schedule finding study.

The first of those trials was run out of the UK, and looked at short regime, and a medium length priming regime, and looked at two dose levels of rPA. There are approximately 100 subjects in each of those four cohorts. The second trial was run out of the U.S., and that looked at a longer regime, and included an AVA control arm in it, which was under the licensed regime, so it was a sub-cut delivery on the zero, fourteen, twenty-eight-day schedule.

The AVA control arm had slightly fewer subjects in it, 40 subjects. The rPA cohorts contained about 80 subjects in each of those. And for the purposes of immunogenicity, we were looking at ELISA and TNA levels. And, in particular, measuring prior to dosing, and two weeks post dose.

So I'm not going to dwell a long time on the safety conclusions, as we are talking more about immunogenicity and correlates here, but I think it's just worthy just to spend a brief moment on it. So over 600 subjects were exposed to the vaccine, and the results show that the vaccine was well-tolerated, and really, we didn't see any significant effects in terms of dosal schedule, the number of doses that were given, any differences between males and females, or any differences between the groups, and there were no vaccine-related serious adverse events that were recorded either. So we were pleased that the vaccine, essentially, did show a good

safety profile.

In terms of the immunogenicity, so starting with the response, and I've just really concentrated here on the selected -- I've been deliberately ambiguous in terms of not giving you the dosing schedule, so I apologize for that now. It's proprietary at this stage, so I'm not really able to answer questions on that, specifically. But what I can say is that the selected dosing regime for the rPA vaccine gave very similar response rates to those that were seen with the AVA vaccine at just over 90 percent of response, response being defined as four times the lower limits of concentration for the ELISA assay that was used.

In terms of the titers that were obtained, again, very similar response in terms of level of titer between the AVA and the chosen rPA dose regime, so there was no statistically significant difference between the two cases of AVA versus rPA. The

analysis, again, being done so that this graph is really showing the titers at two weeks post the third dose.

The other point to make is that we had a fairly, as you hope and expect, a fairly typical normalized distribution in terms of responses that were seen, and titers that were seen across the population in the study, so we saw this fairly typical reverse cumulative distribution plot where you have a nice sigmoidal curve to it, showing that here at where you've got no -- a titer of zero, everybody has a titer of zero or above, and here you have a maximum titer in the last subject with the highest titer there, but in-between you have this nice typical sigmoidal distribution. And this was the same for both ELISA, as well as TNA. I'm just really showing the TNA data here today.

One observation that we did find in this study, which came as a little bit of a surprise to us was that there was a difference

in the level of functionality of the anti-body with time, and this is shown in this graph here. Again, I have deliberately removed the time intervals, and just shown it as four different times during the course of the study that were looked at. So early on in the study, you can see that here, the ratio of ELISA to TNA was almost one-to-one, so if you had an ELISA titer of, for example, 100 micrograms, you would get a TNA PD-50 value of 150. As you progress through the study, and through different dose numbers, as well, that ratio actually changed. And towards the end here, you're at a ratio of about one to six, so your 100 microgram ELISA value now translates to a 600 ED-50 in TNA. And even then, that is different to what is being seen in the animal studies, so that's another thing to bear in mind when we start to look at these correlates, is that the functionality and the way they respond in the TNA assay is different, so I think some of the data that

Louise Pitt showed us this morning showed a ratio of roughly one to ten in the rabbit, and I think the non-human primate studies. This is looking more like a ratio of one to six.

Okay. So I'd now like to spend a little bit of time just going through the non-clinical data. And, again, you have seen some of this already this morning from John Bigger, particularly, in his presentation, so I apologize if it's reiterating what you've already heard some of this morning. I'll try and be brief on the parts that you already know.

Starting with really the animal role, this is what it's all about, how we're going to correlate our vaccine's performance in the animal models to what we're seeing in humans. And Drisilla gave a fine talk this morning going through the animal role, and some of the key -- pointing out some of the key points within that guidance as to what's expected, so I think it's well accepted that

the pathophysiological mechanism of toxicity is well understood, and I hope you can see it on there. There is a reference, which I haven't included in the handout, so as a bit of an afterthought, a few references which are dotted through the presentation. If you can't quite see them, you can either contact me afterwards, or email Freyja. I'm sure they'll be able to send them out.

The other principal matter of importance, of course, is to demonstrate the effect that you're seeing in the animal species is a response that is predictive to that that you see in humans. And this is a point that Ed, particularly, was spending some time on this morning, saying we are not developing a vaccine here for animals. We're developing a vaccine for humans, and it's very important that what we're seeing in the animal models is predictive of what's going on in humans.

So bearing that in mind, there are

some interesting points that need to be taken into consideration. I think Ed also touched on this this morning. The first of those is that there is documentation in the literature that really shows that Anthrax is not 100 percent fatal in humans, and this has a profound effect on the way that you then need to treat and look at the way that you're setting up your animal models. And, in particular, there's a very good review by Holty, which I've referenced here, which is an excellent starting point in terms of a review of inhalational Anthrax cases. And he goes through some of the different levels of lethality that have been seen in human cases.

The antibody is known to be protective, and we have already seen this morning, again from Mark Perry's presentation, that in passive transfer, human IgG when it's transferred into these animal models can be shown to be protected, so that's a very important point. However, vaccination is not

always 100 percent effective. And, in fact, the field evaluation study that was conducted back in the 60s using the existing license vaccine showed that - and that was in a cutaneous environment, rather than aerosolized exposure - but it did show that people who had been vaccinated weren't necessarily all protected. It wasn't 100 percent protection. I think the efficacy came out at about 92 percent, or thereabouts.

Okay. So this logistic regression model, John Bigger spent quite a bit of time going through this earlier this morning, so I'm not going to belabor the point by going through that again. Suffice it to say that the logistic regression model has been developed using the rabbit model, and the thing that I would point out on here, and I think it was also pointed out this morning, is that when you look at this, there is a linear section to the graph which really is between the 20 percent, up to about 80 percent. And

then beyond that, this is where you start to get towards the upper asymptote, and your confidence intervals here start to get a lot wider. And, in fact, if you look at the table that was in John's presentation, that's in your handout pack, you'll see that the confidence intervals, as you get above 80 percent survival, that those confidence intervals become very, very wide. And that has a very significant effect in the way that you then start to determine what's an appropriate titer to be aiming for in your human clinical studies.

So we also covered a bit about passive transfer this morning, and I think all I really wanted to say here was that it has been demonstrated from this morning's data, in fact, that we can show that in principle, human IgG is protective when it's transferred into some of these animal models. However, there are other mechanisms involved in protection, and passive transfer, in itself,

although it can give an estimate to the level of IgG required for protection, it may not necessarily be a precise level, and may have a tendency to over-estimate that level. And as somebody very ably picked up in one of the questions this morning, one of the reasons for this is that when you passively transfer in, of course, there is no reserve of antibody in the naive animals. There is no cell mediated immunity; and, therefore, it's not quite the same situation, as an animal or a person that has actually been vaccinated.

Okay. So the last part of my talk, which I want to spend a little bit of time on is starting to think about how we can start to pull these two sets of data, the clinical data and the non-clinical data together to start to inform the program, to inform the product as to how -- as to what level of response in humans is likely to be predictive of being efficacious, providing survival. And, to my mind, there -- I've covered here three

potential ways that this could be examined, three potential ways of setting a targeted response.

The first of those is a fixed cutoff for survival. The second one is pretty much the same sort of principle, of using a fixed cutoff for survival, but taking into account the confidence intervals, and looking at the lower bound of that confidence interval. And then the third approach is to use a more population-based model, a vaccine efficacy model, as being termed here.

So what I'd like to do now is just spend a little bit of time going through those three approaches, and I apologize to all statisticians, and mathematically inclined people in the audience, because I am not that way inclined, at all, so this is my very simplistic view as to how we can start to -- how these particular methods or approaches can be used to start to draw a correlate, potentially.

So the first of those -- oh, sorry, beg your pardon. Before I get onto that, I should also talk about the similarities between the rabbit model that's being developed, and the human data that's been generated in the clinical trial. And the purpose of this slide, really, is just to show that the type of -- the level of response that we're seeing in terms of TNA, and it's blocked out in blue here along the bottom, so this TNA value from the lowest to the highest, and when that's correlated, or taken over to the human side of what we saw in the clinical trials, from the lowest to the highest, it's covering the majority of the population in human, and we're seeing the majority of the animal model, the rabbit model is producing a similar band of titers. So we're not sort of a million miles away in terms of being able to say that what we're seeing in the animal model is similar to what is being generated in humans.

So the first of these approaches

that I talked about was taking a fixed cutoff for survival, and these are just illustrative numbers that I put in here. They're not actual, they're just ones that I've plucked out, basically, to illustrate the case. And we can see here that, if we look at the -- if we take a predicted level of survival in the rabbit model of say 80 percent that we want to achieve, we take that across and say okay, what's the TNA value that rabbits require? And you can read off here, and say that okay, rabbits that have a TNA value of X or above have an 80 percent chance of survival. We can then take that same TNA value over here on the human Reverse Cumulative Distribution Plot, take it off and read across, and say okay, so 70 percent of the subjects in this case, in this particular population have a TNA value that is greater -- the same as, or greater than a level that is predicted to protect 80 percent of the rabbit study.

However, that does have the

disadvantage that it doesn't take into account any confidence intervals, and so a sort of refinement of that approach could be to look at the lower bound of the confidence interval here, which is slightly confusing, because that actually comes out, of course, as a higher TNA value. So, in other words, to be 95 percent confident that you will protect 70 percent of the rabbits, in this case, you would require a TNA value of Y, or greater. And, again, you can take across, read it off the human Reverse Cumulative Distribution Curve, and it gives you a value. And, not surprisingly, it is lower than the value for the straightforward cutoff survival taken here, because if you look at where that is, if you just take that line up towards where it crosses the GMT values here, you're actually looking at something that's nearer 90 percent, rather than 80 percent, so not surprisingly, the TNA titer is higher.

The third approach is, to use, what

I've termed a population-based approach to vaccine efficacy. And this is where you take your entire population, and here we have the human data, so this is a nice normalized distribution from the Phase II studies, showing a few subjects with low TNA values through the entire set of people in the study, out to a few people with very high TNA values here. And for each of these, you can, effectively, do the same process of going across and reading off probabilities of survival along here. And then from that, you can take an average, so you can average all of those, and work out an expected average probability of survival for the population, again, based on the rabbit model, of course.

So, in conclusion, there are -- of the three approaches that I've outlined here, there are certain advantages, and disadvantages to each of those. So the 80 percent cutoff model that I illustrated, for example, has the advantage of being very nice,

and easy, and simple to follow and understand. However, it does not take account of the confidence intervals, as we discussed. It also does not take account of the fact that somebody who has a titer that is less than the value of your cutoff still has a chance of survival, so it's a bit of a -- if you've got a titer above the threshold, hurray, you're safe forever. And if you've got a value that's below the threshold, and you become exposed, you're doomed. And that's clearly not the case. If you've got a titer that is less than the threshold, you still have a chance of survival, but probability is reduced somewhat.

So the 70 percent level cutoff model that I discussed in the middle here, has the advantage that it does account of that confidence interval, but it does still have the disadvantage that it is a threshold, and it doesn't take account that if you have a value below that threshold, you still have a

chance of survival.

The vaccine efficacy model, on the other hand, does take account of that. And it looks at the entire population, it takes into account the probability of survival, whether you have a low value or a high value, and predicts for that entire population that, whatever it is, 85 percent, 70 percent, 90 percent, whatever it is of that population would survive, given the levels of TNA titers that you would see that are distributed over that population.

So those are perhaps three approaches that we've looked at, that could be taken into account. I think what's going to very interesting over the course of today and tomorrow is to start to look at and discuss whether some of these approaches are appropriate, inappropriate, are there other ways to do it? And I think this gets to the very root of what we're trying to do here in terms of how do we actually start to correlate

what we're seeing with these animal models, and what we're seeing in the human data, and how can we start to use that to predict what would be an acceptable, and a predicted level of survival in humans.

So with the last slide, I would just like to make some acknowledgments, because, of course, it wasn't me that did all this work, not surprisingly. There are a number of other people involved, far too many to mention, but I would like to make a few particular mentions, our Medical Director, Tony Lockett, who has really overseen and driven the clinical trials, and the clinical data here. On the non-clinical side, Kathryn McNeil from AVECIA, and Di Williamson at DSTL, who have been very closely involved in reviewing the non-clinical data. Of course, those non-clinical studies have been run out of Battelle, so the Battelle staff have done an excellent job in terms of generating that data, and presenting it.

We've had some statistical support from Ann Yellowlees, who helped to develop some of these models that we've gone through here, and more laterally, though not mentioned here, Bob Kohberger has also helped with the vaccine efficacy model, in terms of determining that. And lastly, but not least, of course, we've had tremendous support from the group at NIH, and the Inter-Agency Animal Studies Group. So with that, I would like to conclude my talk, and I will take any questions. Thank you.

(Applause.)

DR. CHAWLA: Anil Chawla from Panacea Botec. In your second slide, you said that dose response across the 5-50 microgram range was there. Did you see a plateau 100 microgram?

DR. DUCHARS: Yes. The 100 micrograms was not -- it actually dipped slightly in the 100 micrograms, so whether that was significant or not, it was too few

people in the study to be able to say. It wasn't a very highly path studied, but it didn't show a continuation, an upward trend in terms of dose response.

DR. CHAWLA: And my second question is that when you claim that it's 5 microgram and 50 microgram, did you measure it at the time of administration, if there was any kind of measurement for stability?

DR. DUCHARS: It wasn't measured at the time of administration, no. Clearly, it was measured at the time that the vaccine was made, vaccine is made to GMP, and so it comes with a certificate of analysis with the value of concentration of rPA present in the vaccine. However, we didn't measure at the time, for the same reasons that I think it was John Bigger talked about this morning, that the same difficulties of being able to disassociate the rPA from the alum and be able to measure that separately, so that's something that we have been working on

subsequently since the Phase I study.

DR. CHAWLA: Thank you.

DR. DUCHARS: Okay?

MS. VOLKMANN: Hi. I'm Ariane Volkmann, Bavarian Nordic. Very nice comparison between rabbits and humans. I was just missing the scale. I mean, there was a log scale, but no numbers.

DR. DUCHARS: Correct.

MS. VOLKMANN: So in order to do that comparison you suggested, do we know what numbers we need to protect people?

DR. DUCHARS: Yes, you're right, the scale was left off, and that was quite deliberate, so I'm not really at liberty to say exactly what those values were. What I really wanted to go through today was the different approaches that can be used, and I think -- the actual values themselves are proprietary to our particular product. But I think what we need to do, and it would be useful to get people's thoughts, in

particular, any thoughts from the FDA, from the agency, is in terms of which of those approaches is perhaps the more useful, most appropriate approach to take.

MS. VOLKMANN: But any of the three approaches is only valid or doable if you know the values you need to protect people.

DR. DUCHARS: Is it? I don't know. I'm not sure that you do need to necessarily know the values, because -

MS. VOLKMANN: Well, how would you know what you need, your range? If you have to compare it to, say, let's say 80 percent of the population as protected, you need to have the value to compare.

DR. DUCHARS: Well, for the purposes of today and tomorrow's workshop, we're talking more theoretically about -- it's more of what-if scenario, so what if the value is below an 80 percent level, or above an 80 percent level? So I don't think we need to discuss the actual -- I mean, it would

certainly be nice to be able to discuss the actual levels, I acknowledge, but, unfortunately, I'm not able to, or not today, at least. But I think we can take a theoretical approach, and still get quite a lot of value from it, I hope.

DR. LYNN: Thanks. I think we're about five or ten minutes ahead of schedule, but I suggest we go ahead and take our break, and come back at the allotted time. And at that point, Bob will have one presentation, and then we will set up for the panel discussion. Thank you.

(Whereupon, the proceedings went off the record at 3:00:33 p.m., and went back on the record at 3:30:00 p.m.)

DR. BURNS: Would everyone take their seats. We're going to start again. We're going to end this session with a statistical talk. Bob Kohberger is going to talk about Analysis of Active Immunization Data: Methods to Bridge From Animals to

Humans.

DR. KOHBERGER: Well, let me start with the genesis of this talk. As I've been working with NIH, I knew we were going to get human clinical data. And the question entered my mind, we have this immunogenicity trial with all this human response data, what would a regulatory agency do with it to come up with some conclusion about efficacy? Is it efficacious? And the answer I got back, always, was protective levels, protective levels. And I began to think well, maybe there's a better, and a different way to look at it.

You're going to see that there's quite a bit of overlap between what Matthew talked about, and what I'm going to talk about. And we're not independent, in that Matthew has heard me on various NIH calls and meetings express these opinions, so there will be some overlap. Mine will be more statistical, and to my mind, you'll have a

better foundation of why you do things that you do. But then, again, I'm a statistician, and you actually -- I have an integral in some of the equations, so that if gets you all excited, you'll see that. So it is going to be more statistical.

Now from the start of this, I am going to assume that there is a model that relates immune response, in this case TNA, to the risk of disease or survival, and that model will hold for humans. That's my assumption, and I'm starting with that. And I'm not going to discuss why it's true. I'm saying if this is the model, here's how we can proceed.

So just to refresh your memory, vaccine efficacy, it's 100 times one minus relative risk. This is important, because relative risk, one way of looking at it, it's a probability of the event when you're vaccinated, divided by the probability of the event when you're not vaccinated. It could be

the ratio of two binomials, the ratio of two Poisson rates, hazard ratio, and to do this, you need a true clinical trial where you measure disease rates.

Now there are some different ways of predicting vaccine efficacy. One is the single point method, which is a protective level. And it assumes that if you are above this level, you're protected. If you're below the level, you're not protected. And in the vaccine world, they use protective levels for tetanus, diphtheria, and I mentioned this earlier this morning, H. influenza, hepatitis B. Typically, they're used in comparative trials, a combination vaccine, versus vaccine given separately. They will compare the percent above 1.0 for a combination vaccine, and the vaccines given separately, and if they're equivalent or non-inferior, that's acceptable. So they will compare the percents above this protective level. And the reason they do this is that the percent above the

protective level is essentially equivalent to vaccine efficacy, so if 95 percent are above the protective level, it's 95 percent efficacious. We're going to look more into this.

Now you can also use the continuous relationship, and you've seen this morning this logistic regression. It's one way of doing it, and, again, the examples where it's been used is in human trials, in pneumococcal conjugate, otitis media, colonization, it's been used in pertussis, acellular pertussis. You can also use survival models, Cox regression, other parametric ones, and it's been used in varicella.

Okay. If you're going to use a protective level, and you want to estimate what the protective level is, you have to do an ROC analysis, Receiver Operating Curve. This is sensitivity, and specificity. You have to pick the level that gives you an acceptable sensitivity and specificity, and

look at positive predictive values, negative predictive values. If you're going to use a continuous relationship, and estimate a protective level, you need to fix the probability. Well, what probability of event are you interested in, 90 percent survival, 80 percent? And you can estimate that from a logistic model.

Now I'm going to give an example with real data. It's going to be the same data set, and I'm going to look at protective levels, and I'm going to look at this continuous relationship. The data set is essentially what John Bigger showed you this morning. I may have combined some additional experiments. It's unaudited in that it hasn't gone through all the big QC process, but the logistic regression that you see here on the screen is very similar to what we see in the Anthrax Challenge experiments.

Like John, these are binned values, so this is the actual percent survivals in

this TNA range, actual percent here, and this is the fitted logistic regression. So this is a curve that I'm going to be using throughout. So with this data that we have, from logistic regression, I can predict if I want the probability of death, the TNA value is 176, so 80 percent survival with a TNA level of 176, 10 percent is 397. Notice, these numbers are very similar to what John presented, not exactly the same, but very similar.

If we look at a protective level, these are the levels and sensitivity and specificity for each of these. And, remember, sensitivity is being greater than the level given that you survive, so if we use a protective level of 714, the sensitivity is 15 percent poorer. In other words, many survivors have a value less than 714, and only a few are above it. So if we're going to choose a protective level, here are two choices, 176, 80 percent probability of survival. It's got good predictive value and

specificity. It's got somewhat low sensitivity, it's only 67 percent subjects below the level are going to survive. How about this level this high? Now we've got 95 percent survival, but very low sensitivity, only 15 percent. So how are we going to choose a protective level using typical constructs of sensitivity and specificity. It could be 176, it could be 838.

The next problem is, once you choose the level, now we've got our human subjects, what percent of the subjects should be above this level? Should it be 80, 90, 95? Well, I'm not sure, and a regulatory agency, either in the U.S., or anywhere around the world, is going to have to make two kinds of decisions, what should the protective level be, that's number one. And number two, what percent of the people should be above this level, should we require 80 percent, or just what should we do?

That seems to be a problem. I

mean, now somewhat facetiously in my experience with regulatory agencies, the fewer decisions they have to make, the better off the manufacturer is. You don't want them to have to decide too much, because it takes too long. So can we help them? Can we do something to help the regulatory agency? And this is what I think.

If we use logistic regression now, here's the mathematics, there's an integral sign there. The logistic regression gives us the probability of death at a particular immune response. We also have the probability distribution of the immune responses in humans, so if I multiply and integrate, I can get the probability of death in the population, which very simply is just an average. I mean, what you do for your human responses, at that particular response you have a probability, and you just come up with the average. If the probability of death in an unvaccinated subject is one, which it is in

our challenge studies, then the vaccine efficacy is just one minus this average probability.

So using the logistic regression, I don't have to figure out what the protective level is, I don't have to figure what percent are above it. All I have to know is what vaccine efficacy do you want, 80 percent, or 90 percent. I've cut the decision points in half, which is a good thing.

Now how does this work, and what would it look like? Well, first of all, I took a simulated sample of human immunogenicity, and it's a sample size of 200, which is typical of a lot of the immunogenicity trials. Importantly, the standard deviation is .7, and this is on the basis of a log transformed TNA result. And this came from about three different papers, and I sort of took the average of three. That's about how humans vary in that immune response. And if I have a certain GMT now for

my simulated sample, and my estimated logistic regression, I can predict vaccine efficacy. And from my simulated sample, I can tell you what percent are above a certain level.

Important to look at, look at this one. If we have a GMT of 300, in other words, if the vaccine manufacturer makes a product where the GMT is going to be 300, we would predict 85 percent vaccine efficacy. That looks pretty good. Seventy-six percent are going to be above 176, which is one possible protective level, but only 9 percent are above 838, so just looking at this, I'd say 838 doesn't look like a good value when I'm getting 85 percent vaccine efficacy. But you notice what I'm doing really is putting my bets here on vaccine efficacy. That's what I'm looking at, and that's what I think we should look at to say what a good GMT is. And if you're going to develop a product, you can get 85 percent efficacy with a GMT of 300.

Here are some more protective

levels. The GMT to get 90 percent above this, you'd need 433 as a GMT, and you get a vaccine efficacy of 90, so a manufacturer can go back and forth on this with a protective level and a vaccine efficacy to estimate what the GMT should be in his product.

Now that's all very nice, but what do we do about uncertainty? If you remember back to John's slides, the confidence limits on the logistic regression were pretty wide, and if we estimate a protective level of 176, how certain are we of that? Well, if you're going to do a protective level, Matthew mentioned this, and you put confidence limits on that level, and for 176 the upper limit is 438, for probability of death here in 400, you would need a protective level of 1,300. And if you recall back here, to get a protective level of 1,300, the product is going to have to deliver an average of about -- well, it's over 2,000, isn't it? That's an impossible task, and unnecessary. So in my opinion,

basing this on protective levels with appropriate corrections for uncertainty is not really a feasible approach to product design, or in my opinion, product registration, but the panel and other people can discuss that some more. So that's what happens with uncertainty in the protective level.

With the regression analysis, we can also come up with a confidence limit on vaccine efficacy. And in the notes, and actually I skipped through it on the slides, there's a paper by Ivan Chan from Merck, where he has applied this idea to varicella vaccine. And this idea here of bootstrap confidence limits, it's not my idea, although I like it, and I think it's really good, Ivan is the one that's published it, and it's on one of the slides previously.

Well, the vaccine efficacy we have, the variability in it, it's a function of two things. The variability in our logistic regression, after all, we've only got 20

subjects at each of these doses. We don't have a lot, and those confidence limits were pretty wide.

The second source of variability is in the immunogenicity sample in humans. We have 200 subjects, that's pretty large, but there's still variability there, the sample from the general population. So to get confidence, you can do a bootstrap confidence, which is basically a simulation approach, and it happens in two steps. You estimate the logistic regression sampling from the data that generated the logistic with replacement. Now you have a new logistic regression equation. You estimate vaccine efficacy using that with a bootstrap sample of the immunogenicity sample, and you just do that repeatedly, and what you get now is a population of vaccine efficacy estimates, and the .025, .975 limits are the 95 percent confidence limits from bootstrapping. And you may have to do this 5,000 times, 10,000 times,

but a computer is doing it, so you don't have to worry about it.

So I looked at this, and I just did a single bootstrap. I just used logistic regression. I didn't bootstrap the immunogenicity sample from people, and the reason I didn't was in the interest of time in getting the slides together, number one, but more importantly, in my experience using this, the immunogenicity variability adds very little to the confidence limits, maybe two or three percent. Almost all the variability is coming from the logistic regression, so these are approximate limits that you see here because it's only a single.

Again, my immunogenicity sample, that was the standard deviation. I had 200, and I did the bootstrap with 1,000 replications, and here's what you get. The GMT of this simulated sample is 150. With 150 and our predicted vaccine efficacy, you get 76 percent. I went through the bootstrapping.

The bootstrap median estimate was 76 with confidence limits of 65 to 86. That doesn't look too bad. It certainly looks a lot better than the confidence limits on predicted values. And if you go up to the GMT of 300, so if you have a product that on the average gets a mean response of 300 TNA ED-50, the predicted efficacy is 85 percent, and our confidence limits are 78 to 93. So now, to my mind, I could say to a regulatory agency, or they could say to a manufacturer, rather, just prove that your vaccine efficacy is bigger than 80. Well, the manufacturer knows that he's got to design his product to get above 300, and that's relatively simple criteria, and it's a reasonable GMT.

So, to summarize, if you use a protective level, it's very simple to calculate. All you do is look at the percent above it, and it's easy to understand. It's difficult to set specifications on it. You have to worry about well, what level should

you use? What should the sensitivity specificity be? Then you have to set a specification on what percent of this immunogenicity subjects should be above this.

The third thing is, it doesn't capture all the vaccine's efficacy. As Matthew said, and as I've shown on some of this, you can be below the protective level, and there is still some reduction in risk, so it doesn't capture all the efficacy. If you use a continuous relationship, a logistic model, it's more difficult to calculate. I wouldn't say it's difficult, you just have to be able to program a little bit in various packages to get bootstrap confidence limits, which is more difficult.

Specifications are rather simple. They're on vaccine efficacy, whether it's proving it's bigger than 80 or 90, you just have to choose what you want to do. And it does capture all of the vaccine's efficacy, assuming the model is correct and applicable,

rabbits to humans. So that concludes what I have to say, and any questions?

(Applause.)

DR. LYONS: Robert, is -- I know these are probably calculated on individual experiments, but in general, we get a lot of bang for our buck from vaccines because of herd effect, too. Right? In most cases. The Anthrax is -- well, it's unlikely, as far as we know, there's going to be any herd effect. It's either a yes or no event. Does that change the way you handle your comfort zone, and what you want to get after or not?

DR. KOHBERGER: Well, in general, for things like pneumo, pneumococcal is the one I'm most familiar with, and the herd effect is that a lot of older people now have reduced risks of pneumonia because probably their grandchildren are vaccinated. So if there is a herd effect, all that I've seen is it increases vaccine efficacy.

DR. LYONS: That's right, but it

kind of allows for -- we all know that, so we can accept a lower level, in our minds, anyway, I could. Would you be more constricted in your looking for a higher level of efficacy say for an Anthrax vaccine versus a vaccine that's likely to have a herd effect?

DR. KOHBERGER: I'm going to repeat how I understand your question; that for vaccines that have herd effects, we're willing to take more risks on what efficacy should be because we think that there's going to be a herd effect.

DR. LYONS: Yes.

DR. KOHBERGER: And with a vaccine like Anthrax, where there probably is not a herd effect, I would think that you would be less willing to take risks on efficacy because it's not going to be bigger. Yes. Yes?

DR. NASS: Meryl Nass. I wrote a review article on Anthrax vaccines that was published in March 1999 in Infectious Disease Clinics of North America. And in that

article, I discussed at some length the differences between naturally occurring contagious diseases, and the use of this vaccine, which is intended for a non-natural occurring event that is not contagious person-to-person. And so if you start really looking at the reality of what we're talking about protecting against here, first of all, there will be, of course, no herd immunity. But there will also be confusion in the minds of the vaccinated that they are protected if they come down with a flu-like illness, which is the prodrome of Anthrax, and so you could be causing people to be confused, and their physicians to be confused if you tell them that this is a very effective vaccine.

On the other hand, this will be a political event, and if you're vaccinating people, you cannot tell them we're giving you an Anthrax vaccine. It has an 85 percent chance of protecting you. People will not be very happy, and they will not be marching up

to get the vaccination.

You also have to consider that in biological warfare, you may get altered and/or particularly virulent organisms. Paul Miransev published a paper around 1994 or `95, in which he took PA completely out of the Anthrax bacteria, and put Cerolysin in instead, which allowed the other toxins to enter cells. None of the PA-based vaccines that we've discussed today would be of any use at all against that type of Anthrax construct, so it seems to me that if we had a vaccine that was perfectly effective against all forms of Anthrax that could be designed, and used as a biological weapon, it would be a perfect deterrent because nobody would obviously use Anthrax against a vaccinated population.

What would they do? They would go ahead and make or find something else to use. So what we're really talking about here is a vaccine that will, if it's used, will probably never be needed because it is primarily

serving as a deterrent. But it will only shift the equation in a different direction to a different pathogen.

DR. KOHBERGER: Well, there's only one comment that I think applies to what I've talked about, and that's when I say vaccine efficacy, remember the models that were based here is - don't forget this - it's efficacy just before challenge. I haven't talked -- this model that got set up doesn't talk about duration, doesn't talk about efficacy two years after vaccination. It says that if you are exposed shortly after taking this vaccine, your risk is reduced by 85 percent. Whether that's good or bad, well, after you're exposed, what's you risk if you don't have the vaccine? So that's the only comment I have, and that's the only thing I think applies to what I've said. Anything else?

DR. BURNS: Thanks, Bob. I think we'll now go into our panel discussion, and if I could ask the panel members to join me up

here, and we'll just take about a two-minute break.

(Whereupon, the proceedings went off the record at 3:58:43 p.m., and went back on the record at 4:00:21 p.m.)

DR. BURNS: Okay. I think we will start now on our panel discussion. First, I would like to introduce the members of our panel, starting down here at the far end, we have Steve Self from Fred Hutchison Cancer Research Center; Don Rubin from Harvard University; Emil Gotschlich from Rockefeller University; Pat Ferrieri from the University of Minnesota Medical School; Rick Lyons from University of New Mexico Health Science Center; and Eric Hewlett from University of Virginia School of Medicine. And we want to thank all of you for coming and participating.

We have heard today and outline of strategy that has been being followed in order to get the data to support efficacy of new generation Anthrax vaccines. And we wanted to

hold this workshop as sort of a mid-course reality check to make sure that -- are the studies the appropriate ones, the data that's coming out of them, how do we really now take those data and go from the animals to the humans? When you start getting into the details, it becomes a little more difficult than in 2002 at the workshop when we just had sort of grand ideas how to do it, and sort of glossed the surface of how to do it, but now we have the data, and how do you appropriately extrapolate from animals to humans?

So we have a number of questions that we are going to address to the panel. However, I would like to say if members of the audience have additional comments that they would like to make, just come up to the microphone, or if you have additional questions for the panel for clarification, then just feel free to ask them.

So I'll go through each of the discussion points, and then we'll take them

one-by-one. The first one is comment on the soundness of the design of the animal studies. Since this panel discussion is really focusing only on general use prophylaxis, as far as post exposure prophylaxis, we'll address that tomorrow.

The second discussion point is comment on the strengths and limitations of using toxin neutralizing antibodies to extrapolate titers, to extrapolate animal protection data to efficacy in humans. Then comment on the strengths and limitations of active immunization and passive immunization data in defining the correlate of protection. Comment on potential approaches for inferring protective efficacy in humans from animal data. For example, establishing a protective antibody titer, or cutoff level, use of alternate statistical methods to estimate predicted vaccine efficacy, et cetera. We heard a number of different approaches today. And then, finally, what additional data, if

any, are needed to strengthen the extrapolation of GUP protection data in animals to efficacy in humans?

So we'll just start with the first discussion point. We heard about a number of animal protection GUP studies today, and we saw what the general design was. And we would just like the panel to discuss, or to comment on the soundness of the design. Is anything critical missing, et cetera. Does anybody have any -- Pat?

DR. FERRIERI: Shall I start by kicking off on this question? I'd like to take a very general and brief step here, by commenting on the historical past, in the past, say five to seven years, I got initiated in this through an Institute of Medicine Committee, as did Dr. Gotschlich, sitting to my right. And at that time, there were many, many things that hadn't been done. And then I participated in the Blue Ribbon Panel NIH convened, so I got to hear about all the

developments of the two companies that have presented today. And now I'm hearing further work since then, so I'm very pleased with the general direction of the design of the studies that have permitted us to see dose responses in animal models. We've been able to see timing and intervals numbers of immunizations, and some of the work compared to the only currently approved vaccine, the AVA. And most importantly, for me, was presentation of data using passive protection of antibodies from either one species, and then the protection in another model, something that some of us argued for very forcefully with CDC, if I recall correctly, several years ago at the Institute of Medicine. So, in general, I see us moving forward, seeing all of you move forward in a relatively sound way. But I think that I would like other members of the panel to comment on their take on it, the positives and the negatives on this point, Drisilla.

DR. BURNS: Erik?

DR. HEWLETT: Yes. Thank you, Drisilla. I need to start with a disclaimer. Rick Lyons and I are pathogenesis types, and not biostatisticians. And Drisilla told us not to talk about pathogenesis, so we're a little stuck here.

DR. BURNS: No, I did not. (Laughing.)

DR. HEWLETT: And I also can't help but taking my experience with pertussis into consideration when I think about this problem. And what struck me about the design of these studies, and the data that are collected is that I believe relevant here, as in the case with pertussis, individuals are immunized. They then are exposed, and have an anamnestic antibody response by virtue of having been immunized, and those individuals then, depending on that antibody response, don't get sick. They get rid of the infection.

I think what's missing here from my

mind is we're looking at antibody levels at a bunch of different times after immunization, and before challenge, but not looking at what happens in individual animals that are challenged, and then either survive, or don't survive. And I acknowledge the difficulty of measuring antibodies in the non-survivors, but the point is that the magnitude of that response, seems to me, is a very important variable here that's more relevant perhaps than what the titer was some time several weeks earlier. And I understand Conrad may have some data like that from humans, and maybe you can talk about it, either now or later.

DR. BURNS: Conrad, would you like to comment?

DR. QUINN: Conrad Quinn, CDC. We have previously published on follow-up immunological studies on the survivors from the 2001 air attacks, and we have demonstrated that Anti-PA specific IgG is circulating and

detectible up to 16 months post infection. And, also, memory B cells are circulating and functional, and detectible up to 14 or 16 months after infection. In recent tests, surviving animals, be they naive and survive due to innate immunity or some other mechanism as yet unknown, and vaccinated animals that have survived challenge, we have followed the antibody responses in those animals, and we see good anamnestic responses. The maximum responses being similar to those seen at week 30 in vaccinated animals, so those are the extent of our data at the moment. I believe others may be looking at phenotypic cellular responses post challenge, but I don't have those data.

DR. BURNS: Conrad, can I just follow-up on that? I mean, did you do a careful look to see if the anamnestic response correlated well with vaccine titers? I guess you said 30 weeks, was there a good correlation with the animals that were

protected?

DR. QUINN: Yes.

DR. BURNS: Okay.

DR. QUINN: Those are the data that David Madigan showed this morning, and in terms of both the ELISA measured IgG levels, and the TNA measured levels at weeks 30 and 34, the maximum responses are well correlated with survival.

DR. BURNS: I'm sorry, but the anamnestic response, did that correlate with an earlier time point, too?

DR. QUINN: We haven't looked at those data.

DR. BURNS: Okay.

DR. QUINN: We're still analyzing those data. So that's a good question, and we will make sure we get an answer to that.

DR. LYONS: Actually, Conrad, I think I saw on David's slides that, interestingly, the two -- I think there are two primates that were vaccinated on the low

level, and they died. But, to me, it looked like their amnestic response was quite good, actually, so it would be worth taking a look at that to see how that falls out.

DR. QUINN: We certainly will. Nothing -- by simple inspection, nothing has jumped out and bitten us, and said this is it. It's quite clear that animals that were vaccinated, you can detect the onset of the response at day five and day seven quite clearly, and it peaks at day 14. In animals that survive due to innate responses, we don't see a measurable response until about day 14, which one would expect from first exposure to infection, so there's no correlation there. The speed of the onset of the amamnestic response, we haven't looked at that.

DR. FERRIERI: I was going to just say that one of the specific points I would make that I heard is being planned, and would really give us the kind of information we need is the length of protection, because if you

scrutinize a lot of the animal data carefully, you see the titers decrease sometimes within a matter of four weeks, and so how long is protection? I'm happy to know that such studies in animals are being planned. This is still part of the general design of the GUP studies. Emil is being uncharacteristically silent, so I'm waiting for a bomb to fall.

DR. GOTSCHLICH: First of all, I'd like to second what Pat has said, in terms of the evolution of the field from the moment that I got dragged into it, as did she. And there is obviously a great deal of data that has been assembled, or gathered. And for a non-statistician, and one that seems to be relatively impenetrable to becoming one, there is a very interesting discussion been raised today on more sophisticated ways of integrating this data into what is really the thing that is needed here, a robust understanding of the new response in terms of protection, but what has not yet been done,

and which are sort of missing, is to hear what actual -- what would happen if these newer techniques were applied to the existing data. And, perhaps I've missed it. There was certainly some in David's talk, but not in an overall way.

DR. SELF: I guess maybe I'll chime in here, because I disagree a bit with what you said. I think it's a very interesting problem trying to integrate thresholds or survival curves derived from animal models with human immunogenicity data, and how to synthesize that, but that's not really the hardest problem here. I mean, the hardest problem here, I think, is that what's been referred this morning to that leap of faith. We've got some data from rabbit studies. I think the design for the question here are quite fine.

I found the analysis and presentation difficult to follow. I think that just some very simple things could be

done in terms of having a simple standard analytic method that's applied to all of the data, so that they can then be displayed on, at least, a similar scale and compared. This is certainly going to be required when the non-human primate data are available. And I think it's all going to be about the sources of variation that we see within species, and between species, and how to display that, and understand that. And it's not going to be a highly technical discussion, I think, involving -- I think it will involve some very deep concepts, and I'm sure Don is going to talk about some of that, but in terms of the analytic methods, it's not going to be complex. So that's kind of where I'm seeing the important focus needing to be.

DR. BURNS: I mean, I think that's an interesting point about variability between individuals of a species, versus between species. And that is something that really hasn't been looked at very carefully that I

know of.

DR. SELF: So just the rabbit data that were shown this morning, there were variations in regimen, challenge dose, timing, there are a variety of things, and I couldn't quite put together -- I mean, let's say that you just use a logistic curve as sort of the fundamental analytic unit, and you take, pick it, your favorite antibody level that corresponds to 90 percent survival, just to make it really specific and simple.

It was hard for me to take those estimates with the intervals about them indicating what statistical precision, or lack thereof is, and line those up, and see how consistent they are. I think I saw some consistency, but it just -- we just need to do that. And once that's done, to bring the same sort of data in for non-human primates.

Now I think there are going to be two kind of patterns that might emerge once you bring that second kind of data in. Either

things are going to line up pretty well, in which case, yes, maybe then that leap to humans, you start getting some comfort about. If they don't, then I think it's important to start thinking about other non-human primate animal models. If you have a second species, non-human primate species, does that line up with the macaque data? Does it not? That's the variability that's going to either give you comfort, or no comfort at all in leaping to humans.

DR. BURNS: I mean, do you think you would have to have a third species, or could you take a conservative approach? I'm from the FDA, so that's what we always do. And just take the highest antibody level, and say that that's -- and extrapolate that to humans?

DR. SELF: Well, we heard this morning that legally another species isn't required by the rule, and again, I think it will depend on what those data show, just how

variable it is within, and between species. At the very least, it would be nice to explore some of that variability within the macaques if there's any question, rather than having a single point, data point there. So I think we just have to start seeing the data and talking it through.

DR. FERRIERI: My impression is that relatively small numbers are included in all the non-human primate studies that were presented, and I appreciate the difficulty of boosting up those numbers, but that would be very helpful, and I think essential in order to draw any firm conclusions in extrapolating from the animals to humans. It was intriguing data, but not quite there for me.

DR. BURNS: Okay. Does anybody who is actually doing the studies want to comment on boosting up numbers, on the feasibility?

DR. NUZUM: See if I can do this again. I think the simple answer is yes, they are feasible. And I think we can get

reasonable numbers with reasonable powering. As we do more studies, and get more data, we can focus in on fewer larger groups, rather than more smaller groups, as we do dose ranging, and data gathering. So I think it is feasible, in general. It's just that we aren't, in NHP group in studies, in general, we just aren't as far along, as far as having as many clean and complete data sets, so that's why we aren't talking about NHP data on this workshop, but we definitely intend to do that. And we agree that that NHP data will be important.

And just as a general guideline for our rabbit studies, some of the larger groups, the largest groups are usually Ns of 20. For NHP studies, they're typically six to ten per group, but they could be larger if we used smaller, I mean, fewer groups, so it's just a matter of refining studies as we go along.

DR. LYONS: I just want to comment that I really appreciated Robert's discussion

on the statistics. I hadn't really thought about it, but it seems like, and I think the designs were fine for the question being asked regarding trying to show correlates of protection with a TNA. And I think the passive transfer is very strong, myself, for showing that that does correlate with protection, because that's sort of a worst case scenario, as far as I'm concerned. But it seems like, looking at Robert's terminology and definitions, which I appreciate very much, Robert, it seems like we're someplace between four and three. And this whole process of developing a correlate or a surrogate, to me, is where research comes in. It's not going to be a static process. I realize most research is done trying to prove -- get to two. After about 30 years you might get there, but unlikely. So I think the question is, in my point of view, are we satisfied with this as a reasonable correlate at this point? And knowing that the work is going to go on, and

we're going to keep testing theories that will, hopefully, get to two, but it may not. And I think that's the real question today, is are we comfortable with the TNA, or an equivalent for a reasonable correlate. And I think, personally, I think we're there.

DR. BURNS: Maybe that just leads us on to the next comment, where we really can start talking about the strengths and limitations of using TNA antibody titers to extrapolate animal protection data to efficacy in humans. And I think that there's a couple of issues here that I would like to hear the panel's thoughts on.

We heard about TNAs, we heard about ELISAs, and I think that one of the major issues here are, are these ways of measuring antibodies independent of species enough that we could use them to extrapolate from one species to another? And I think that's a very important point that I'd like to hear people's thoughts on.

DR. RUBIN: I'd like to make a more general comment, if I may.

DR. BURNS: Okay. Sorry.

DR. RUBIN: I feel like I have a lot I want to say, but I'm not sure how to say it coherently for this audience, so I was sitting here trying to think about that, but I'll try. Robert, in the last presentation, pointed out that what we critically want is probability of survival given, and I'm going this in a particular way, if we knew sort of the dose of the vaccine, the immunogenicity, which is a function of the dose of the vaccine in humans. And we'll never get that, because we're not going to go around challenging humans in any kind of randomized experiment.

But what we do have are probability of survival I guess in a variety of animal models, macaques, rabbits, maybe some other animals, guinea pigs, as a function of those same things, immunogenicity, which is a function of dose immunogenicity as measured by

IgG, for example, and that particular animal, in that particular animal.

Now there are two problems with trying to go from one probability to another probability. And the most obvious is the dependence upon a different animal. We want it in humans, we're never going to get it in humans. We have data in different species. That's always going to be a leap of faith, because we're never going to get the animal data, we're never going to get the human data. It certainly would be nice to have it in species that are close to humans, like macaques. It would also be nice to have it in a variety of species that sort of close to humans, and see that those probabilities don't change. They don't change across species that are close to humans, maybe they won't change from species to humans. That's an extrapolation, but it's the only thing we can do.

There's another problem, and that

actually came up this morning. I think you mentioned it, that this idea of the meaning of dose, if it's the same dose, has a different meaning for different animals. Why is a dose that's right for a 200-pound man in the military the same as the right dose for a 2 kilogram macaque? Does it make sense? And we talked about that five years ago as an issue.

Well, how do you get around that problem? Well, you do have some chance of getting at data on that problem. What you want to see is in the animal data where you do have survival, and dose, and immunogenicity measured, is you'd like to see, in some sense, that survival only depends upon immunogenicity. It doesn't depend upon dose.

In order to make that formal, there's a lot -- there are a tremendous number of mistakes in statistics, and in biostatistics even now being made about this issue of some time it's called surrogates, sometimes it's called direct and indirect

causal effects. And it's very, very subtle problem, because even the great statistician, Ronald Fisher, made mistakes on this throughout his whole life, and "Statistical Methods to Research Workers", when first published in 1925, to the last edition, fourteen editions later, and "Design of Experiments", published in 1935, to the last edition that was published, made the same mistake, misunderstanding this issue. So the fact is that it's not an obvious issue at all.

Actually, could I possibly look at page 6 of Louise's presentation? Is it possible to put that up?

DR. BURNS: Page 6.

DR. RUBIN: There were two figures on that page.

DR. BURNS: I'll be working on it.

DR. RUBIN: What that figure shows is, it's an experiment where this -- I think with rabbits, where you randomized dose, and you measured immunogenicity. And what you see

-- and then there are dark diamonds for the animals that survived, and open diamonds for the animals that - or the other way around - dark diamonds, I think the animals that lived, and open diamonds, the animals that died.

DR. BURNS: Is that the one?

DR. RUBIN: Yes. That will do. Maybe the next one is a little bit clearer.

DR. BURNS: The next one?

DR. RUBIN: Yes. Okay. So what you'll notice in that is that what you're randomizing to the different animal groups are the horizontal axis, so there are different dose levels. And they start from the highest to the lowest dose level. But you'll notice that there is an overlap between the groups in their immunogenicity levels in the vertical axis. And you see, also, that there's a survival benefit, so it's easy to get inferentially the benefit of the thing that's being randomly assigned, dose, on survival. You just compare the survival benefits in

those different vertical bands.

But in order to get an inference that immunogenicity is causing that, you have to make an explicit or implicit assumption, and I always like making the assumptions explicitly, so you know what you're doing. You have to make the assumption that within a level of dose, nature is randomizing immunogenicity level. And, if so, then you can draw an inference about the causal effect of immunogenicity on survival, and then at least know that in this animal model, the probability of survival depends upon immunogenicity, and doesn't depend upon this thing called dose that has different meanings across different species, because they're way different amounts. They metabolize it differently, all the rest of that stuff.

But if you can make this assumption that nature has randomized for you immunogenicity for different levels of dose, you can make the inference. So those are two

very different kinds of assumptions that are needed in order to make this bridge from the animal data to human data. First, you have to make this assumption that it's not going to depend upon the species, if it depends upon just immunogenicity, and then try to show that it only depends upon immunogenicity within the animal data. And that's a difficult thing to do, because there's no direct evidence in the data to support that conclusion. That also has to be something of a leap of faith, but less of a leap of faith than bridging from a different species to humans.

Now in that regard, the graphical approach, which Robert talked about in his first presentation, I find hopelessly seductive, and deceptive. I think it's absolutely the wrong thing to do. It's absolutely -- it's one of the reasons why people, brilliant people like Fisher got the wrong answer. David Cox has made the same mistake, nothing but the -- he's the smartest

man I've ever met in my life, but he gets that wrong. And that's why sort of Prentice's Fourth Criterion, that you have to have perfect prediction, essentially, that will never hold. Well, Steve, you have one example where it holds, I guess.

DR. SELF: I have one example.

DR. RUBIN: One example, but you can't hope for that in this field, that you get perfect prediction of survival, that people -- those above this, everyone survives, below that, everybody dies. That's hopeless, so you have to, instead, buy into nature's randomization at some level.

And just a comment on this graphical approach from Judah Pearl's, that's on number 20 of Robert's, I think,

DR. BURNS: Of his?

DR. RUBIN: Not page 20, but number 20.

DR. BURNS: Okay. Second?

DR. RUBIN: Of his first.

DR. BURNS: His first.

DR. RUBIN: And this graphic came right out of this Judah Pearl article that was published in Biometrica maybe a dozen years ago, something like that, where you use these arrows to represent cause and effect.

DR. BURNS: This one, Bob? The other one?

DR. RUBIN: And it's really pretty to look at, but it doesn't work. And the reason why it doesn't work is the middle point there, which is like the number of worms after I guess you spray or you don't spray a fertilizer or insecticide, something that's actually an outcome variable; and, therefore, it's two variables. It has one value if you're assigned to the active treatment, another value if you're assigned to the control treatment. So, in general, these pictures are just simply deceptive. I think Steve agrees with me.

DR. SELF: Yes.

DR. RUBIN: And you just have to think about it the correct way. You can't think about it this way at all. In fact, the two greatest experts on causal graphs don't agree with how to interpret such things, Judah Pearl and Stefan Lourdes in England.

Now one more thing that I want to say, that I recently worked on a problem very similar to the Anthrax one for Novartis with Lou Shiner and Jerry Needleman for a submission to FDA, where you had to do bridging, not bridging across species, but bridging between adults and children in an anti-epileptic drug that had been approved in adults for adjunctive therapy and monotherapy, both based on randomized trials. It had been approved for children as an adjunctive therapy in randomized trials. And Novartis wanted to get it approved as monotherapy in children, but it's considered unethical to do a randomized trial, to a kid, you're six years old, your parents say we should randomize, to

take you off any treatment at all, and treat this drug versus placebo, unethical.

But, nevertheless, it was approved by FDA about two years ago for monotherapy in children based on arguments that are like the ones I was giving earlier, on these assumptions, and based on, basically, models of dose response after conditioning on lots of covariates. One thing that makes it clear to me in the humans, in order to make these kinds of models correct and apply them correctly, you're going to have to have lots of covariate data. Maybe rabbits in one species don't vary much in age, height, weight, the way they metabolize vaccines, and so forth, but people do a lot, for example, depending upon what other drugs they're taking, how old they are, whether they're pre or post menopausal. All those things matter for people, and you get different survival.

You may -- you'll get different -- you would think that you would get

different survival curves, and that will be a trick to try to get the survival given immunogenicity levels for humans, it's going to have to also depend upon covariates in order to make those models fly. So, in a general sense, I do agree with what Robert was saying in the last session, but I don't think - and this agrees with what Steve was saying, I think - that the issues are not in the technical details of models, where it's logistic regression, probit regression, tobit regression, there's zillions of regression models, and those are technical details where they're smoothed or not. Those are all details only statisticians could love.

I think that the real issue, though, is the conceptual ones, how you formulate models, and enough variables, how you collect enough data that make these models plausible to scientists. And I believe there's still some work that needs to be done.

And I'll end with one last, more of

a question. For example, since we have to make an assumption somewhere that nature is randomizing immunogenicity, why haven't there been studies done with animals, like rabbits, or macaques, where you randomize dose, but you do more than that. That's just an initial dose, and then you randomize immunogenicity, and you titrate to that level of immunogenicity. So you have two kinds of experiments, one where you measure survival in a randomized dose experiment, and you try to relate survival to immunogenicity. And an other experiment where you randomize immunogenicity in a titration experiment, and you measure survivor's immunogenicity where you get true dose immunogenicity, and see if you get the same answers from both of them. That would be direct evidence on this other assumption. So I'll end with that question.

DR. BURNS: Okay. Thanks.

DR. RUBIN: Very, very long comment. Sorry.

DR. BURNS: Pat, did you -- okay.

DR. FERRIERI: Well, I -- those are stunning comments, and the audience and FDA seem, and NIH does not seem willing to take you up. I'm glad about the diagram, because I would never be able to construct a causal diagram myself.

DR. RUBIN: And just realize they're always wrong. No, they're not always wrong, they're probably sometimes right. You just don't know when they are, and when they're not.

DR. KOHBERGER: Don, two comments. I'm glad to hear you say that about causal diagrams. I mean, I included it in my talk because it's oftentimes referred to a way of dealing with causality. And the fact that it's wrong, I mean, I don't have a problem. Now I'll take it out of my slides, since nobody wants to see it.

DR. RUBIN: No one wants to see it, anyway.

DR. KOHBERGER: That's fine. I never quite understood how to set up these functional equation models, anyway, so it makes me feel good.

I think I know what you're going to say on this, but the question of dose immunogenicity. I mean, one way of dealing with this is looking at fitting a model with survival, and immunogenicity, and see if there's a significant effective dose above the immune response. And I think you're probably going to say since it's not randomized, that's not a

valid question-

DR. RUBIN: That's right. And, actually, I've written about three articles with little simple examples showing how that gets completely the wrong answer. And that's an example that Fisher got wrong, and it's in his paper that was published, a couple of papers that were published, and one is discussed by Stefan Lourdes, sort of trying to

-- who's one of these big fans of the graphical approach, but it just doesn't work.

DR. KOHBERGER: So I figured you were going to say that, so the question I have then is, I can't think of a way to use principle stratification in terms of that kind of immunogenicity data to look at dose. Just off-hand, do you think it can be done? And, if so, it means I just have to think a little bit harder.

DR. RUBIN: Yes, it can be done, but see, the trick is -- the problem with the approach when you just like regress out, regress survival on immunogenicity, and see that it doesn't depend upon dose. So dose is one variable, it can be like zero one, say, two level dose experiment. But immunogenicity is the observed value of immunogenicity, that's what you're doing the regression on, and for half the rabbits, let's say, it's immunogenicity under the active high level of treatment, and for the other half of the

rabbits, immunogenicity under another level of treatment. So you're regressing on just one -- replace one variable by half of one variable, and half of the other variable; and, therefore, you can get the wrong answer all the time. You may get lucky and get the right answer, for example, if immunogenicity is not affected by dose, then you get the right answer. But let's hope immunogenicity is affected by dose. Right? And then it's two different variables, and you have half the animals in one variable, and half the animals with the other variable, so you're doing the wrong thing. And it's very easy to create simple examples where it fails, even though Prentice, his criteria -- you're doing exactly what Prentice says, except when he adds the thing you have to have perfect prediction, that problem goes away.

DR. SELF: So to the perfect prediction, so we saw in data from the passive transfer experiments that I think shows pretty

clearly that antibody levels we're measuring isn't the whole story. There's a lot of other response; and, yet, we also saw data that our ability to measure the cellular response by ELI-spot is just hideous.

There are other ways to measure cellular response. How important is it to try and capture other aspects of immune response that aren't currently captured in these antibody measurements that we know is contributing, or likely to contribute to -

DR. RUBIN: I mean, my own feeling is that those things are very important. And, also, if you can find covariates, things that are measured prior to randomization, that can help a lot, because if you can find other things, even on rabbits. I don't know what they would be, but measurements on rabbits that -- I'm sorry. We don't really have to, page 6 of Louise's thing where she had -- there were two different doses, but they had the same immunogenicity, but one of them

died, and one of them lived, even though they had the same immunogenicity, but at different doses. Well, maybe that's because the vaccine did something else besides the measured immunogenicity, or maybe it was just random. And which it is makes a big difference. Right? But you can't see it from the data that's collected. It's not there.

Now maybe if you found out one of the rabbits was old, and one was young, and the old one died, and the young one lived, maybe that's the reason. Or maybe the one who lived had been exposed to something, or maybe the one who lived was male, and the one who died was female. But those kind of descriptors can help understand it. If you have one that lived, and one that died, and they had the same level of immunogenicity, but different levels of dose to get to that level of immunogenicity, how do you know it doesn't depend upon the dose? The dose is doing something else, and that's exactly the point

you're raising.

DR. FERRIERI: This happens all the time in animal studies, that you have these unpredictables, and so although you may not like it, you're not able to really control for it. The likelihood is, it is random, and it may be important, and it may be one of the other variables that we would love to understand from a basic immunological point of view. But when you have small numbers in those vaccine groups, you're going to have these outliers that go in opposite directions, and I would think that the people who have done the work would be able to substantiate that point from FDA, NIH, whatever.

DR. LYONS: Yes. Don, I'm just -- I hear and understand what you're saying, but I'm concerned with sort of reality, because within an out-bred population of rabbits, or primates, you're dealing with probably multiple, multiple covariates that we don't understand, so just the genetic

susceptibility to the toxin, for instance, and for Cell-Ts, and neutrophils function to clean up the bugs, and things like that, that for a given animal you can set up so many different assays, but it might be actually different for each individual animal. And rather than -- I mean, it could create just a morass of not even information, just data that you can't even use, so that's where I get concerned about going too far down that road right now, because we have limited tools to even understand some of the questions.

DR. RUBIN: But at some level, you're going to have to make some assumption about what these other variables do, either say they don't do anything, or nature is randomizing them all, and that's -- I'm saying you have to be explicit about that. And to the extent that you can get some information that may -- oh, now I understand. That's right. These are -- of course he died, because he's got this other problem. I didn't

hear any discussion about that.

DR. HEWETT: I think we're also focusing -- you're focusing on the host response, the immunogenicity to a particular molecule, and I hate to bring this up, Drisilla, but - we have a standing joke - that there are other virulence factors here. We acknowledge the fact that we're talking about toxin neutralization with using PA as the antigen. But there clearly are other virulence factors that may come into play. If you have enough antibody to neutralize half of the toxin, but not all of it, then the remaining toxin may not be enough to kill the animal or the person, and then the other virulence factors can come into play, and have synergistic effects that we're not dealing with at all. So I think we're not going to identify those now, necessarily. We just have to acknowledge that they're there.

DR. BURNS: Larry.

DR. WINBERRY: Yes. I just had one,

more of a pragmatic comment.

DR. BURNS: Would you say your name first.

DR. WINBERRY: Hi. I'm Larry.

DR. BURNS: Larry Winberry.

DR. WINBERRY: Larry Winberry. Yes, from BCG. We're talking about the variables that contribute to the immunogenicity, but one aspect is the challenge, the number of spores. You might have two animals with the same immune response, but they may get a different spore load even though you're controlling LD 50's, et cetera. But unless you've seen the challenge and understand that, that's a calculation. And then the other is the biological variability in terms of that spore germination. Two animals may get the same number of spores, one may get a germination burst, whereas one may not. So the overall challenge to those individual animals may differ. Whereas, I think, there is probably more control in terms of the dosing with the

vaccine.

DR. HEWLETT: But that probably is pretty much randomized though.

DR. FERRIERI: Thank you very much, Larry, because that's a point I raised earlier from the floor this morning. That despite your telling us what the challenge was, this was not going to be perfect each time.

DR. BURNS: So it sounds to me like it's going to be difficult to do all of the experiments to check all of the variable that need to be checked. And the question is are there some reasonable assumptions that can be made from other systems that we know about or are there - for dose, is there nothing that we can rely upon and it really has to be looked at?

DR. LYONS: I think there probably are reasonable assumptions, as long as you explicate them, that FDA, in my experience is willing to be reasonable and listen.

DR. BURNS: Ok. On that note, I do

want to get on with the next - we do need to move on a little bit. Now, TNA titers - we talked a little bit here about extrapolation from animals to humans, but I think this is more a nuts and bolts question. Does the TNA assay - is it species independent from what you've heard, so that you could use it to extrapolate or do you see any limitations there?

DR. LYONS: I think it's definitely species independent. Now, whether you can use it to extrapolate, I'll have to throw it down to Don. But it is, the advantage is that it is measuring a functional endpoint rather than something that requires a species-specific reporter system. So, it ignores that and just measures activity or lack thereof. And so I think that's a very nice endpoint.

DR. BURNS: And now we - Pat.

DR. FERRIERI: I was just going to say that I like it very much, but I do like having both expressions of the immunogenicity

and although the ELISA is disconnected at times from the functional assay results, I favor the TNA immensely. There is nothing like ever - nothing is ever going to be better for any vaccine assessment than having functional assay whether your endpoint is kill of bacterium or, in this case, killing of cells.

DR. BURNS: Any other comments on assays and read outs. Yes.

DR. VOLKMANN: Yes, I have a comment to that, because, what about systems where you get perfect protection in the absence of any antibodies, just by T cells. So, yes, you're right it's a functional assay. But it measures only one function of complex immune system.

DR. FERRIERI: Well, you're absolutely correct, but from a practical standpoint based on the choice of this as the primary virulence factor, this is as good as it's going to get in assessing the value of the PA vaccine, in my opinion. I love all of

the other parameters of assessing the immune system, but it's not going to be possible to examine all of those other variables.

DR. GOTSCHLICH: I think one has to be careful that one is not beguiled by the word "functional". The issue here is that yes, it is very attractive, and it is very, very limited. It applies to this particular antigen, only. But I would like to like to always see it accompanied by a quantitative immune response, because that can be related to other things, other antigens, other things, and I would not like to see this field stuck in the mire of bactericidal reactions, which in meningococcal fields remains mired. So let us be careful here.

DR. BURNS: Okay. Any other -

DR. FERRIERI: Well, we're not dispensing with a quantitative antibody here. They're linked together, and I would encourage that they're inextricably linked, and let's go forward together, at least from

my point of view.

DR. LYNN: I just want to make one comment regarding the TNA, and that is, I think it's a misnomer to think of it as not a quantitative assay. In fact, I think it's a beautifully quantitative assay. And, again, looking at the values we get among laboratories, we're getting very similar values, and we also can apply a standard and do the same kinds of normalization that you would do in an ELISA, so from that standpoint, I would just argue it is a quantitative assay.

DR. FERRIERI: And we're not going to get into antibody avidity today apropos of the ELISA results and the quantitative-specific antibody.

DR. BURNS: Okay. Let's move on. We touched on this, but perhaps we -- there might be a few more comments about this. Comment on the strengths and limitations of active immunization, and passive immunization data in defining the correlate of protection.

Any takers?

DR. LYONS: Okay. I'll take one shot. I think it would be very interesting, if you're looking for experiments to do.

(Laughter.)

DR. LYONS: You know, it would be very interesting to take a monoclonal antibody that doesn't neutralize, and one that does in a passive transfer experiment in the rabbit and see what happens. I mean, I think that would be -- that would help close the door on, at least at some level, what we're talking about, because then we're not worried about interfering antibodies in a mixed population, these kind of things. You're talking about strictly a monoclonal, it either neutralizes, or it doesn't. And I think the available tools are out there, I think. And that would be worth adding to the data, I think, and that would help. And that's what I think the beauty of passive transfer is, it really sets up the worst scenario, and yet, if you can

protect in that scenario, it's hard to argue with, I think. I mean, it's very hard to argue with.

DR. BURNS: So do you think the data that we saw today was a promising start, as far as -

DR. LYONS: Oh, yes. I was just suggesting an extension on that.

DR. BURNS: Yes?

MR. SUTER: Yes. I think I would like to reiterate what I said this morning or afternoon, I can't remember. I mean, an antibody has very functional capabilities, neutralization is one. We can measure that by TNA. It can be taken up by a number of cells. And I think I would really like to see an FAB, whether it's neutralizing it or not, if it's monoclonal or not, I don't really care. But I would like to see the functional neutralization part disassociated from the exceed part, which I think is very important, which you do not measure with TNA.

DR. BURNS: Would anybody like to -- who does those experiments like to comment on the difficulties and feasibility, or any technical problems? No? I do know that it was difficult enough to do the experiment that was done, to make FAB fragments adds a level of complexity. I mean, it is a wonderful idea. The question, I think, deals more with feasibility on this.

DR. FERRIERI: One more point on the strength of the passive immunization, is that from my simple way of looking at it, it permits a better definition of the quantity of antibody that may be protective in cross-species studies, et cetera. And it just generates a more powerful argument at the level of protection, I think.

DR. SELF: Yes. I'm probably out of my depth here, but I would disagree. I mean, it seems to me that those studies tell you something qualitative about mechanism, and sort of confirm something like that, but to

the extent that there aren't all of the other aspects of adaptive immune response that are going on, I think makes it very difficult to use those to quantitate any sort of threshold that's applicable in a vaccination setting. And we saw the data where the thresholds in the one experiment that we saw were much, much higher than what we saw in the vaccination experiment. Anyway, maybe I'll -

DR. LYONS: I see your point, but I think it really enhances the argument that it is a good correlate of protection, because it's the only thing that's there. Now whether or not -- it's a minimum thing you need to be protected.

DR. SELF: Right. I was with you until the quantitation -

DR. LYONS: Okay.

DR. SELF: -- part in the threshold. Then that's where I -

DR. BURNS: Which maybe I could just ask a little bit, to extend this

conversation a little bit, because it seems that the passive immunization studies may over-estimate the amount of antibody that might be needed, but with the active immunization studies, under-estimate, or are they a more accurate reflection? I mean, we're making this leap, again, from animals to humans. Do you think active is sufficient, and having the passive study in there was very reassuring for exactly the reasons that Rick talked about?

DR. FERRIERI: I think the passive is absolutely necessary here, just as we argued several years ago. I understand his point. It's a more refined way of looking at all of this stuff, but I do not think that it's adequate to only do the active studies, active immunization. Emil, do you have -

DR. GOTSCHLICH: Well, first of all, the purity of the data that you get from the passive protection studies has already been alluded to. Then there are also, of

course, other things that can come up down the line in terms of the therapeutic potential of such antibodies. And for that, we also need passive protection studies, although they're not necessarily germane to the topic that we're dealing with today.

I still feel very much at sea with the data that has been presented, and the fact that it has not been presented in the formulations that are being so elegantly discussed by the two gentlemen here on my left, who think that I disagree with them, but, in fact, I do not. It's the fact that I have not seen the data formulated in the way that they would wish it to be formulated, and it seems to me some of the data here could be formulated in this way. And would, therefore, be more intelligible, and be more digestible.

DR. RUBIN: Actually, one comment that I'm going to make about that is, it seems to me that the passive immunization is really much closer to the thing I was talking about

in titrating to immunization, because you actually get to titrate that level. You put in that number of antibodies, rather than letting the vaccine do what it's going to do. So I think that the passive immunization is really -- if you see the same relationship short-term, because we know that they'll go away in time, between survival and immunization there, as you do after vaccinations, but we see a little bit less. Right? So that probably means the vaccine is doing something else beyond the antibodies.

DR. FERRIERI: Or?

DR. RUBIN: We don't know what. That's right. But there's got to be something else there, we think.

DR. FERRIERI: Perhaps.

DR. RUBIN: Can't measure yet.

DR. HEWETT: Or the additional response -- the anamnestic response, those animals clearly got passive immunization, but then developed their own immune response, and

they hadn't seen the antigen before. So that makes it all the more important that if they have seen it before, they get additional -- likely get additional benefit from that.

DR. FERRIERI: Well, there is the concept of the capsule of the organism is a very important virulence factor under various conditions, and I'm speaking a little out of my field here, because I don't do research on Anthrax, but there could be up-regulation of capsule a la Larry's comments on spores, the nature of germination, all of those things that do not predict the numbers, but then capsule begins to play a critical role as an antiphagocytic virulence factor. And, so, yes, there are many, many other things about the organism that we're not even discussing, that's not appropriate here, but that influences outcome, in my opinion.

DR. LYNN: I'm Freyja Lynn. I wanted to just clarify one point from a couple of things I just heard. So if we can show, or

can we show through passive immunization that the antibody level is dose-independent? So can we passively transfer different levels of antibody, show that we get different levels of survival, and, therefore, establish the antibody that way as being dose-independent? Is that feasible?

DR. RUBIN: Well, I don't think that's what we saw. Right? We saw survival - correct me if I'm wrong, because I've never seen this before today - but I thought we saw that survival was better with active immunization, than with passive immunization.

DR. LYNN: Right. I'm thinking about it in a two-stage process. In other words, if we can show that antibody does protect independent of the dose of vaccine, so that it is an important aspect of the immune response, and then go back to active immunization to get the absolute level, is that a feasible approach? Because I don't think we're ever going to get away from the

dose-dependent issue.

DR. RUBIN: It's certainly a piece of it, but it's not complete without some other assumptions. But they get better and better. I mean, as you drive the noise level, the understandable noise level smaller and smaller, the question of the validity of the model becomes less and less important, so you understand more and more.

DR. NUZUM: So let me try to take another shot at what I think Freyja is -- and I'm trying to understand this, too, so I think you're saying we need to somehow randomize the immune response independent of dose. So based on the dose-dependent GUP studies we're doing, we're learning where the linear part of the curve is, and so why couldn't we do -- pick the midpoint in that curve, one dose, instead of using five groups of ten rabbits, one group of 50, and we know we have variability within the group, so one dose of 50 rabbits, and, as you say, let nature do the randomization.

DR. RUBIN: Well, but is nature doing the randomization? At one dose, is the one who has the highest immune response only randomly different from the one who has the lowest immune response? Probably not, so -- in fact, Steve just said absolutely not. Now I don't agree with that, so what you have to do is why not actively randomize different immune responses to titrate to? And then if you get the same dose, same survival immunogenicity level curves as you do when you're randomizing dose, then you've learned a tremendous amount.

DR. GOTSCHLICH: Could I ask you a question? I'm really not very clear on what you're talking about.

(Laughter.)

DR. GOTSCHLICH: Could you explain to me exactly how you believe the dose changes the antibody that is produced by the animal?

DR. RUBIN: How the dose changes -

DR. GOTSCHLICH: Changes the

quality or the nature of the antibody that is being produced by the animal. I don't quite understand.

DR. RUBIN: Well, isn't it true that as the dose goes up in a randomized experiment, the immunogenicity level goes up? The immunogenicity level goes up, the antibody level.

DR. GOTSCHLICH: The antibody level.

DR. RUBIN: Yes, the antibody level goes up, the higher the -

DR. GOTSCHLICH: Well, that certainly was shown by all the experiments -

DR. RUBIN: That's what I thought. I thought I saw that. Yes. You're saying no?

DR. FERRIERI: For some antigens, that is not the case.

DR. RUBIN: Okay.

DR. FERRIERI: Where it is not a principle in doing just response curves that

that is always fulfilled. This is somewhat different than what we're seeing with Anthrax today, but it is not a universal principle, that as you increase the dose, that the animal is going to make more antibody.

DR. RUBIN: No, that I understand.

DR. GOTSCHLICH: But the issue that I'm not clear on is, why do you believe that there is a problem in the amount of antibody that is produced by a lower dose versus a higher dose of antigen, if, in fact, you, at the end of the day, wind up with two animals that got the same amount -- that have the same amount of antibody in their system?

DR. RUBIN: Okay. If they have the same amount of antibody, and they got -- but they got different doses.

DR. GOTSCHLICH: Yes. Why is that antibody different?

DR. RUBIN: Okay. So do you believe that those two -- well, I will tell you in this Tryptizol, which is anti-epileptic

drug study, it could be very different, because they had different body weights, they were different sexes, they're all different kind of descriptors that if you now controlled what you blocked for those things, you saw different relationships.

DR. LYONS: But I think drugs are very different than -

DR. RUBIN: It may be.

DR. LYONS: Okay.

DR. RUBIN: This is not an area that I know a lot about.

DR. LYONS: Okay. Okay.

DR. RUBIN: I'm saying that's the issue that you're going to have -- but that's the kind of argument that you have to make. And if you were to say that if two animals have the same immunogenicity level, even though they got that from different doses, that they're equally protected, we don't think that's true. Right? We saw data that suggests it's not true.

DR. SELF: So another example how they might not be randomized, same dose, different antibody levels, are they randomized? Well, one with the higher level might generate a much better T-cell response, and that leads to higher antibody levels. The other one has a crummy cellular response, leading to lower antibody levels. The quality is very -

DR. RUBIN: But I think -

DR. SELF: So they're -

DR. RUBIN: No, I understand what you're saying, but to me, that -- you're getting -- we're mixing sort of formulation issues that really drive a very robust immune response, and alter the immunogenicity of the antigens, versus doing these control studies, where we are making the animal produce less by giving a low antigen immunization, versus a high antigen immunization. And we're giving it under a controlled circumstance, so I think, one -- I see what you're saying, but I

think they're addressing different issues, myself. I think, I agree. And, to me, that's -- I would call that randomization, whether or not an animal takes -- has a robust T-cell response, genetics are doing the randomization, because we're looking at out-bred population here. I mean, we can't control that. There's no way. That is a random event in nature, the genetics.

DR. BURNS: Emil, you've done a lot of correlate work. Emil, in other systems where there's a correlate of protection, it's assumed that the dose doesn't matter. Correct? It would just be the level of antibody that's important. That's what's being assumed in other systems.

DR. GOTSCHLICH: Let me stick to my own formulation. At the -- we have an antigen. We know what the antigen is. We have an animal. It's a black box. Out of that animal comes an antibody. We actually believe that the antibody is more or less a

chemical entity, and that was the issue that I was trying to get to, that at the end of the day, we come up with a chemical entity, and it wasn't clear to me why we were concerned with how much antigen was necessary to put into the black box to get the same amount of antibody out. That was the only -- because I couldn't quite understand the importance of also worrying about the dose. That's all.

DR. RUBIN: Well, if the scientists are going to tell me, if the doctors are going to tell me, the scientists that it makes absolutely no difference how you got that level of immunogenicity, whether you got no -- you got the level of immunogenicity because you were born with it, or because you got the maximum dose vaccine, it makes absolutely no difference to survival, I'll believe you, but you have to argue that with people more knowledgeable than I am about it. Okay? I'm willing to accept it, but then I'll ask how do you measure immunogenicity exactly? Now maybe

it's this true value of immunogenicity, but we've heard lots of ways of measuring immunogenicity, some of which are highly correlated with each other, and some which are not. And if you're saying all those things are identical, that I don't believe.

DR. GOTSCHLICH: Well, I think the task before us is to try -- as I understand it, the only thing that we have is the antibody response, which can be measured in a number of ways, but whether that can be in some way correlated with the amount of protection that it can afford. And that, I think, is the main task before us. There is no evidence that was produced for any underlying cellular immune phenomena beyond the fact that, obviously, cellular immune phenomena drive the antibody in this box, but no other correlation could be gotten.

DR. RUBIN: But how is it measured, is it peak, is it average over some period of time. If it's the average over some period of

time, for how long? At what point in time is it measured?

DR. GOTSCHLICH: That gets us back to exactly the thing that I mentioned to you before, is that we have not seen an integrated set of data in a digestible form.

DR. RUBIN: Right.

DR. NASS: I have some comments about dose. Meryl Nass. There are two papers by Lincoln, et al, from 1967 and `68, in which pure PA, as pure as they made it back then, was injected into the bloodstream of monkeys. And in one study they showed that all brain electrical activity ceased for several minutes in some of those monkeys, and in the other study, they showed that certain blood chemistries had profound changes with the injecting of PA. Those studies have never been repeated in the open literature since then, but since paracelpsis, we are aware that every substance has a dose that makes it a good thing, or a poison. And in the case of

this vaccine, and it's very important to take the safety data in order to determine efficacy, because there may be certain doses that are safe, and other doses that are not safe, and the safe doses need to be efficacious.

Now the GAO showed us in 2000 or 2001 that due to manufacturing changes at the time of the Gulf War, the concentration of PA in now BioThrax, previously AVA, increased approximately 100 times, and there are a number of people who suspect that that increase in PA concentration, or other materials in the vaccine, which also increased in concentration, may have something to do with the fact that it appears the vaccine is causing a much higher rate of adverse reactions now than it did before.

The data that it increased in strength came from Fort Dietrich, as did the Lincoln papers come from Fort Dietrich, so there probably are people here who can speak

to this, but this -

DR. BURNS: I think we really need to keep on subject today, which is immunogenicity and efficacy. And I appreciate your comments, but it really is -- of course, safety is a huge part of these vaccines, but it's just not the topic today. And we have a lot of other things that we're covering, so I'd like to keep to just immunogenicity and efficacy for the time being. Okay?

DR. NASS: That's okay, but you're -- I mean we are talking about specific doses of PA here, and we've pulled safety out of it, and it's critical, before we say 25 micrograms of PA is the right dose, to know that that dose is okay.

DR. BURNS: I appreciate your comments, but I think we do want to move on to talk about statistical approaches for inferring protective efficacy in humans from animal data. Do we have any comments, especially from the far end of the table, on

this, beyond what's been said?

DR. SELF: Well, I guess I'll just pile in. A lot of the last discussion has been around how to -- building more sort of detailed and accurate causal models within a species, but at the end of the day, we're going to run into the limit of our ability to measure things, that it's not going to be perfect, and the balancing kind of data that we can get to that, given sort of the imperfect within-species model, is that across species. And so, the more data that we can get across species, however imperfect, I think that will, ultimately, be the basis for an empirical, more empirical predictive model that will be the basis for this extrapolation to humans. And so, we just need to kind of balance how much we invest in rabbits, and how much we invest in getting across a few -- building a few bridges across species.

DR. BURNS: Any other thoughts? I mean, I don't want to belabor things, but

-- yes?

MR. SUTER: You said earlier, and I think this is a very important point, we need to have standardized assays so we can compare actually the assays, but I think what we also need is a comparison between two species, which we can challenge. And I think death and survival is a very good read-out. And I think this is the first step to see whether we can even then extrapolate from this data to human. And I think before we can extrapolate without having these data from rabbit directly to human, we need to have two species where we can compare these different assays, and then we will see whether antibodies really make a difference or not.

DR. FERRIERI: Regarding establishing a protective antibody titer for this part of the question, I would just say that there are several vaccines out on the market that have only been out there within the past 10 years, where the waiver is, we

don't know what the protective antibody titer is, the concentration that is protective, pneumococcal polysaccharides being an example. Good is good. I love high. I don't like low, or I'm not sure, is 1.5 micrograms of some given antibody good enough? It may not be if you have a huge challenge, but usually, with the rarest exception, there's nothing bad about high antibody titers.

DR. BURNS: Steve, did you have a comment? Anything else on statistical issues that -- anybody? Okay. Then the final question, what additional data, if any, are needed to strengthen the extrapolation of GUP protection data in animals to efficacy in humans? And we've heard some good ideas come out already.

Are there any other need to know types of data, or nice to know, even, just hear your ideas on that.

DR. LYONS: I was actually talking to Erik about this, because this is the first

time I've ever seen sort of so many statistical people in the same room with biologists, and it's a little scary to me.

(Laughter.)

DR. LYONS: Particularly, talking to Don there. I wouldn't want to meet him in an alley, but I think -- and this is true just for emerging infections, and challenge for bio threats, in general. I think it would be a reasonable idea to get -- to have some sort of meeting of types to discuss these issues, because it's the same thing for all of the organisms. And whether or not we can even talk about causality, I don't know, but you have your dominant pathways, and then you have your tangential ones that, I think, is what Don is talking about, the ones that really add value to the findings, and are likely to be very important; although, the level of which we don't know. But it would be good that the statistical people understand our limitations, and we understand their concerns, and come to

sort of a meeting of the minds, so that we have a plan in mind for any pathogen that these -- have a focus plan, but then we're bringing -- we're, at least, keeping the data that might help the final analysis, and it just doesn't go out with the bath water, so to speak, because a lot of data is lost, just because it's negative data, and so it never gets out there. But there's a lot of important data out there.

DR. BURNS: Anybody else? Any final comments from anybody on the panel?

DR. FERRIERI: Well, I do have a question for you, Drisilla, about the antigenic composition of PA that I probably should have done some research on a long time ago; and that is, trying to better understand variability, and epitopes that may then stimulate protective antibodies, and how much variation there is, if any, that's been detected to-date from one strain to another. We're always working with Ames, or Sterns, or

whatever, and I just wonder if any of you worry about that.

In the case of Emil's favorite organism, with some of the outer membrane proteins, an alteration in one epitope can shift the level of immunity within a given population over a period of years, examples, outbreaks in the Netherlands where existing antibodies would not protect because of that alteration in an epitope of that protein. And, so, I have a gap in my knowledge and understanding PA, its variability, et cetera.

DR. BURNS: My understanding is PA does not vary all that much. It's relatively conserved, but Erik, or Rick, do you have other -- or people in the audience? We've got a lot of people in the audience. Anybody have any good recollections? Judy?

DR. HEWETT: I can't really make any specific comments, but my understanding is that there are a variety of monoclones to PA that have different activities, so that would

probably be the easiest way to look at that.

DR. BURNS: Right. We do know the sequences of PA from different strains. And from what I've seen, it hasn't been that different, but please correct me if I'm wrong, somebody.

DR. FERRIERI: My interest in it has to do with projecting, or cause being the role of bio terrorism, the risk of bio terrorisms, the ability to manipulate the organism, et cetera. Are we really infatuated with something that could fool us and change?

DR. BURNS: Conrad?

DR. QUINN: Conrad Quinn, CDC. The existing literature indicates that the protective antigen gene is very clonal, is very little, if any clonal variance between different strains, and that translates literally into the protein sequence. There are very few differences between the different geographic isolates that have been looked at to-date.

DR. BURNS: And I think, Pat, from me being in the toxin field for many years, to have an active toxin requires a lot of different components of the protein working together. And, therefore, making drastic changes that might change the antigenicity a lot, I would think would be very difficult to accomplish. Erik, do you have any other thoughts on that?

DR. HEWLETT: Back to the pertussis analogy again, the micro heterogeneity that exists doesn't really affect -- that does exist, but doesn't really affect the function, so I think it's precluded, those bigger changes are precluded by virtue of the need to have the function.

DR. BURNS: Bob?

DR. KOHBERGER: I'm going to take a real gamble, and re-express what Don said. What we've seen is, if we have an animal that we vaccinate them, and they achieve an immune response of 400, they're protected; yet, if

that 400 is given passively to an animal, it's not protective. That's what the passive and the active show, so we know that antibody alone doesn't tell the whole story.

So the next question is, suppose I give at a dose and get 400, and now I give half a dose, and yet that animal gets 400, this different animal gets 400, are they going to be equally protected, or is there something about the dose itself that impacts protection? If we can make an assumption, and maybe we just have to make this very explicit, that whatever the immune system is doing actively to generate a TNA level of 400, it doesn't matter what the dose is. There's a black box going on underneath there. There's cellular things, there's all sorts of things that immunologists could tell me what's happening, but as long as actively they get the 400, the two 400s are the same. Does that sort of explain what you're saying, Don?

DR. RUBIN: What I said? Yes, sir.

DR. KOHBERGER: Rick, does that help?

DR. LYONS: Yes. The only caveat I would say is that I think you have to be careful about during the active immunization, because as Erik was saying, the robust amnestic response comes up so quickly when the host sees it, that that's going on in the background, and we would -- it still could be, and I, personally, believe that it is dominantly antibody that's doing this. And the reason active works better is that the baseline 400 we're measuring really looks at, and I think Conrad suggested this today, and I agree, it sort of represents a population of B-cells operating in the background ready to turn on at a moment's notice, and it comes roaring on within hours, certainly. And that's our -- then you get titers that skyrocket, as opposed to passive, that it's just being absorbed, and being sucked out of the system.

DR. KOHBERGER: Just one -- so you're saying with active immunization, it doesn't matter how we get to 400, because all these other things that are going on.

DR. LYONS: Right.

DR. KOHBERGER: That's fine. That's an explicit assumption we need to make this a almost a -

DR. FERRIERI: But there is a caveat to this, and that is upon repeat immunizations with the same protein vaccine, you may generate different populations of antibodies, so all the antibodies are not equal, and we're getting a polyclonal response. And from one person to another, the polyclonality may vary, from my perspective, which may be false. But it's a very exciting, dynamic heterogeneous population of antibodies that may be generated from the first immunization, as well as maybe more so upon re-immunization, which is certainly the plan. One dose isn't going to cut it.

DR. BURNS: Conrad?

DR. QUINN: We tried to address this in the CDC macaque study, where we get different levels of antigen. The first hypothesis was if we get different levels of antigen, we could get different magnitudes of response, we will be able to module that immune response. I think David Madigan showed that we achieved that. He also showed that within each dilution of the vaccine where the antigen load is normalized, as far as is possible, the variance by animal was significant, so we have this randomization, if you like, or perhaps we're approaching it.

He also showed that irrespective of what level of antigen they got, if they got above 250 at week 34, they had a 90 percent chance of protection, or expectation of protection, so we are taking those steps in those directions. David also pointed out we had 136 animals in the study, but the statistical power was low, so how many animals

are we going to need to do these types of randomizations? It's large.

DR. RUBIN: May I make a couple of comments? One is that I -- we have to be very careful here that we are comparing on the one hand the titers that have been obtained with the standard vaccine, the AVA vaccine. And on the other hand, we're also -- we're trying to compare that to titers that have been obtained with the recombinant protein. And we really do have to keep the two apart, because we do not know to what extent other small amounts of antigen exist in the AVA vaccine, and do have some effect, so we will have to analyze this data somewhat separately.

The other thing that I would like to make a comment on is immunological memory, that's not what's come up. And I will use the example of the polysaccharides, and that is everybody had great hopes for immunological memory that was supposed to be engendered by the conjugate vaccines to be an

important part of the protection. And a the end of the day, children are protected, as long as they antibody aboard. They may easily be prime -- they may be prime, they may easily respond to the next injection, but they're not protected unless they have antibody aboard. And that's been shown now for the Group C meningococcus, that's been shown for the amophylis, and it will probably be shown for all the others. Why is this important at this point in time? Because I think that an Anthrax infection is a lot more like what you would get in a pneumococcal or meningococcal, or amophylis-type infection, very acute and rapid disease where there isn't time for the child to play around with its immune system in order to combat this infection, neither is there with somebody who inhales a good dose of Anthrax spores, so I think we really will have to focus here on the antibody that is aboard at the time of the infection.

DR. BURNS: Okay.

DR. QUINN: One more comment. I didn't want to imply that the response at week 30 was a threshold. I think it's a surrogate for the state of readiness of the immune system. It's not a threshold, it's a surrogate for the state of readiness or correlate, whichever the right word is.

DR. BURNS: Anything else? On that note, I think --

DR. HEWLETT: Just let me say, Emil, I agree with you, but I think that we -- that the diseases are different, but I think we probably need to have those data to see whether having antibodies -- the passive antibodies that are on their way down, and the active antibodies that are on their way up, how much difference that will make, because I think there are circumstances in which it could make a difference in terms of a rapid response. I think we just need to know that information. I don't think we can infer that.

DR. GOTSCHLICH: I agree with you.

DR. HEWLETT: Okay.

DR. BURNS: Okay. I would like to thank the panel, and the audience, and we'll see you tomorrow morning.

(Applause.)

(Whereupon, the proceedings went off the record at 5:30:25 p.m.)

UNITED STATES OF AMERICA
DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOOD AND DRUG ADMINISTRATION

CENTER FOR BIOLOGICS EVALUATION AND RESEARCH

ANTHRAX VACCINES: BRIDGING CORRELATES OF PROTECTION IN ANIMALS TO IMMUNOGENICITY IN HUMANS

Friday, November 9, 2007

The workshop began at 8:30 a.m. in the grand ballroom of the Hilton Washington North, 620 Perry Parkway, Gaithersburg, MD.

T-A-B-L-E O-F C-O-N-T-E-N-T-S

Session 5: Animal Models for Post-Exposure Prophylaxis

NIAID Perspective on Rabbit Post-Exposure Prophylaxis Model Development, Judy Hewitt, Office of Biodefense Research Affairs, Division of Microbiology and Infectious Diseases, NIAID.....5

Development of a Rabbit Post-Exposure Prophylaxis Model for Testing Combinations of Antibiotics and Vaccines against Inhalational Anthrax, Jason Mott, Battelle Biomedical Research Center 14

Session 6: Panel Discussion on Post-exposure Prophylaxis, Moderator: Ed Nuzum, NIAID

Post-Exposure Prophylaxis: Issues to Consider when Developing a Correlate of Protection for the PEP Model, Bruce Meade, Meade Biologics, LLC 80

Session 7: Meeting Summary, Moderator: Ed Nuzum, NIAID.....135

Summary comments from panelists.....140

Adjourn

P-R-O-C-E-E-D-I-N-G-S

8:34 a.m.

DR. MEADE: Good morning. It's about 8:30. Can we go ahead and get started? Good morning. My name is Bruce Meade. Today I am serving in the function as a consultant to DMID and will help moderate the morning session on post-exposure prophylaxis. Before we get started Freyja asked me to make three announcements and hopefully I will get them correct and, if not, she'll make sure I get them right.

First is I understand there are some extra copies of the meeting materials, the handouts, the slides that you all received. I think there are a few extra copies that will be out in the lobby. Feel free to take one extra copy for a colleague if you would like one.

I understand, again as you heard, that the transcripts will be available, my understanding is, sometime within the next month or so. The instructions on how to obtain those will be available on the FDA website. I guess you can start checking that

in a few weeks for that information.

I understand those of you who need taxis should make arrangements directly with the hotel staff. Individuals meeting are really feared up to do that so work directly with the hotel staff.

Again, I think just to briefly introduce this morning's sessions we are going to change direction a bit and deal with another important aspect of the anthrax and that is on the post-exposure models and approach and approaches to dealing with vaccination after event if a possible event should occur.

There will be three talks. Judy Hewitt, Jason Mott from Battelle, and then I will finish up with mostly asking you some questions at that point. The first talk will be Judy Hewitt from DMID who will talk about the development of the rabbit pet model.

DR. HEWITT: Thanks, Bruce. Good morning, everyone. What I want to do today is tell you a little bit of the story about how

we developed this model. I'm not going to get into any data. I'm leaving that to Jason.

First I would like to start off with some of the assumptions that we made going into this program. In the beginning we anticipated that we would need two models for licensure, the rabbit and the nonhuman primate. That was really based on the 2002 workshop and discussions.

Back in 2003 when we started this effort the post-exposure prophylaxis, or PEP indication, was in our minds going to be the second indication that would come after licensure of the vaccine for the GUP indication.

Also, at that time we anticipated developing a nonhuman primate model ahead of a rabbit model and that was based on some preliminary studies that have been done by Friedlander in the early '90s predominately aimed at licensure of various antibiotics for anthrax, but there was also a component of those studies that had some vaccine added to

the antibiotic regimen.

Because the antibiotic Friedlander used at that time was cipro we assumed that would be the antibiotic that we would use for our model development efforts.

Of course, in science there are always changes to your plan so PEP became the highest priority around 2003 when the bioshield legislation was first proposed and it was clear that there was an intention to stockpile the vaccine and it would be used basically for emergency use or under emergency use authorization.

The push was to get the PEP model developed first rather than having that PEP be a follow-on indication. So we had a proposal of how we were going to go about developing the PEP model and with the importance and urgency of developing this model we deleted one study that had originally been in our plan and that was to determine the start time of the antibiotics and the vaccine regimen.

Instead of performing a study to

actually decide what was the most appropriate time to start antibiotics and vaccination, we just discussed it, picked a time, and went with it.

Of course, the antibiotic changed. I should back up and say we started our efforts in parallel with both rabbits and nonhuman primates. You are going to hear about the rabbit model today because that has essentially been completed. We are still working on the nonhuman primate model.

In the rabbit model we changed the antibiotic to levo. That was based on the rabbit's tolerance of the antibiotic. Then, of course, as is typical in science, we got some unanticipated results.

Basically did a study where we got some very promising results but it fell short of statistical significance. That required a little bit of refinement and repeating a study so that added a little bit of time to our plan. At any rate, things moved on.

The other point I would like to

emphasize here is that the development of this model the goal was really to be able to demonstrate added value of vaccination. It was not strictly to use the human regimen in terms of the antibiotic and the vaccine.

In fact, I think until the human regimen or human dose of the vaccine is really determined or humanized dose in rabbits if you will, these studies will have to be repeated. At any rate, we feel that the model is very well developed and ready for final drug product to be tested at the appropriate doses.

This just gives you sort of an overall summary of the sort of level of effort that went into this project. I have some proposed and actual statistics here. The duration of the end-life studies we anticipated to do these in very rapid order but it actually took us much longer.

My second line here, the duration with the reports, is really important because it is one thing to do the studies and gather the data but then you also have to get the

reports into FDA so that they can look at the data in a very rigorous manner. It has been a real considerable effort to get all these reports completed and into FDA.

Originally the original plan was seven studies. We dropped the one right away, as I mentioned, but because of the unanticipated results we actually ended up doing nine studies in the end. The number of rabbits you can see doubled basically. I think this is kind of an important point.

It gives us a huge amount of confidence in this model in that quite a few studies have been done. They have behaved very well from study to study. There is actually a large number of animals in these studies.

Of course, originally we were only going to test two vaccines and we ended up testing three. As I said, they behave very similarly and Jason will show you that data.

I want to introduce this concept of TRLs or technology readiness levels. It's

something that the Government has used in the past to assess the readiness of particular countermeasures or, you know, in the military things like tanks. Are they ready to be bought and purchased by the Government.

In an effort to support planning for bioshield stockpile procurements people are always asking the question, "Do we have the animal models that are going to be needed to get these products licensed?"

We developed a series of levels here and this is just sort of a very high-level summary of what the animal model TRLs represent with one being your very most basic studies and 9 being GLP studies that are going to be sufficient for pivotal studies for licensure and all the various steps along the way always building on your model, getting closer and closer to your end target.

So a couple of conclusions then about our model development program here. I gave you a sense of sort of investment and effort spent on this program was.

We have basically progressed this rabbit PEP model from something like a TRL 2 or 3 all the way up to an 8 or 9 depending on whether -- I mean, in my opinion the model actually is sufficient to support pivotal studies and it is a matter of the product being at that level as well, final drug product consistency lots.

All of the studies that we did except the very first one were GLP and the result is a very well-characterized model. All of these reports have been filed. We have a master file with FDA so all of the animal model development work that we've done is filed with FDA through our master file.

The studies that have single vaccine efficacy testing in them have been filed through the various sponsor INDs. Of course then we grant the IND holders letters of cross reference to reference our master file for the information that they require to support their pivotal studies in the end.

I have to end with this slide which

is a picture of the nine final study reports. This is all very neat and tidy. These are all completed. They are actually all now in the hands of FDA as of this week.

The fun is going to begin in terms of looking at this model and the data and assessing whether products are ready to be -- whether studies are ready to be considered pivotal to support the PEP indication.

With that, I am finished and I'll take any questions.

DR. NASS: I am unaware of strong evidence that treatment is needed in the time frame that a vaccine is going to become effective post-exposure. I know that was our assumption at the time the letters were sent but of the 30,000 people who may have been exposed who did not take long-term antibiotics and did not take vaccination, none of them became ill. So what is the data that shows that we actually need post-exposure vaccination?

DR. HEWITT: Well, I'm not here to

address what the need is. I'm here basically to tell you that we have a model that can support that. In terms of stockpiling a vaccine the intention is that it would be used only in a post-event scenario. The need is an unknown at this point.

DR. NASS: Well, in order to get something licensed you have to have an indication so I would think it would be a critical question for you to answer.

DR. HEWITT: Right. We have a model that can support -- one model that can help support licensure for this indication.

DR. NASS: You don't have an indication.

DR. HEWITT: I appreciate your comment.

DR. MEADE: Our next speaker is Jason Mott from the Battelle Biomedical Research Center in Columbus. He is going to talk to us about all of the very difficult and important work that they have done with this post-exposure model.

DR. MOTT: First of all, I would like to thank Judy and Ed and Freyja for the opportunity to be here and everybody else for coming to see this talk. As Bruce mentioned, I will be discussing a lot of data that was derived over the last couple of years on the development of a post-exposure model in the rabbit.

As Judy mentioned, the post-exposure is one scenario that we were looking at that has been placed in a high priority just in case we do have another accidental or deliberate release of Bacillus anthracis spores to the public.

The whole premise of our studies were based on a combination therapy to look at the ability of the antibiotics to kill the spores that germinate after an exposure as well as a vaccine to develop a protective immune response to take care of the latent spores that were left in the lungs.

The current antibiotic regimen is 60 days. As we have heard, compliance is

often an issue so certainly the added benefit of the vaccine will help those who do not take a full 60 days of antibiotics. They are reported to be less than 50 percent compliance following those last attacks in 2001.

The rabbit model, as Judy mentioned before, was worked through over a period of about four years with countless hours and discussions with an animal working group to help us show added benefit of the combined therapies with success being defined as increased survival in the rabbits following a potential lethal challenge.

Reproducibility was obviously critical for us and we were looking to establish a model that could be applicable to new vaccines or therapeutics. A lot of the initial goals in the original proposal were developed and individual studies were designed to help us evaluate each one of those goals.

As Judy mentioned, one of the goals which was determining optimum antibiotic start time was not performed. However, we still

need to look at optimum antibiotic dose to allow us to get partial survival or death in rabbits post-challenge and post-treatment.

The studies I'm going to go through will help to address each one of these goals that you see here. The knowledge of these first set of goals was then used to do a combined study with antibiotics and vaccine with an aerosol challenge in the rabbits.

I have a lot of data to go through in a short period of time so if I don't cover something and you have questions, just find me or e-mail me at some point to help clarify.

The first study was a toxicology tolerance testing and pharmacokinetics study in the rabbits using two doses of antibiotic. It's a nonchallenged study. We use levofloxacin in these studies. As Judy mentioned, cipro was dropped out due to tolerance issues.

The study design is shown in the chart where we had two separate antibiotic regimens that were modeled to give us plasma

levels that would be similar to what we would expect in humans.

There was a split regimen of 40 and 10 and then another split regimen of 20 and five mg/kg given about 12 hours across 14 days. We did take blood for plasma antibody levels. As you can see it bears time points as well as looking at CBC and clinical chemistry for the toxicity issues.

What we did find was that with levofloxacin there were no toxicity issues. We did not see any clinical signs during treatment. No lethargy, hematuria, respiratory distress or weight loss. We also did not see any significant changes in the CBC or the clinical chemistry data. Then on gross necropsy and histopathology there were also no signs attributable to the administration of the antibiotic.

The pharmacokinetic data did show that the Cmax or the P concentrations were found at about a half an hour to three hours post antibiotic treatment. At the high dose

regimen we did see that the Cmax was at least 10-fold higher than the reported MIC50 of .12 micrograms per mil.

We also determined an AUC/MIC ratio at these two doses and found that at the high dose there was a ratio of greater than 125 which indicates a very high therapeutic potential. We also chose that dose for subsequent studies because we would expect that would be a dose that may interfere with vaccination if that is going to occur.

With that said this chart shows a little bit of the actual data from the two groups in comparison to some data that is in the literature for humans. We can see at the levofloxacin split regimen of 40 and 10 we have a Cmax of about 3.21 and AUC of 15.8 and a half-life of 2.6 hours.

With the dose that was half of that we see a similar half-life. However, we see a Cmax and AUC of about half that like we would expect. In comparison to the humans we don't see an exact plasma level that we would see in

humans but, as Judy mentioned before, that really wasn't the complete goal here. It was just to try to make sure that we maintain antibiotic levels throughout the studies at levels that were above the MIC50 or MIC90.

Then we moved on to a second study which also was a nonchallenged study. It was a concurrent antibiotic therapy and vaccination study that also looked at various vaccine regimens and doses.

The objective here obviously was to assess the tolerance and toxicity of the dual treatment with antibiotics and immunization as well as to evaluate immune responses in the various regimens in order to determine a rapid response that we could expect to be utilized in subsequent challenge studies.

We also looked at the peak and trough flows of antibiotics. Once again, we did look at hematology and clinical chemistry data on these animals. This is a breakdown of what the study looked like and you can see that we had nine groups of animals with up to

24 animals in a group with various vaccination days all the way to zero and 21 and various vaccine doses of 10 or 50 micrograms of RPA.

We also had in group 9 one group that had the high dose of antibiotics added for 14 days with zero, seven, and 10 microgram vaccine dose. We also took blood at various time points, as you would expect, for ELISA and TNA analysis and plasma antibiotic analysis as well. What we saw here was that dual administration of the antibiotics and the vaccination did not have an effect on each other.

The levels of the ELISAs and the TNAs that were seen in comparable groups as well as the antibiotic levels did not change from what we had expected from the previous study. We also found that two doses was better than one as we expected and that boosting at day zero and seven produced a higher response than zero and five.

The 50 microgram of RPA group given on days zero and seven did produce a higher

level of ELISAs and TNAs. However, it wasn't significant when you compared that to a similar group receiving 10 micrograms. We did not have any tolerance for toxicity issues in the rabbits, again, and we did have peak levofloxacin levels similar to what we had in the first study.

This is just one of the figures that was put together for the study that further is evidence of the immune response. The 50 micrograms of RPA given on days zero and seven on the top in red and you can see, as I mentioned before, there was a slightly higher level of immune response for the ELISAs and the TNAs. This is specifically for the GMC and for the ELISA response.

The other two lines that we can see here are zero and seven regimen with 10 micrograms rather than 50. In one of these groups, this purple group here, was given antibiotics at the same time so we can see that there was no ill effects of the dual treatment on these animals for the immune

response.

We then moved on to a post-challenge antibiotic duration study. The main objective here was to determine a length of antibiotic to be given to these animals post-treatment that would allow us to have no deaths during the treatment but a less-than-optimum survival or less than 100 percent survival once the antibiotics were removed.

We also did the normal assessment of immune assays hematology clinical chemistry, bacteremias as well. The study design here was to take New Zealand white rabbits challenged with an aerosol LD50 dose of approximately 200 LD50s on day zero and then administer antibiotics for five, seven, or nine days.

What you will see is a little different here that we used levofloxacin at 50 mg/kg once a day rather than a split dose. Judy alluded to this earlier that we did have some issues part way through so I'm concentrating on the study that we used to

determine antibiotic dose for the more pivotal studies.

The animals were bled for plasma antibiotic level as well as ELISAs and TNAs. These animals were watched for 28 days post-antibiotic treatment. We also did gross necropsies found dead to help us determine the cause of death was anthrax and make sure of that.

Results showed that observations on the animals as far as clinical signs were absent until the antibiotics were stopped. Most commonly we saw lethargy and anorexia, normal clinical signs. We did not see anything on clinical chemistry data. We also saw minimal changes in the CBCs once the antibiotics were stopped.

The typical changes were increases in neutrophils as well as decreases in lymphocytes post-exposure. Pharmacokinetics were as expected. We did have levels similar to the previous studies with the levofloxacin. As far as survival the important part once

the animals were on antibiotics and post-challenge we did not see any deaths until approximately several days post-antibiotic treatment.

When we look at this table we see the results from the five, seven, and nine-day treatment groups and we see that with the five and seven we had over 90 percent mortality. With nine days we had about 70 percent. What you will notice here, too, is that when we look at the time to death, the time to death in those groups was obviously delayed.

There were no deaths during the antibiotic treatment. You can see that compared to the control nontreated groups we got a time to death that we expected would be good to allow a protective immune response from a combined treatment study. You can see there was a large range in death in some of the groups with one of the animals dying out close to 30 days.

We used the data from these first three studies to help us design the final more

pivotal study where we combined the antibiotics and the vaccine together post-exposure. The main goal here was to determine if the combined therapy allowed the animals to survive long enough post-antibiotic treatment for that vaccine to kick in to have an effective immune response.

We also did the normal battery of tests as far as plasma antibiotic levels, ELISAs, TNAs, clin chem, CBCs, temperatures. The study design here was to take New Zealand white rabbits again and challenge on day zero with 200 LD50 of spores.

The antibiotics were given, once again, once a day at about six to 12 hours post-challenge at 50 mg/kg for seven days. Then they were vaccinated. Here it says with "with what?" I'm not sure that was really supposed to be in there. What it really means is that we used one of three vaccines in this study. As we go on we'll see that we repeated the study three times.

We did the complete gross

necropsies as well. This is the basic study design of these three studies with the vaccine. What we see here are five groups of animals treated. The first three groups were treated with a combination of antibiotics and vaccine, the 50 mg/kg antibiotic for the seven days.

Then for the vaccine we used what I am describing here as a high, medium, or low dose for the studies. Group four shows an antibiotic only. Group five are the untreated controls. The vaccinations were also started at six to 12 hours post-challenge at the same time as the first antibiotic dose.

What we have done, as I mentioned, we repeated the study no less than three times using three different vaccines with the 50 mg/kg levo dose. We have seen very good reproducibility. All nontreated control animals have died through these studies. That's 28 out of 28. You can see that the time to death is very short of 2.4 to 4.1 days.

Mortality was absent in all the treated groups until cessation of antibiotic just as the previous study that I had described with a range of 11 to 28 days post-challenge. The vaccine combination groups did show that they had a significant amount of protection in combination with the antibiotics or with the antibiotic only.

Approximately 54 percent of animals in the antibiotic only group died post-challenge where only 9 percent in the vaccine groups. We can see that there was a slight dose dependency with four, seven, or 15 percent in the high, medium, and low doses respectively.

This is a figure that shows the mortality rates of the three studies in one. As we expected, you can see there was a rapid death rate in those animals that were not treated with a 2.4 to 4.1 average, as I mentioned before.

There is a slight grouping here of the three antibiotic only groups from the

three studies at about the 50 percent death rate. Then there is significant changes here of the survival in the combined groups.

Bacteremia was also looked at in these studies and we did see that a very high percentage of animals that died on the study post-challenge or post-treatment were bacteremic. All over it was 93 percent. What I do want to note is that while the animals were on antibiotics very few, one out of 162 animals, showed a positive bacteremia.

In addition, only six out of 173 survivors showed a positive bacteremia at any time point. The bacteremias were not only taken at the time of death but also at various time points throughout the studies.

Anorexia was the most common clinical sign and that was more prevalent prior to death and post-treatment if they were treated. Temperature changes were also noted. We took temperatures twice daily from little transponder chips that were implanted into the rabbits. However, if you saw a change in

temperature of 2 to 2.5 degrees it was very predictive of mortality in those animals.

Generally animals that were on antibiotics and vaccines, either of those groups did not show any changes in temperature while they were on it but prior to death if they did die it did become prevalent.

A mean peak in trough levels of levofloxacin were routinely greater than the MIC90 at all time points that they were collected except for the first initial trough collection. As far as the immunogenicity the anti-PA and ELISA and TNA was detectable by 10 days post-vaccination in those treated groups.

As we expected, antibiotic only groups if they did survive did not show a development of a protective immune response. It was significantly different than those animals that received a combination of antibiotics and vaccine for the levels.

What we have done is take all this data to help us define the model which we hope will be used to design studies testing other

vaccines or antibiotics as needed. Certainly for submission to other regulatory agencies such as the FDA we would have to do this similar study in this model that we feel is very well characterized at this point to use their final product formulation.

I do have a couple more slides showing some of the information. This is similar to the earlier survival or mortality graph that I showed with Kaplan-Meier survival curves. It shows step by step when animals died for each of the groups. It looks very similar to the earlier.

We do have a geometric mean curve for the ELISA data for groups one through three for the three studies. There does appear to be a slight dose escalation there. The one complicating factor for these studies versus other normal vaccine studies is that we did have the challenge so we have to throw that in there.

It makes it a lot harder to interpret the levels of the immune response as

well. This looks very similar to the geometric mean curve for the ED50 for the TNA data for the same animals.

There were some trends that were seen during these studies that we have taken for development of other rabbit anthrax models. Bacteremia is certain, as I mentioned, a very good predictor of death in these animals with over 90 percent of animals that died showing a positive bacteremia at some point.

Temperature is also something that we feel very confident in as a predictor of death with animals that show a 2 to 2.5 degree increase or decrease in the body temperature just prior to death. It's very predictive. It usually happens within two to three days prior to the animal dying.

Then for hematology we do see increases in neutrophils, decreases in lymphocytes and then the CRP reacts to protein which is an acute phase reactant. Usually it's going to be something that increases its

levels during inflammation and infection and that is another thing we've gone to.

Baseline levels of the CRP in these animals are normally less than detectable limits. However, as these animals get sick to the point of dying the levels constantly increase at the time of death.

This is just a slide to show that the proposed path really wasn't what happened over these last couple years. Originally we did have six studies that would be done and two of those at the end would be very similar.

As Judy mentioned, we had a small hitch in the road here that made us go back, reevaluate what we were doing from the previous study, run another study until we could get to the point of having that combination model that we felt confident in.

I would like to thank NIAID, Judy and Ed and Freyja as well as others, the companies that provided vaccine, BEI Resources for providing us all of the reagents for our ELISAs and TNAs, a whole multitude of

technicians and study directors and stats, QA at Battelle.

One person that isn't on here, too, is Jim Eastep that helped us design these initial models. He's not at our facility anymore but he should be recognized as well.

If you have any questions, I'll take them now.

PARTICIPANT: In looking at the Kaplan-Meier curves there seems to be a gap in the protection. I think the trick here is kind of a handoff between the protection provided by the antibiotic and that by the vaccine. From days 10 to 15 there is a little window there where you see deaths in the most efficacious groups.

It would suggest that window really hasn't been completely closed. I am wondering what the thinking was to stop at seven days the antibiotic given what you know about peak responses to zero to seven vaccination a nd so on.

DR. MOTT: The normal time of death

in the control animals being 2.4 to four days basically. We went back and looked at the data of animals that were treated with antibiotics we did see a slight delay in the time to death in those groups.

Especially it seemed to increase with the increasing time of antibiotics so five versus seven versus nine. If you look back at that previous data there was an animal in that five group that was at 30 days that really expanded that group out to look like the seven days.

We picked that time of seven because of the significant death that we saw in that group as well as the time to death which we suspected based on previous data that showed that the accelerated vaccine regimen would provide us some immune response by 10 days. We were seeing times of death of 12 days, 13 days as an average and we suspected that would be a good way forward.

DR. SELF: In retrospect, would you think that a ten-day course of antibiotics

would close that?

DR. HEWITT: I'll answer that. If we extend the antibiotics to 10 days, then the antibiotic only control group would have been a higher percentage of survival, more like 80, 90 percent. We would have had to have very large groups in order to get a statistical difference between the antibiotic only and the antibiotic plus vaccine.

I agree completely that there is a gap there and basically we chose the seven days because we were still able to demonstrate the statistical difference between those groups.

DR. SELF: I understand the arguments about study design, still the regimen to me is somewhere along the course length. Do you know what I mean?

DR. HEWITT: Right, but I think what you will run into then is that the antibiotics themselves will be so protected that you may not be able to demonstrate added value of vaccination.

DR. MOTT: In some of the data --

DR. HEWITT: I agree this is artificial.

DR. MOTT: Some of the data that you don't see in a previous study indicates that as well from that last figure where I showed we kind of took one way versus the other. An earlier study did have a longer administration of antibiotics that had a lot higher survival as well.

DR. SELF: So then a ten-day course prior to vaccination, wouldn't that be credible?

DR. HEWITT: Potentially, yes, in rabbits.

DR. CHAWLA: Just to add to what he is saying, usually most of the vaccines taken 15 to 20 days to really give the effect. I have been working with vaccines for the last 20 years and it seemed that the response of bacterial vaccines. It takes 15 to 20 days for the immune response to come so it will be a good idea to see the studies for at least 15

days.

My second question is related to antibody which is the antibody in human. Is it levofloxacin or ciprofloxacin?

DR. MOTT: They are both approved but for at least for our studies we have to go to the levo based on tolerance issues for the rabbits.

DR. CHAWLA: But, as you know, yesterday we discussed many times that we are focusing on human vaccine, not a experiment in animals. The focus should be use the antibody in human beings.

DR. MOTT: The antibiotic or the antibody?

DR. CHAWLA: Yes, antibiotic which can be used --

DR. MOTT: We do use the other antibiotic with nonhuman primates but for the rabbits based on the tolerance issues we couldn't use it for an extended period of time.

DR. CHAWLA: Is it going to be

levofloxacin as the antibiotic of choice in humans in post-exposure cases in combination with --

DR. MOTT: I don't think I can make that call. That is not up to me based on these models.

DR. CHAWLA: Okay. My third question is that there is a lot of potential for exposure in terms of GA exposure so any plans to have a post-exposure model for GA in addition to aerosol challenge? Because it is easier, I would say, for anybody to expose human beings to GA exposure instead of aerosol.

DR. MOTT: I don't know of any right now.

DR. CHAWLA: Let's say the spores mix with water or food.

DR. MOTT: We have been focusing on the aerosol.

DR. CHAWLA: Because that could be kind of an easier route of exposure than creating aerosols, etc. A study should be

designed for that kind of exposure also.

DR. FRIEDLANDER: Just a comment in reference to the design of the experiment and the question that was asked before. As I understand it, the purpose of this is to try to demonstrate that vaccination can enhance the efficacy of post-exposure so it is very dependent upon the animal model.

It is dependent upon the dose that you use and the retention in the particular animal, some of which is unknown and has to be determined empirically as was done here.

That is to say, how long do you need to treat such that when you stop you will get sufficient mortality. If you don't, you can't obviously demonstrate the added benefit of vaccination. That may different depending on the model.

DR. MEADE: In the back.

DR. BIGGER: John Bigger, Battelle. Jason, I just wanted to ask was there any indication that survivorship after cessation of antibiotics was immune response dependent?

Was N bright enough to do a logistic regression on this like the day 10 ELISA data comparing survivorship to immune response?

DR. MOTT: Say that again?

DR. BIGGER: You know, Mark and I both presented survivorship based on immune response in our talks and in our case we had very defined control and a range of immune responses and our end was pretty high in our survivor versus death rate. You may not have that luxury here but maybe at day 10 if you were to look at the immune response in the animals at day 10 it might be a predictor of survivorship or not.

DR. MOTT: I don't know that we did that analysis.

DR. MEADE: We are going to talk about that a little bit in the next talk or, at least, explore that a little bit in the next talk as well as in the panel discussion. It's a question that is being thought about, yes.

In the middle, please.

DR. GLYNN: Kate Glynn, CDC. I found this model very intriguing and I think very useful as a first step. It clearly does show that the addition of vaccine on top of a microbial agent is beneficial.

I guess what I would like to her is how you think the next step -- how this model could actually stand up to the next step which is talking about how post-exposure prophylaxis regimen using vaccine and antimicrobial agents would over the estimated potential incubation period for spores actually provide the protection that we're talking about for PEP and whether this model could be used to extend.

I don't know if that may be getting at what some of the other people are asking as well. It shows very nicely this particular issue but whether this model could actually apply to the broader question.

DR. MOTT: I certainly think this model is shown to be very well characterized in reproducible. Do I think it could be used

towards the next step, something that we can interpret the data here to be extended to humans? I do. I realize that not everything is optimal here to look like a human infection.,

DR. GLYNN: No, I'm sorry. Just to be clear, not to just look more like humans but the situation we are trying to address in humans if that model could be recreated -- I mean, that situation could be addressed using this model, the animal model.

DR. MOTT: As a deliberate release, for example.

DR. GLYNN: And the longer-term protection so this addresses short-term.

DR. MOTT: I don't know whether or not this would provide long-term protection. I would guess that it would. I know one issue with the rabbits that isn't going to be the same in humans is that the period to treat to keep this model going versus getting into an area where it's therapeutic is a very little window so I don't think we can extend it out

much further.

Obviously with the primates they seem to have an extended period of illness by a couple days. I think once we get that going maybe we can extrapolate from that model a lot better towards what we would see in humans.

DR. MEADE: Pat.

DR. FERRIERI: Pat Ferrieri, University of Minnesota. This is very nice preliminary data. I gather we are not going to hear anything on the nonhuman primate along the same lines of combination of antibiotic with vaccine.

My interest among others is the pharmacokinetics of the antibiotics in the different animals. I'm sorry that your rabbits couldn't tolerate the ciprofloxacin because I think it is viewed as the preferable agent in the human situation over levofloxacin.

Levofloxacin human serum levels were much higher than the rabbit area under the curve so I guess it doesn't completely

address the issue of whether the antibiotic would suppress the immune response.

Have you give any thought to that or maybe someone else in the room has data on the ciprofloxacin in the nonhuman primate that would address the issue of suppression of immune response. Nonhuman primate levels may be more comparable theoretically to ours with cipro or levo.

DR. MOTT: Well, having run a study with primates and the vaccines and antibiotic in combination, although the data isn't being presented here I can tell you that it is similar to the rabbits that we don't see any changes in the levels of the immune response. The antibiotic levels in the primates do tend to be higher with the cipro but we haven't seen any interactions that show it does inhibit the response.

DR. FERRIERI: One other quick question. I assume that the ciprofloxacin or levofloxacin does not interfere with the TNA, the ability to conduct the assays, the TNA?

DR. MOTT: I think I would give that off to someone in the middle area.

Kristen.

DR. CLEMENT: We looked at that during our validations studies, the effect of various types of rabbit matrices and one of them was rabbit serum from rabbits that had not just been -- it wasn't a spiked situation where we look rabbits spiked with levo.

We actually gave rabbits the highest dose that has been used on studies of levo and drew their serum to use in our TNA validation studies. The end result is there is no effect of the levo metabolites.

DR. FRIEDLANDER: Just to address this question that was raised about whether in the primate ciprofloxacin will interfere with the immune response. We published a paper last year on the nonhuman primate in which vaccination occurred post-exposure along with ciprofloxacin.

Those survivors developed an immune response and were resistant to rechallenge

although we did not compare the levels to a group that just received vaccine alone. The antibiotic did not inhibit the ability of the vaccine to induce a protective immune response.

DR. MEADE: Okay. In this last talk before the discussion I wanted to -- I was asked to at least talk through some of the issues related to use of the PEP model in the area of correlates. Since the purpose of the meeting is sort of translational, how do we think about applying the PEP model to humans?

The question has already come up. There's a number of issues that have been raised and I think I just wanted to go through a few other points that we thought about or at least questions we've asked. I won't have any questions. Again, my job was really to raise questions and hopefully lead into some of the panel discussion issues that we brought up.

I certainly won't raise all of the questions and issues but, again, I did want to highlight a few of the complexities that we

have thought about. A lot of these -- many of these come from hours of discussions with Drisilla Burns and many other people in these groups. This is reflecting the thoughts.

Again, just to focus back on the issue of the meeting in terms of the animal rule and, again, what we are focusing on primarily is the fourth item. The fourth issue is the information that will allow the decision to be made in humans and how do we begin to think about linking those models.

Again, I think this was raised or has already been brought up but, again, I just wanted to refocus and make sure that everyone is clear what the issue is that we are really talking about, the management of individuals after exposure or possible exposure to Bacillus anthracis spores.

Again, either some post event which I define as maybe or presumed event, possible exposure, or exposure. I don't really see that those are different. In most cases you're not going to know if there was an

exposure or not so I perceive those as really the same thing but to distinguish it from treatment.

This is really not the intent to be dealing with a treatment after onset of symptoms. That's the models that were being designed to deal with those steps. I think there is no doubt that the primary approach will be antibiotics and the whole issue is to evaluate vaccines as a supplement to that. I don't think there has been any intent or belief that antibiotics would not be part of the approach.

Again, I think just to reiterate what I think was one of Judy's comments is the animal studies for post-exposure will build on the pre-exposure or general use studies we talked about yesterday, the active and passive immunization studies. They are really intended to build on all of the information that has been gathered and the really nice data that we heard yesterday.

Again, you saw a number of these

curves yesterday where the general approach to coming up with some estimates of what concentrations of antibody would be protective.

Again, you saw a number of these curves where you have plot the antibody of the titer as a function for individuals whether or not they died or survived and then logistic regression or other models you can make a plot of the probability of survival as a function of TNA and then, again, with confidence intervals you can make the kind of modeling and come up with estimates that we heard about a number of times yesterday.

I will speak only for myself but when we were thinking about this at the beginning I sort of naively thought you can do this with PEP, too. Why not sort of think about the same thing. What's the problem? Go ahead and take the TNA and make these plots.

The more I thought about it the more I realized that is really hard and has more problems and issues. I'm going to spend

the next few slides talking about what I think are the reasons that's problematic and needs someone far smarter than I maybe to help put this all together. Again, I think that will be the focus on some of the discussions.

Again, just to remind you, I'm going to build on the data that Jason has just presented but, again, just to simply it, we are talking about studies where there is a group that receives antibiotics alone, one that receives antibiotics and vaccine. You obviously have a control group that receives neither just to be sure your challenge is behaving properly.

After some minimal amount of time needed for the vaccine to elicit an immune response the treatment is stopped in the groups one and two and then you are going to monitor survival over time which, again, is the data that Jason presented. Then you are going to compare the group that received antibiotics and vaccine to that who received antibiotics alone.

I think I had group three in the handouts incorrectly thanks to Bob Kohberger. He noted my error. Again, that is sort of simple. What I've done here is maybe inappropriately from a purist point of view but, again, I was trying to keep things simple.

As I put together the three experiments that Jason just presented to keep it simple because the specifics matter less than the general format of the data. Again, just to highlight the basic data, those animals that receive no antibiotic at all and no vaccine die very quickly.

Again, those animals in blue -- I'm color blind so I try to do colors that I can see but I don't necessarily know what to call them. I'll keep pointing and hopefully the shapes will help. The blue line is the groups that receive the antibiotic with no vaccine.

Again, it just shows, as Jason said, that there are no deaths during the time when antibiotics are around and beginning a

few days after the antibiotics are terminated the deaths occur over a fairly long window of time which again differentiates that from a lot of the studies you saw in the pre-exposure is that the deaths occur in a much shorter window of time.

Again, they are occurring from 11 to 28 or something like that. Again, the three groups. I pulled the three vaccines, the low dose, the mid dose, and the high dose to show the time at which deaths are occurring. You clearly saw added value from that perspective.

It clearly shows that the vaccine is adding value to the system. Again, my simple-minded thoughts at the beginning were the ideals situation as you would estimate the antibody titers and animals protected by the vaccine.

You compare those eventually to those levels achieved by humans receiving the vaccine administered on the schedule intended for the post-exposure prophylaxis. Then you

would use approaches we talked about yesterday and sort of make some estimates about the efficacy of vaccine for post-exposure indication.

However, again, as I have said, as I thought about this estimation of a protected level from the pet model, at least in my mind, it's a lot more complex than with the pre-exposure models we heard yesterday. Again, I just want to talk through some of the ones that I've thought about. Again, I think we have already heard a few others that I think we can talk about further.

One is the fact that the pet model is a very dynamic situation. There are many things happening at the same time and I'll illustrate that. Again, you have an infection going on at the same time you have immunization.

Again, as has already been alluded to, and Jason mentioned this, too, I mean, the fact that you're immunizing in the presence of organisms I think there are questions that the

antibody response could be influenced by infection.

Then there are certainly -- again, because of the nature and reality of this I think there is greater uncertainty in the statistical modeling, at least based on the sample size as we have so far and I'll illustrate at least what I mean by that as an nonstatitician.

Again, just summarizing the data just to give the shapes of the curves more than any details, this is the antibody response from the three PEP studies, pooled the TNA as a function of days.

There is a dose effective in going from low, middle, high but the more important things is the kinetics, that the immunizations are given zero or seven. Beginning a few days after that seven-day dose there is a very rapid rise of antibodies during the period of the study.

Again, just to highlight what at least my perceptions or uncertainties

regarding the PEP model, the first issue is the fact that it's a very dynamic situation.

As you saw, some of the immunized animals are going to die following removal of the antibiotic and they die, at least the antibiotic only control group, died at a very wide interval of time from days 11 to 28. As I showed in the last curve, the antibody is very rapidly increasing during that same interval.

To relate, again, to the question John Bigger is asking is what TNA value should be used in some of the modeling to define some protective level or try to make some estimations of what TNA value is relevant.

Again, I think, at least my simple-minded view, for those animals that die you presumably could use the TNA value in the last sample prior to death. That is one option I think you could consider but I don't have -- I have far more trouble with what would one do with the animals that survive.

To illustrate to put the two graphs

on the same slide that I just showed, for animals that survived in the challenge what point should you be using for doing this modeling? Again, I think one could make several proposals or, at least, we have kicked around several proposals.

One is why don't you choose day seven, the day the antibiotics are discontinued or maybe day 10 or the day of the first step, something like that, that phase. Clearly if you look at the antibody response curve, that is very early on. It's very low responses.

At least, from my perspective, you could be well beyond her estimating the value that is relevant during the window of time that matters. I guess you could also propose that you look at the peak or maybe the last day in the control group, something down here.

I think, again, that is very likely to be overestimating the protective values. Again, because it's basically going from all to nothing during that same window, you can

sort of choose anything you want. I think that is where I'm based.

Choosing a logic for choosing a time or, at least, modeling a time, is something I think we hopefully will have some discussion about. I'm a little uncertain as to how one would make that and defend it in terms of a modeling point of view.

I think, again, the second issue that I think in the first animal model working group discussions on this, I think Jim Eastep raised this question that you are immunizing in the presence of an infection.

I mean, clearly there are antigens from the organism that could be there or the infection process could be influencing immune response in a positive or negative way. I think it's different. I think there is some evidence from studies.

The studies were never designed and empowered to look at the immune response in this way but I think there is some evidence that the antibody concentrations may be

higher, at least in some of the rabbits, that receive the vaccine and challenge and others receive vaccine only.

I think to answer this question, if it's necessary to be answered, you look at the TNA titers in the two groups, one that had the antibiotic and vaccine and challenge versus those that got antibiotic and vaccine in the absence of challenge. There are some animals in that comparison.

Again, I think they certainly were not powered to look at differences but there is certainly some evidence that some individuals did show an increase that certainly could be attributed to the presence of the challenge.

The issue of three deals with statistical issues. I need to give a disclaimer that I'm not a statistician. I should have brought the picture but a statistical colleague of mine at my retirement event last winter gave me the book, "Statistics for Dummies."

I think there was some comment about my statistical experience in there. Again, hopefully I can at least illustrate the questions and hopefully will turn it over to statisticians who can turn this into more details.

The issue is that the survivors, at least using my terminology, represent true positives and false positives in that they are 40 to 50 percent of the unimmunized animals survive.

We had, again, looking at the animals immunized with antibiotics about half the animals are surviving, again, in the absence of a vaccine. For any given animal vaccinated that survives, can you differentiate whether or not it survived because of vaccine or because of -- or would have survived anyway.

I think what you're doing is population siding. You are clearly showing added value so that clearly shows added value vaccine. In terms of modeling I think you

have the concern that half the animals will survive anyway.

Than, again, the other point, at least with respect and comparison to the pre-exposure studies we talked about yesterday, relatively few immunized animals in these PEP studies died. Again, we are talking this group here and I think I did the counting and there were 162 total animals in these groups of which 14 survived.

We are talking about relatively small number of animals -- excuse me, that died. I'm sorry. Fourteen animals that died in the immunized group. I think generally we are doing modeling using the responses in immunized animals so I think, at least at this point, we have relatively few animals to be looking at.

Again, from a modeling point of view I think, again, we saw these yesterday. In the pre-exposure models, at least in the rabbit, all the low responders all died and all of the high responders lived so it's

relatively easy to -- you don't need very many animals at the asymptotes when it's an all or nothing event and you can really focus your studies on the range of interest where you are going from, the death to survival.

With the PEP studies we are really dealing with a lower asymptote in this 40 to 50 range. Again, my picture of it is graphed in here and if you are doing a similar curve comparing it to the rabbit GUP model, here are two curves which have the same E50 slope and asymptote. For the rabbit PEP model you are going to be fitting data from this kind of curve.

Again, you may not need that many animals here where there is the full protection but down here you are going to leave a lot more animals down here in this range here to really give any precision of your estimate of the values to come up with the shape of this curve.

Again, the fact that it is a much shallower curve, again, I think you generally

need more animals and more data when you have a flatter curve and you have basically less of a range differentiating between 50 percent and 100 percent. You are going to need more animals and really have quite a bit more data in here to really come up with a model with any precision.

Again, I think, just to summarize at least my understanding of the challenges to statistical modeling is that because 40 to 50 percent of the antibiotic treated animals survived without the vaccine, there is greater uncertainty in us estimating the lower asymptote in E50 and to achieve comparable precision in estimation of a protective level PEP studies would need to include many more animals than GUP studies. I noted on Judy's slide they have already done 700 animals, I think, so far. We are talking a very large resource commitment.

Again, I think, in the studies done to date I think there's relatively few animals that died, immunized animals that died so

future PEP studies, again, if one were trying to model, would need to include a wider range of doses including more sub-optimal doses and adequate modeling requires a sufficient number of surviving and nonsurviving animals from the immunized groups in order to do those models.

To summarize, the complexities in estimating a protective level from the PEP studies the fact that the PEP model is a dynamic situation which antibody concentration is rapidly increasing. Secondly, the antibody response could be influenced by the infection and then the fact basically of the nature of the model suggest you need more animals to do the modeling.

Again, I think just to go back one step, again, I think even to do this modeling -- one point I want to make is even to do this modeling one would have to make a decision about what, in fact, time point TNA you are going to model. I think that is in addition to the fact that you need more animals you would have to make a decision and come up with

rationale for how one would do that.

At least from my view, defining a protective level for TNA based on the PEP studies it appears to be in my mind challenging resource intensive. I think given that I just really wanted to raise some questions and some thoughts that I think will lead into the panel discussion and, I think, what alternatives should we be thinking about if we are unable to obtain good estimates in protective levels.

I think one of the things that we need to keep in mind is that as the studies proceed, data is going to become available for different vaccines, different species, different models, the general use of pre-exposure, the passive protection, post-exposure.

I think one question to ask in this context is really how -- I guess my approach to PEP, or at least one approach to PEP, would be how should the data from the various models be integrated to make judgments regarding the

value of immunization in the PEP application.

I think the question merits discussion. Again, there are some very bright people in the room that may have some thought about this and have some other interesting idea but can a protective level be defined for the PEP model.

I think importantly it should at least be asked is it necessary to divide a protective level for the PEP model. Is that necessary. Again, I think where we are now, what other data should be gathered in PEP studies. And, at least from one perspective, is it reasonable and appropriate to assume that a TNA determined to be protective through the pre-exposure GUP study would be protected in the PEP scenarios.

Or is there any reason to believe that if it's protective in one it wouldn't be protective in the other. Hopefully that will be discussed. I think, just to conclude, I really wanted to come back to a reminder of

how the studies were designed is the PEP were really designed to achieve the primary goal of demonstrating added value of vaccine.

Can you, in fact, demonstrate and certainly in the rabbit model as designed has clearly shown that with a lot of work involved as Jason and Judy have shown it was very difficult to get that. I have done a lot of animal modeling work in my life and I think this was more complex and more difficult than anything just because there are so many variables and such a fine tuning required.

It was a very difficult but, again, remarkably reproducible at this point and clearly shows the added value of the vaccine. I think at least one approach is to emphasize the data from the GUP studies when defining TNAs that can be used to predict protection as opposed to trying to do it in the model that may be very complex.

Anyway, I would be happy to respond to any questions at this point.

DR. CHAWLA: Anil Chawla. The

causing mortality in rabbits in terms of LD50 and in terms of how quickly it can kill rabbits, will it have the same kind of mortality in LD50 monkeys also or nonhuman primates?

DR. MEADE: I'm sorry?

DR. CHAWLA: The strain of challenge, the challenge strain that causes mortality in rabbits, let's say one species of animal, will it have the same kind of mortality or LD50 nonhuman primates?

DR. MEADE: I think I can safely say I don't know. Again, I am here from a data analysis perspective. I think there are other people in the room that have a lot more information about the strains and the models but, again, I think this was really focusing on what has been developed so far which is the rabbit model.

DR. CHAWLA: Can somebody from the audience answer this, the LD50 in rabbits and monkeys the same, the mortality the same from a particular strain?

DR. MEADE: I think the question is is the LD50 for this particular strain at least similar in rabbits and nonhuman primates.

DR. PITT: Similar.

DR. MEADE: Similar is the answer from Louise Pitt.

DR. CHAWLA: There is no need to adapt this strain to a particular species, the strain which is killing rabbits very quickly like within 48 hours in 200 LD50 need to be adapted in monkeys before it is used as a challenge strain?

DR. PITT: No.

DR. MEADE: Larry.

DR. WINBERRY: Larry Winberry with Biologics Consulting Group. Enjoyed the talk, Bruce. Good to see you.

DR. MEADE: But --

DR. WINBERRY: No, no. Actually, we're on the same page in terms of being statistically impaired. In looking at the modeling it would seem that from the vaccine

doses that were used in the study it would probably add value if we looked at a more fractional dose to get a lower protection. It seems weighted towards the higher end of your survival curve.

As you indicated, there is a weakness in the lower asymptote whereas if we were to perhaps exchange one of the higher doses which seem to have less discrimination capability for a third dose that perhaps was lower, we could fill in some of that data as well.

Then the other point would be it's problematic, as you indicated to identify where you take an antibody or TNA level as a predictor. I'm wondering when I looked at the antibody response curves, although there is a difference in peak height, you also see a difference in obviously the rate of change for the different doses.

I'm wondering if instead of an absolute cutoff one looked at, say a velocity or a rate of change in antibody titer as a

predictor of survival since that seems to be dose dependent as well.

DR. MEADE: I think two comments. One is in terms of the kinetics. In order to really fine tune the kinetics you would probably have to take more samples. You also have relatively few data points.

When you get a slope the question is is it the time which immuno response begins and the slope and the velocity is the same. There are multiple issues that require a lot more fine tuning. At least my perception, and I won't speak for anyone, is that the status of the PEP model in the rabbits is the groundwork has been laid for some of the more definitive experiments.

A lot of work went into this to show that it is reproducible, predictable, definable. Then I think the definitive studies are obviously going to be discussed at length between what are the final and pivotal studies and I know I won't be part of them.

DR. WINBERRY: I was very impressed

with the work that was done by Jason and his colleagues. Excellent studies.

DR. MACALUSO: Tony Macaluso from BARDA. I'm curious if these and other studies result in a PEP indication, I'm wondering how narrow or broad that indication will be. Will it apply only to cipro or to drugs in the same antibiotic class or to any antibiotic? I think this is a real concern as we think about possibilities of events with multiple drug resistant Bacillus anthracis.

DR. MEADE: I will choose not to respond because I know I don't have that information but I think that is the kind of question that should hopefully continue to be raised during the discussion. I think that are some of the issues that need to be further discussed.

DR. SUTER: Very simple question for you. When you challenge these rabbits and you put them on antibiotics where do you actually see the Bacillus? Is it still in the lung or is it transported somewhere else?

DR. MEADE: Is there someone from Battelle that would choose to respond to that? I think the question is where are the organisms following the infection. Can you speak into the microphone?

DR. MOTT: The bacteremias taken during these studies are from the blood itself. We also will sometimes do tissues if needed --

AUDIENCE MEMBER: Can't hear you.

DR. MOTT: The bacteremias that we take during these studies are from blood samples taken from ear veins, for example. Also tissues at the time of necropsy if needed to show bacteremia but we don't specifically look in the lungs post-challenge through lavages or any type of direct sampling that way.

DR. SUTER: I think it's interesting to compare rabbits with monkeys, for instance, because the anatomy of the lung is totally different because you have balls which you do not have in monkeys. I think it

would be important to see how the processing of these Bacilli are occurring between the different animals because you might see differences in the immune responses between rabbits and monkeys.

DR. FUNNELL: Can I just --

DR. MOTT: Sure, it would be interesting. We just haven't had a chance or opportunity to do that based on the design of these studies and the main goals at this time.

DR. FUNNELL: Can I just follow on from that discussion? You mentioned in your talk that a fluctuation in temperature and an increase in CRP was something that you found in the rabbit studies. Can anybody comment about those observations in NHPs?

DR. MOTT: We do look at those parameters in HP protocols as well and we do see changes in blood parameters. The temperatures may not be as good in the primates but we do see changes in the CRP as well.

Typically it's going to be baseline

values for hematology as well as CRP before challenge but post-challenge or at times when these animals show illness you do see changes in those parameters that we've used to help develop the models as well.

DR. KELLER: James Keller, FDA. I don't know anything about anthrax so perhaps I shouldn't be asking the question but it seemed like in a post-exposure situation a lot of the complexities for active immunization would go away if you used a immunoglobulin combination therapy with antibiotics. Was that addressed earlier?

I didn't catch the reason why we are going with active immunization and a post-exposure instead of trying more directly. I think a lot of complexities would go away with animal variation, variability to immune response if you gave a single injection of an immunoglobulin to the PA.

DR. MEADE: Again, I know I'm not the right person to respond other than I know that I think many options are being

considered. I think those studies evaluating all of the options for treatment and vaccine options and I think the status is to collect appropriate data that allows judgments to make.

I think others will make some decisions involving practicality, cost, and a lot of complex issues which I think will ultimately have to be taken into account. It is a broader level question that I think we can focus today.

DR. NASS: In response to that question -- I'm Meryl Nass -- there have been several companies that have developed monoclonal antibodies for use. One down the street in the small rodent model and showed that the use of the monoclonal was effective and also served as a vaccine so that when animals were rechallenged six or 12 months later they were protected.

There also has been polyclonal serum collected from a number of military vaccines that has been stockpiled by CDC so

that there is some ability to treat using that. In China polyclonal animal hyperimmune serum are used to treat humans who develop anthrax.

DR. MEADE: Thank you. Okay. I think I will turn this over to Freyja.

DR. LYNN: We're making up for all the extra time we had yesterday by running late this morning. I think in the interest of efficiency and good use of time why don't we go ahead and take a 20-minute break now and that way we can come back after the break and have the discussions and the wrap-up. It is now close enough to 10:00 o'clock. Please reconvene at 10:20. Thank you.

(Whereupon, at 10:00 a.m. off the record until 10:17 a.m.)

DR. NUZUM: Okay, so we are going to start our last session now. The agenda has changed a little bit. We are going to combine the two discussion sessions, one for PEP to conclude the previous session and then we'll have a final wrap-up, general discussion.

These are the panel discussion points for PEP. They are very similar to what Bruce had concluded with. I thought I would go over a few notes just from the questions that I took.

Maybe just to start the discussion, there was one on the challenge route, why we do aerosol. Keep in mind these studies are funded with biodefense dollars which means we have to develop products for biodefense indication.

As Drisilla pointed out, the animal rule requires that we do studies that support the label indication, and the biodefense threat is considered to be aerosol so we do aerosol challenge studies.

I think there was a question on why we are doing a vaccine for PEP or what the indication and so forth is. That largely has to go to compliance. I think the license regimen for antibiotics post-exposure is 60 days. I think in the last incident downtown they actually prescribed 120 days. I don't

think any of us thinks that people will take antibiotics that long.

The other component of that is antibiotic -- just from the standpoint of how much is in the inventory. If antibiotic duration could be shortened. The license regimen could be shortened and the antibiotics will go further. Those are a couple of reasons why we are pursuing this.

There was a question on long-term protection. This is probably addressed in Bruce's slides, but I think the next step for this is fairly straightforward. This is where we combine the GUP data where we have a sense of what protective levels are in a post-event scenario.

If you take an antibiotic and the vaccine and then we do studies, we'll know what the antibiotics in humans are. We can relate that back to the efficacy data we get from the GUP studies. I think it's going to show people are protected in the long-term. There is really nothing left to do. I think

that should fall out pretty well.

Kind of along those lines Bruce made a point post-event versus post-exposure, and he said they are essentially the same. I certainly agree with him that they are the same in that we almost never will know if it is truly post event or post-exposure.

However, I think we need to keep in mind they are very different scenarios. From a regulatory label point of view, yes, it's post-exposure because that is something you can clinically address. You can clinically test and we can do studies to design that.

In actual fact we won't know if people have been exposed or not. In that case the duration of immunity is important. I think we just have to keep in mind -- I think my main point there is that they are different and because when we are designing studies and the questions we ask in the studies are vastly different.

As with yesterday there is clearly a desire expressed to have NHP data and there

are questions about the antibiotic used. We definitely agree we need an NHP PEP model. We will do Cipro in NHPs and that should help address the antibiotic question.

Keeping in mind, and I know it has been said several times already this morning, the purpose of these studies are to show value added of the vaccine. It's not to emulate the antibiotic used or the antibiotic regimen that's used in humans and rabbits. That is not the point. The point was to get a model where they are partially protected by antibiotics so that we can show the added value of the vaccine.

That said, going back to my point yesterday that these animal studies to support the animal rule are really about people. Even in these studies we try to use -- and if they are refined and we learn more about the GUP, the dosing and so forth, we will come up with what I refer to as a humanized vaccine dose such that the immune response elicited in rabbits is similar to the immune response in

humans.

Similarly the antibiotic regimen we use in these models will be based on blood levels that don't exceed blood levels expected in people. That is kind of where we have -- we have to design the model to answer the question we want, but in so doing we keep in mind the human clinical situation.

There was a question on the antibiotic that will be given with the vaccine. Remember that whenever you start talking about antibiotics post-exposure you are talking about CDER. When we are talking about vaccine use it's CBER.

So our emphasis is just on the vaccine at this point and presumably once it has shown value and an indication is approved for PEP, that will simply be added to the current license. This is my opinion. This is a very regulatory oriented issue so I can be wrong and people can correct me, but this is my look on it.

What we are doing now is to get the

PEP indication for the vaccine that will be simply laid on top of the license regimen for antibiotics. If there are other studies down the road to change the antibiotic regimen, the post-exposure antibiotic regimen, that will then involve CDER.

It will probably another round of studies and discussions and so forth. I see it as a staged effect where we bring the vaccine in post-exposure. If there is a need to shorten the life of the regimen, that will be a whole other effort.

The vaccine, I would presume, would be able to be added to any of the antibiotics in the stockpile. I mean, we know very well their effect on the organism. All we need to show is that we can add the vaccine to any antibiotic or one or two antibiotics. I think at some point it would be safe to say it will work with any antibiotics.

Okay. I think that is all of my notes on the PEP. Let's start with this first bullet point and that is to comment on the

soundness of the design of animal studies to demonstrate the value added of vaccine for PEP indication. I'll do as Drisilla did yesterday and just open it to the panel if anybody would like to comment.

DR. LYONS: Yes. I just wanted to congratulate Judy and her team and Battelle and their team for doing a fabulous job. I mean, I don't know how many people can appreciate it, but these are extraordinarily difficult studies, and to do it as well as they are done you get lots of kudos there.

I think there is some confusion a little bit out there. Just remember this is a very artificial situation that they created to ask the question do you get an enhancement or not. Normally, as Steve was noting, people wouldn't be treated for seven days and then be vaccinated.

Those kind of things wouldn't happen. In order to see it statistically, you have to create the artificial situation to see it, so let's be clear on that. That is what

the design is supposed to do, bring out that phenotype of enhanced protection by the vaccine in this artificial setting, and it did very well, I thought. I mean, I thought it did that.

I think where a lot of the -- and then it comes a little bit to making that leap of faith. I think that is what everybody is struggling with to a certain degree. Some of the questions I heard had to do with things that we are very unlikely to ever be able to know like this question of where is a spore sitting.

The real question is what is the reason for the breakthroughs in nature. Why do people sit there with the spores in the lung and then it breaks through. Well, it's probably due to both host events, the genetics of the host, and certainly some of the pathogens in trying to make some prediction about that rare event after long-term treatment with antibiotics is going to be probably not in reality statistically ever

approachable.

I think we need to move on beyond those kind of things that are very important scientific questions that people will be working on a long time. But related to product development it is almost impossible. We just have to accept that fact that it occurs and move on. I think this design is an outstanding design and really brings it out.

DR. NUZUM: Thank you.

DR. FERRIERI: I'll comment next. I think the design is quite good in terms of our understanding of the rabbit model with set parameters, and it's a good starting point, but I think we need a lot more data and it may or may not be feasible to answer the questions you want answered. I think it's really tough.

The chaos of the immune response under conditions where you have antibiotics on board and you have the vaccine and some of the animals are going to make more antibody than others. I think it will be a very random event whether it's the rabbit or the nonhuman

primate. I think we have to accept that we are not going to be able to control things as we would like to do.

Ultimately I think the model can be transferred to nonhuman primate. It would be great if we had had more data to see data, but I understand the limitations.

Again, it's going to come down to the numbers and how can you maximize this part of your question and the design of the study to the nonhuman primate without wasting animals unnecessarily and can you strategize by maybe doing a few studies to zero in on the correct dose, for example, and capitalize on the data that has been presented with the rabbit to try to streamline the studies in the nonhuman primate expanding the numbers of those animals in a way that can give us more precise information on the immune response as best as we can determine it with more than one event going on, the immune response to the challenge, to the bacterial challenge that was given, the spores that were given as well as

the vaccine immune response. I think it is very complex, as you have all stated, but I think we need to streamline it.

I do feel these are valuable studies so I disagree with the member of the audience who doesn't think there is an indication. I think we should be prepared. I strongly support the efforts, and I want this to be part of the public record here. I strongly support the effort to come up with a reasonable approach that will give us added protection to antibiotics. I think this should be done.

DR. RUBIN: I would like to second the plaudits that were given for the execution of the study and the comments that Pat just made on the desirability of having this ability to combine antibiotics and vaccine in a situation where a real serious need arises.

Starting at that point I would like to highlight a complication of the animal model that was not explicitly considered. I would like to just develop that a bit. As you

saw, if you take these animals with antibiotics, as you go on the mortality of these animals without any vaccine decreases steadily. By nine days the mortality is 71 percent and certainly we know that if we had treated these animals further, at some point of treatment their mortality would have been nil.

You ask yourself what actually did you do when you did that. The answer to that, at least in my mind, is that basically what you did is you slowly decreased the challenge dose. All of these spores that were inhaled at first, 200 LD50s, were sitting there and occasionally one would rear its ugly head and germinate and would promptly be killed by the antibiotics.

Slowly there was a decay of the existing spores so that you could actually say, with not much confidence, but you could say that actually at 10 days the animals really had a chance of only 1 LD50 because only 50 percent of them died.

I think that if you keep in mind that the challenge is markedly altered during this process, it becomes even more difficult to interpret the immunological data in a sense of what do you need to protect. And in my own mind, I think these two processes should be separated. I would recommend that they be separated, that we get good experiments to tell us how long do you have to treat an animal with antibiotics to have such a probability, one or another probability of them surviving.

From the other studies that we have discussed yesterday we get the data of what is necessary in order to protect an animal irrespective of preceding antibiotics and how fast that response can be achieved.

DR. SELF: So it was very helpful to hear that this is really a test of concept sort of model, but it does raise the issue that we are not now just trying to bridge across species, but we are now also adding another bridge that we are building, a

completely different regimen, a completely different nature of exposure in some ways.

And all of this in the context of what seems to be a very complex temporal sequence of events. We didn't hear much today about the complexity of -- the underlying biological complexity of that sequence. We saw some kinetics.

The comment about decreasing dose was really interesting, and it raises some questions in my mind about the potential for thinking about modulating the challenge dose maybe in some experimental designs. I'm not quite sure what that design would look like but it raises that interesting possibility.

But I think, as I understand it right now, the potential for formally identifying some sort of intermediate outcome or surrogate and using that to motivate what would be a regimen, a combination regimen that would be used in people, that seems very remote to me right now.

It strikes me that an important

test of concept is made in rabbits that can and should be replicated in nonhuman primates, but I'm having a hard time seeing how there is going to be much more precision in characterizing the added benefit of a vaccine that would be relevant to the human setting.

It may be at the end of the day that the efficacy of antibiotics as it is delivered in humans would be seen as the basis that there is very plausible, based on these tests of concepts, added benefit from the vaccine obviously weighed by whatever sort of adverse experiences vaccination might give.

But that would at the end probably be the ultimate basis for any sort of recommendation for a combined vaccine. I'm just trying to be very practical about the limitations of this model in this complex setting.

DR. RUBIN: I appreciate all the comments that have been made so far and basically agree with all of them. It's a very complex situation, and specific suggestions

really are impossible in this kind of context right now, but maybe some overview comments may be helpful, overview comments about statistical approaches.

Much can often be clarified in complicated situations by embedding what you observe in data in a missing data context. That is, when you have an observed data set think about what you wish you had that would answer the question. It is a conceptual thought experiment.

Also it turns out that there has been tremendous progress made in the last few decades in dealing with those kinds of models and dealing with models with a tremendous amount of missing data. These models have made their way into astrophysics, into bioinformatics, into medicine, and so reviewing missing data very broadly.

In fact, there are some single papers in statistics that when you Google them they've had millions of hits that have been written in the last few years. That just

indicates the, sort of the generality and popularity of these approaches. Maybe one example to just indicate.

If, for example, we wanted to think of these randomized experiments where you're randomizing dose of yesterday, but you sort of wish you had been able to randomize immunogenicity which you just can measure.

You can figure that in a missing data context in terms of noncompliance which is another term that came up today. There has been tremendous progress made on noncompliance in the last decade in statistical analysis with noncompliance.

What you can think of as is you wish you could randomize immunogenicity, but you can't or you didn't, but when you gave a high dose, you were trying to encourage high immunogencity. And so rabbits that were given a high dose and had relatively low immunogenicity were sort of noncompliant in that sense.

Rabbits that were given a low dose

and had high immunogencity were sort of noncompliant in that sense. They didn't follow their encouragement. Those kinds of models have been used now for a decade and with great success in a variety of situations.

Although I agree with Steve that it seems to be a hopelessly complex situation because there are so many different features in this PEP environment, although I'm not sanguine about it, I'm more enthusiastic I think, perhaps, about the possibility of making progress if the problem is laid out correctly first and then see what you are missing.

Because very often when you do lay it out in this missing data context you see there are some key pieces that are missing, but they are not so hard to get by doing the right experiment. Until you lay it out seeing what all the pieces you need are, it's impossible.

You are sort of thrashing around so it's like you wish you had this and wish you

had this. You don't have the full picture yet so ideally we'd have a meeting with people who understand immunology, which I don't, and some statisticians. Maybe we would lay out what we are missing and what we would like to do and maybe even discover that some of the missing pieces can be discovered in relatively simple experiments.

DR. FERRIERI: I would like to follow up on Don's remarks. There is a wealth of fantastic expertise in this audience. What I'm missing, Ed, right now is the plan moving forward who is going to be doing the nonhuman primate studies? Will it be one of the sponsors? Is some of it being done by NIH or FDA at Fort Detrick? Where is this work going to be done?

Is everyone talking to each other to plan the most precise experimental studies so the next time we may convene here the presentation of the data will be done in a way that is laid out comprehensively and in a manner that will permit at least me to digest

it more easily?

DR. NUZUM: I would comment on that. Certainly NIH is planning to move forward with the NHP studies as we've done with the rabbit studies. However, it is getting more complicated as there are more funded companies developing multiple products, monoclonals, polyclonals, vaccines. It is an area that we need to look at in the government.

There's a lot of discussion on that very point. I kind of alluded to it in my talk that is where we need planning to maximize resource utilization. That is a very good point.

DR. SELF: This is a bit of a tangent, but it might be a good time to interject. Yesterday there was a lot of data presented but not all sort of analyzed in the same day and aligned.

I guess one of the things that I would encourage you to do would be to set up some process for maybe an independent

statistician, maybe Bob or maybe somebody else, to get those data sets, analyze them in a consistent way and bring those results together as rapidly as possible. There's a lot of, I think, knowledge that is there in that data that hasn't been extracted yet, and it would be very simple to do that.

DR. NUZUM: Thank you. Good suggestion. Maybe we probably should move to the next bullet. Very good comments. The next bullet is take into account the complexity of the PEP system.

Discuss whether protective antibody level can be estimated from animal PEP studies that can be extrapolated to humans. If so, how might this be done? If not, discuss the strategies for linking animal protection data to efficacy in humans for the PEP indication.

Any comments? It sounds like there is general agreement that everyone acknowledges this is a very complex model. I guess my feeling, and certainly Bruce was very clear, but there is a lot of other data

besides death we can look at, bacteremia, temperature, clinical science, other endpoints. I think there are ways to strengthen the model in the absence of a clear-cut correlate is my opinion. Would anyone comment?

DR. FERRIERI: I would just kick it off and see how people react. It's not going to be possible in my opinion to have a very precise understanding of the protective level in any of the animal models and be able to extrapolate it to humans where the complexity of the immune response is going to be even more unregulated, if you will, or unpredictable, whatever language you wish to use, randomized, etc.

I think we need to be very clear that there are limitations here. We accept those limitations and that we don't drive this into a hole continuing to expect more than is possible from the model in terms of having "the antibody level," the TNA precise level.

A term I use in clinical medicine

at times when I'm seeing certain complex patients. We have to guesstimate. Perhaps the best we can do is that rather than something very precise. There will be a range, and maybe that will be acceptable.

DR. LYONS: Thank you. Ed, one of the things that I was thinking about, at least for those of us who trained in the older days, I mean, when one was exposed to hepatitis and hadn't been vaccinated you were basically treated and vaccinated at the same time. I mean, at least, if I remember correctly that was the plan.

It would be interesting to go back and see where that strategy actually came from and what they used to try to -- what was their correlate for protection in that situation. I don't even know if there was one. It might have just been it doesn't hurt to just do it all and it was done in sort an off-label basis.

I don't know. The precedent for doing this post-exposure is there in other

diseases, and hepatitis B is probably the one I'm familiar with. And so it might be worth going back to look and see how they thought about that process.

DR. SELF: I agree with the comments about not getting too precise, but it does occur to me that if in the context of this model you build the same sort of logistic curves relating survival to TNA levels or whatever your favorite antibody measurement is to see if those even approximate those curves in the gut model, that would add a little more comfort at the end of the day even if not much more precision to calibrate what the effect would be ultimately in humans.

DR. FERRIERI: I agree, Steve. I'm not trying to convey that I'm not going to apply rigor and try to squeeze the most out of all of this. I have a high standard and high expectation of all of you who will be doing the work.

To follow up on this point about making corollaries with the hepatitis model,

we have to remember that this is an overwhelming infection. Hepatitis doesn't get into the lung causing acute inflammatory response. The whole pathobiology of the infection is quite different with anthrax in the rapidity and the damage that is done.

One has to -- you have to intervene rapidly. Hence, the necessity of having the combined approach, in my opinion. The schedule has not been commented upon, and zero and seven days were used in the rabbits. I don't know what the plan is in the nonhuman primate, but I would assume it is somewhat similar.

I think zero and seven sounds good to me. I wouldn't be playing around with longer stretches necessarily of the possibility in humans of revaccinating again after seven days to perhaps shut down the need for continuing antibiotics for three months, four months, six months, whatever.

Consideration should be given in the NHP model to perhaps continue in the

survivors to do that to boost the immune response. I would be interested in reactions from those who will be doing the studies.

DR. NUZUM: You mean boost with another vaccination longer out?

DR. FERRIERI: Perhaps three weeks later to do it again. Zero, seven, and then perhaps three weeks later or whatever. Has that been considered?

DR. NUZUM: Well, not at this point, but it is a good point. Again, I guess I would say we are kind of still in the basic model development. We have what we think is the basic model, and certainly there is room for tweaking and refinement.

But at the end of the day, and I think it was mentioned yesterday, the human regimen will depend on human data. We can look at as much of that as we think is practical in animals, but if the human response isn't the same as animals, then that is going to drive the final choices.

DR. FERRIERI: Of course.

DR. HEWLETT: We are separating the post-exposure from the general use, I realize, for specific reasons, but I want to go back to what Emil said. I think that we're talking here -- even though it's animals and not humans, we are talking about -- there is no doubt that antibiotics work as long as you keep taking them.

There is no doubt that the immunization works depending on repetitive onset. It seems to me that we are talking about protocol optimization.

Looking at those and getting the best -- especially since there wasn't any indication of interference, at least in the studies that were demonstrated, optimizing those protocols ahead of time and then combining them because you have so many variables right now if you do it that way and then put them together in a way to show that in fact -- because the idea obviously is to minimize the duration of antibiotic use and to optimize the rate of onset of the immune

response. Those can be worked out individually.

DR. NUZUM: But it seems like that's what we did. We did do separate antibiotic-only studies and then used the information from the GUP studies to select vaccine dose. So, I think I am missing something.

DR. HEWITT: Judy Hewitt. I can add to that. I think maybe another sort of perspective comment on that. The pressure to develop a post-exposure model several years ago was so great that we felt that we couldn't spend the time to carefully to optimize both of those situations, so we sort of moved forward in a way that sort of got us to good enough as quickly as possible.

Having said that, you know, we dropped the start-time study in the very beginning. We just chose a start time. We are in the process now of going back and looking at that, did we pick the proper start time.

That data is not ready to present yet, but we are going back now and looking at some of that optimization, those kinds of studies. I think it's not that we are ignoring that. It's just that we sort of had to -- there was too much pressure to get there too quickly.

DR. HEWLETT: I understand those constraints. I think this takes me back to a question I wanted to ask a little bit earlier. We are talking about the ideal and the practical, and you are dealing with both of those at the same time.

I've gotten progressively confused as we went through all of the data here being asked questions like what additional data are needed when I don't really have an understanding of what the endpoint is. I realize licensing of vaccine and knowing the ideals, but we are not going to come up with a V&A cutoff number.

We are not going to come up with any of those specifics from this information,

and I'm trying to get a sense of what it is that we are trying to get to by asking the questions of what additional things do we need. Do you understand what I mean?

DR. HEWITT: Absolutely. One example when the question was first brought up, what do we wish we had, we have, for example, in the PEP studies on the survivors, the rabbits that survived post-challenge, we have weekly ELISA and TNA values, so we can't really measure the kinetics of those responses on a scale any shorter than week by week.

I wish I had interim data points between those weekly samples. Sitting here and having this discussion brings that very clear to my mind that maybe we need to do a study where we look at more of those interim data points.

We don't need data points out at 28 days and beyond. Let's get some more data points early on in the response. I mean, I think there really are some good pieces of data that we can sort of focus and try and get

next.

DR. HEWLETT: But how much? Talking to the regulatory people, how much do they need to have because we're not going to set levels here and now. We just want to be adequately prepared when a licensed application comes so that you can judge properly where it is. Again, this is the combination of the practical and the ideal. That is the intersection point.

DR. BURNS: Exactly. Drisilla Burns. What we want to hear from you, the panel, from the audience, is to get scientific input on how meaningful you think these models are in regards to being able to support efficacy of a vaccine. You have seen the model presented as it is with the types of data that are being generated.

So the endpoint is can the model, in your viewpoint, support efficacy. Any given vaccine could give a different result in the model, but the model itself, how good is it, how strong is it. Do you need something

else in your mind that would be -- and we could do the need to know, nice to know sort of -- two categories. Does that answer your question?

DR. GOTSCHLICH: Yes. May I ask you another question? Why is this being separated into a different box? Let's just assume for the sake of argument that we have an effective vaccine. Why do you need a different box for an effective vaccine after a person has been exposed to anthrax and in order to be certain that that person did not get the disease was treated with antibiotics? Why do you need a separate box?

DR. BURNS: Well, I think in general it is a different indication that would go on the label.

DR. GOTSCHLICH: Why is it a different indication? The indication is to prevent anthrax in a person. You have two treatments.

DR. BURNS: And one indication is you are exposed after you get the vaccine and

one is before. Now, I would love to hear any comments you have.

DR. GOTSCHLICH: I'm still not clear why it is a different indication. You have a human being that requires the vaccine.

DR. BURNS: Yes.

DR. GOTSCHLICH: In one case prospectively, in this particular case, there is an additional reason why that human being requires an effective vaccine. Why is it a different indication?

DR. BURNS: I think that we want to hear your thoughts on why you think it's not a different indication. We would like to hear your thoughts. What is the science that would say that if you have protection before exposure you are definitely going to have protection after exposure? I would just like to hear your thoughts on that.

DR. GOTSCHLICH: Well, I will repeat them very briefly. Basically it seems to me physicians have two things in their hand. They can treat the disease with

antibiotics or the prospective disease with antibiotics and that is fundamentally slowly eliminating the challenge dose, or if there is no preexisting challenge dose in them, give them vaccine, and if they are challenged they will have enough antibodies aboard to be able to handle their challenge. I see these as completely separate processes.

DR. CLIFFORD: Julie Clifford, FDA, CBER, Office of Vaccines. I'm not sure I'm understanding exactly the question, but I'm going to offer something that might help.

In terms of the vaccine or any vaccine for anthrax, the indication and usage section of the labeling, once it's approved, will say that it offers or provides protection after a certain number of doses based on the clinical data or the animal model, you know, the combined licensure package.

But in a post-exposure setting to someone who is naive to vaccine, the chances of them achieving what we would hope to be identified as a protective level of antibody

or an ability of their immune system to respond enough to protect them may not occur until after for this disease, in particular if they have already succumbed. That is why it would be seen as a separate indication. The preexposure indication would require a certain number of doses based upon clinical and nonclinical data to support licensure whereas a post-exposure we are still trying to get a handle on how you identify the number of doses and schedule in a post-exposure setting and how somebody could be protected.

Right now we have a separate component of antibiotics that affords kind of an immediate protection, and we are struggling with how to define the use of a vaccine that would add an added component in conjunction with the antibiotics. Does that help?

DR. GOTSCHLICH: Thank you.

DR. CHAWLA: I will just give an example. Maybe it can help. We have a rabies example where the preexposure indication is something different. You follow a different

immunization schedule. And in post-exposure where it is human with immunoglobulins or without immunoglobulins in post-exposure you follow different kind of schedule.

So it is -- if a vaccine has to be shown efficacious in both the conditions preexposure or post-exposure, a different set of animal data or efficacy data is required.

DR. NUZUM: I had made -- and I skipped over because there was a question earlier about incorporating either monoclonals or polyclonals in these models. And that is being done. Remember again, and it was shown, there is preexposure, post-exposure treatment.

We are not talking about treatment models. Those are also being worked, but it is a whole different data set you start with. It's a different model set up and endpoint. Yes, that's being done but, you know, we're not talking about treatment.

You have been standing there a while. Go ahead.

DR. FRIEDLANDER: I would just

reiterate what Rick said that this is an excellent approach in beginning to address this important issue. I would also reiterate, I think, what Emil said. In its most simplistic sense I think you could make the argument that you have made, what is the difference between pre and post-exposure.

Why not just consider that you are using the antibiotics only to get the individual to the time when you have already established in a preexposure model that the vaccine is effective.

Assuming or demonstrating that the antibiotics do not impede the development of the immune response, you basically use the same regimen you have used or you could modify the regimen but use the same regimen you used as an indication.

What are you doing when you treat -- when you vaccinate post-exposure? You are not using the vaccine for the immediate exposure. You are using the vaccine for the exposure that occurs once you stop the

antibiotics.

If you know in your mind what that time is when you develop a protective immune response, that is the time that you need post-exposure to be on antibiotics. I think you can argue that case.

DR. NUZUM: So, very simply, I think I would summarize this by saying if the vaccine is shown to be safe and effective on its own, antibiotic is shown to be safe and effective on its own, and then you can show that combining the two don't provide any adverse effect, what is the problem with using both of them? But then I think we are talking about clinical trials, not animal studies, to really show in people that there isn't an adverse effect.

DR. FRIEDLANDER: Absolutely. You would have to show, I think, that it's safe and that it doesn't impede the immune response. Those are fairly easy studies to do.

I would like to just make a couple

points that came up before in terms of this business about the different models and the duration of antibiotics because, first of all, we know they vary in different animals, and there are functions, as we said before, of the dose, of the challenge dose, and of what the determinants are for germination or retention of spores in different species.

We don't know that. Anecdotally you can think they differ in the rabbit and the nonhuman primate, but we don't know that. But this idea of finding a measure of antibody that we believe confers protection is, I think, very important because if you knew that, you could tell someone when they could go off antibiotics if they had been exposed and, say, developed some infection.

Or if they were vaccinated, or if they were treated for established disease and developed an immune response you could say, "You don't need to stay on antibiotics any longer because you have now been immunized whether from active infection or from vaccine.

I think --

DR. NUZUM: Or both.

DR. FRIEDLANDER: Or both. Exactly. That is exactly right because we think that is what happens.

DR. NUZUM: If it's truly post-event, and that is what I was trying to get to and I got off track on post-event, post-exposure. If it's truly post-exposure, your vaccine is actually boosting infection. If it's post-event but not actual exposure, then there is no danger anyway.

DR. FRIEDLANDER: You are absolutely right.

DR. NUZUM: You won't know.

DR. FRIEDLANDER: But whether you develop an immune response, we know that is a function of the challenge dose when you are on antibiotics. Some animals with a high challenge dose on antibiotics who are protected develop an immune response.

Others, when the dose is low or the duration is long enough but the challenge dose

is low, the antibiotics completely suppress the immune response and that was the genesis for the whole argument for vaccination. So I think this is the right way.

I would point out, as I mentioned before, concurrently with this there was a preliminary experiment that was done in the nonhuman primate, as you know, that gives credence to the idea that vaccination coupled with antibiotics would enhance protection.

DR. HEWLETT: Aren't there three variables here for the vaccine? There is the rate of onset, the magnitude, and the duration. The idea here with the post-exposure is to have the rate of onset faster to get to the same magnitude and you don't care about -- maybe you don't care about duration.

But it may be that you could have the same regimen that would accomplish both of those. Maybe you all know that already, but if you could accomplish that and had long enough duration to be useful, then maybe you

only need one regimen.

DR. NUZUM: Again, the regimen will be dependent on clinical studies.

DR. FRIEDLANDER: And it depends on the model. In the nonhuman primate we know that the current regimen, at least with 14 days of antibiotics, was sufficient. Now that needs to be expanded on.

DR. NUZUM: Okay. Quickly.

DR. NASS: I think another difference between the pre and post-exposure models is that when you are vaccinating people pre-exposure you are going to have antibody levels going up rapidly after their doses and then falling off over a period of time, and they may be exposed at any time during that fall-off period.

And so you generally will not expect the vaccine to be 100 percent efficacious at any point in time later or until the next booster dose. However, if you are using it post-exposure and you are talking about taking people off antibiotics because

you vaccinated them and you think they may have spores in their lungs that have not yet germinated and that they may or may not have an adequate endogenous immune system to fight off those spores like the 88 year-old-woman in Connecticut who probably got just a couple of spores but succumbed none the less because she didn't have much of an immune system.

You are required to have a vaccine that you will get high titers because you will be at the beginning of the immunization period, but you are looking for 100 percent protection if you are going to pull people off antibiotics. I think your expectation of the vaccine is much higher for post-exposure.

DR. KAMMANADIMINITI: Srinivas from Cangene. I have one general question. I understand that in order to be able to demonstrate added benefit of vaccine you are not giving antibiotic treatment to its full potential.

DR. NUZUM: Right.

DR. KAMMANADIMINITI: My question

is is this approach acceptable to the regulatory agencies for licensure of vaccine for this indication?

DR. NUZUM: I think what I can say on that is these models have been developed with active interagency participation. These aren't done in isolation but what will ultimately be acceptable to FDA is the data that comes to them.

The reason we are doing this is it provides the public to see what we are doing. FDA gets feedback and really a large purpose of this meeting is to get public feedback that helps FDA determine if this is satisfactory what we are doing.

DR. KAMMANADIMINITI: So it is not yet clear if it is acceptable or not?

DR. NUZUM: It's clear to me that it's not unacceptable.

DR. KAMMANADIMINITI: Thanks.

DR. HEWLETT: When you say not using the antibiotic to its full potential, are you talking about just continuing it

longer? I think we are here again at the intersection between the practical and the ideal, the theoretical, because people won't take the antibiotic for prolonged periods. There are complications of doing that. It is the reality of where is that balance point. I don't think it is fair to say you are not using the antibiotic to its full potential because that is difficult to do in reality.

DR. KAMMANADIMINITI: So we can simulate a clinical scenario artificially in the animal model.

DR. HEWLETT: As close as we can get. With the stipulations on the antibiotic pharmacokinetics and the immune response we elicit, you know, that's where we try to bridge as much as possible to humans.

At the end of the day you've got to do -- we've got to do -- I mean, the rabbit -- as Jim Eastep would say if he was here, rabbit is not man and neither is a nonhuman primate. In animals we simulate it the best we can to answer the question we want.

DR. NUZUM: I want to give Conrad credit. I made the comment about the clinical trial to look at vaccine and antibody in combination. This has actually been something that has been discussed for some time. Actually during the break he brought it up so I thought it was a good place to interject that.

Let's make sure this last bullet is covered and then we'll move on to the general wrap-up unless there are other comments on the second bullet. We have already touched on this a little bit regarding what additional data, if any, would be supportive.

There is a comment that to help answer that question we really need to know what the end game is. Certainly that would help. I think it is hard to know what that is. This is the first time we -- as far as I am aware this is the first time vaccines have gone this far using the animal rule.

It's new for all of us including FDA so I think it's very hard to give the end

game at this point. I think meetings like this and getting the data out and kind of giving everyone an idea of what we are doing is the best way to help us get to that end game.

I would maybe qualify this question a little and say in the near-term, or in the next steps, or what should we do immediately that we aren't doing to provide data that would be supportive or any ideas on near-term or next steps.

DR. RUBIN: I think one of the first things to do is to try to structure the data that you have in a factorial setting. You have so many factors running around. You have species. You have the vaccine, the dose, the time and duration, antibiotic dose, time and duration, challenge dose.

You get a million factors. Perhaps a dozen factors running around. As you structure the data that has been obtained so far in a factorial setting, you will see that you have some of these cells covered and some

of them not covered, I think. I'm just curious.

Has anyone ever thought of trying to design one study that covered several animal species and varied lots of things at the same time and did a fractional replication? Does anyone know what a fractional replication is besides us two?

When you have many factors, let's suppose you have 12 factors like a factors vaccine and each one has two levels, the high/low, antibiotic high/low, animal species, rabbit, nonhuman primate. Another factor, another factor, another factor.

So you have 10 factors, each one at two levels and you have over 1,000 different conditions. Does that say you have to have data in each of those 1,000 conditions in order to understand almost everything that is going on? No.

There is a famous quote that goes back about 75 years to Fisher that says something like, "Nature will reveal her

answers most quickly if you ask many questions at the same time." This whole scientific idea of doing one question and then wait for the next one and ask that, that is totally retarded.

There is literally at least a half century of work in industrial experimental design coming to the United States and post-War World II Japan that supports that idea that you don't do it that way. I'm wondering here where we have all these factors where is the experimental design guy or are all these people making up these designs learning it on the fly?

If you don't have this real experimental design guy who knows what fractional replication is and how to analyze fractional replicated data and knows about aliasing and all these other technical terms, you're not doing it right.

DR. SELF: You've got to mention the assumptions underlying this, right?

DR. RUBIN: Yes. The assumptions

always are that there are very high order interactions like 10-way interactions are, even if they exist, are un-understandable anyway which is true. There is a 10-way interaction between species and vaccine dose and antibody level, and time -- you never can make use of it in any case. It basically says nature maybe has like three or four-way interactions and that's it. The answer will be revealed most quickly.

DR. KOHBERGER: Don, the issue until this point, and the reason it wasn't done like that was choosing the levels of the factors was unknown. We weren't quite sure what the antibiotic duration should be. We could have chosen 30 days and 20 days. If we had done that the experiment would have been noninformative because they all would have survived.

Until this point I don't think there was sufficient information to know what factors were important and what levels should be important and what the endpoint should be.

Now, I think it has come along far enough that thinking about one big grand experiment is possible now where six months to a year ago it wasn't.

DR. SELF: In terms of additional data one thing that occurs to me that would be interesting and useful would be to follow up on this idea of dose and antibiotics, sort of reducing the challenge dose to look at the short regimen, zero/seven, the timing of the challenge dose that might be varied around a range of times reflecting either shorter courses of antibiotics planned or realized by sort of noncompliance of humans to the 60-day course and to look at those times of the challenge at a range of doses that might be commensurate with what you would expect given a course of antibiotics delivered up to that time point.

That might look at the specific regimen that is being described, the short course. It would look at antibodies and drawing those curves in the absence of

potential interference for antibiotics so it would be comparable to bridge to some of the GUP data and then could also be used perhaps to compare to those curves that you would draw in the context of the combined regimens. I think that would be an experiment that would maybe link to both sides of this gulf.

DR. NUZUM: Thank you. I have to go back to the pathophysiology, I'm afraid, just for a minute because I've been thinking about the issue of with all the data that have been collected, and I think the studies are really remarkable, I keep wondering if there might be a stratification.

Certainly patients with anthrax have a number of different disease manifestations. Some proportion of them have meningitis, for example. I keep wondering if in the animals that have died have a higher proportion of meningitis than in the whole population.

When we are talking about TNA levels that there might be levels that are

related to certain disease manifestations but if your level is here and not here, you are more likely to get meningitis or some other -- liver damage or one of the other manifestations of the illness.

Not to propose more studies, and I'm not sure whether this is in Drisilla's need to know or nice to know, maybe it could go from one to the other, but with the data that have been collected already with the post mortem information and clinical chemistries and liver enzymes I wonder if it would be possible to look at that level of stratification.

DR. NUZUM: I'm trying to think. I think we have done some of that, but I don't know that meningitis is specifically one of the endpoints we've looked at. We do have -- it's a good point. We do have a lot of data like that. We could definitely look at that.

Are there any other comments on this last bullet? Otherwise, we'll go into the general wrap-up.

What I wanted to do here was I have some comments, summary comments, that I've written down, but I also wanted the panelists to each have a last word. The question is do I go first or would you prefer to go first? You want me to go first? Okay.

DR. HEWLETT: You go first.

DR. NUZUM: That's fine. On one hand, I didn't want to steal anybody's thunder but, on the other hand, this way you have the last word.

I think overall one thing we wanted out of this meeting was a sense of whether or not we are on the right track. I think we can say we have heard no major disagreements with the path we're on either for GUP or PEP.

Our basic principle for protective titers for GUP and to show added value of vaccine for PEP seems to be relatively well accepted. I think there was lots of good discussion, questions, comments on what we are doing, and they will provide some good guidance going forward.

The second point I wanted to mention is clearly there is a need for the NHP data, both GUP and PEP. There is also a need for complete characterization of the immune response, cell mediated data, innate immunity, adaptive -- other ways we can characterize adaptive immunity and so forth. Even looking at genetics perhaps.

Actually, some of those efforts are ongoing, and we couldn't present all that here, and that wasn't the purpose today. But certainly to support the animal rule requirements we know we have to look at all of these endpoints to the extent we can.

There was a lot of enthusiasm for passive transfer data and that that work should continue. There was a suggestion to use monoclonal antibodies. I have had an off-line discussion or two that indicates the conclusion may not be universal but certainly among the panelists there was much interest in the passive antibody work.

It may lend another off-line

discussion I had on passive work. Even Freyja, I think, mentioned this. The passive work may be a way to provide that randomization of the immune response that was discussed at length.

Another good suggestion I thought, Rick mentioned this, was the idea of a subgroup. I have biologists and statisticians together to not only address this specific anthrax model issue but the animal rule in general, other counter measures and so forth. I think that is a very good suggestion.

There are two groups that don't get together that often. I think if we have heard anything today or in this workshop, there is a need for those groups to be together because this model development to support the animal rule is certainly going to be heavy on statistics and require good statistical analysis and the right statistical analysis.

There was a long discussion about randomization of immune response so that we can evaluate what is going on independent of

vaccine dose, double randomization studies and so forth. I don't really know what that means, but I'm sure Bob can help me.

I also heard -- I didn't hear consensus on that point because I think the biologists in the group felt -- there is a sense that if you get similar titers, similar protective titers in multiple studies, multiple species, different labs, different vaccines that is very important. How you get to that titer is less important that you get to that titer by a variety of different means.

It was a very good discussion and I think very thought provoking for all of us. Clearly an area where I think the statisticians and the biologists are a little -- not completely in line.

I think Judy and Jason made the point this morning that best laid plans invariably don't go as you hope. Things take longer than you expect. All of us in science know that is the way it is. We have hard proof here that, yes, it's reality.

Regarding the PEP discussion, I think Bruce made a very good case. A clear-cut correlate will be very difficult to get. I think there is general agreement, or I have a sense there is general agreement amongst the panel that that is the case.

There were some good suggestions on how we might modify our studies. Thinking about what is missing might help guide our studies. Was it fractional --

DR. RUBIN: Fractional replication.

DR. NUZUM: Fractional replication. Okay. I mean, there is a lot here I don't understand statistically. That is the reason for this subgroup of biologists and statisticians getting together. It's too soon to know how that would happen, but I think it is a very good idea.

Okay. Those are my summary points. Maybe we will just start at that end of the table. If you have any wrap-up points, we would be glad to hear them.

DR. SELF: Okay. I think that even

though there is a lot of detail being reported in understanding the rabbit model, ultimately the most important information is going to come from meta-analyses across these studies and across species.

I would say that a mechanism to much more rapidly bring data from different studies together, analyze it in a standardized way that doesn't have to be overly complex and present it in that sort of aggregate is probably the most important thing that I see missing today and would probably be the most helpful in planning the next steps forward.

I think the ideas that Don is talking about specifically in terms of being very explicit about what assumptions are being made is going to be useful and I'm not going to say anything more because I wouldn't dare with Don sitting to my left.

I think the GUP studies are on track, but I just have a general sense that there is a little stall because these data haven't been brought together. I suspect

there is one that sounds like a very good data set from nonhuman primates.

I worry that once that is lined up with what we know from the rabbits that there is going to be enough of a discord to have to consider seriously some experiments in another species. I know this is a big effort and a big issue but to get that sense sooner rather than later I think would be important.

Then finally for the PEP, as the model is currently configured I think it provides great test of concept information and perhaps some context, but I am worried that the problem is being too narrowly defined. I think it was defined earlier as just what a vaccine can add to a given regimen of antibiotics.

The suggestion at the end of the table to broaden that field and talk about combination regimens and be able to look at both pieces opens up, I think, a whole series of important questions that could be addressed better by modeling and ultimately might have

more relevance for humans.

DR. NUZUM: Thank you.

DR. RUBIN: The first thing that I would suggest is to have a glossary of what these three-letter acronyms mean. I would figure it out if it is ever written down just by having one page that I can look them up. I'm serious about that. It would be very, very helpful. It's not a very statistical suggestion, though.

I think the first activity that would be very useful and this is very similar to a comment made, I don't remember by whom but it was already made here, to embed all the current data into one conceptual large factorial experiment where you could weigh out all the factors.

Here are the species, here are types of vaccinations, for example, active, passive, so those are different levels, and where you have data and where you do not have data and what species, and antibiotics given also at the same time situation, challenge

doses, and so forth.

It's very hard, at least for me, to keep all the different studies in mind, all the different levels, and all these different things without having them laid out in some organized way. If they are laid out in some organized way, then often you will see that there are certain cells that are missing where you don't have data and why not.

Maybe it was just an oversight. Maybe it was unimportant. Maybe it is unimportant. To keep it all straight I think you really have to get it organized. Always realize that the intention, everybody keeps emphasizing this but I'll emphasize it again, is to extrapolate the results to this one layer of this big factorial table, it's three by five by 10 by whatever it is, to one layer called species that is called human.

The more similarity you see in main effects and interactions and relationships that you can observe in humans and in other species, the more confidence you will have

that these extrapolations to humans where you don't have data on humans, where you can't get data on humans will also be similar.

If you find big changes across species and think that it varies even with humans, then that is evidence that you're not there yet. But to the extent that you can get this table of data in such a way that you see understandable relationships, I think that will be very important to do.

I'm not at a level where I have enough feeling for what that big data set with all sorts of missing cells looks like. But always realize you are going to have to make assumptions, and whenever you are making assumptions just be explicit about them because then you can have a debate among scientists. Not amongst statisticians but among scientists who really understand whether those assumptions you are making are plausible.

The final comment is that even though I've been making fairly strong

suggestions, I think what has been done so far is really a great job. Very carefully done experiments with lots of thought behind them.

I think in some cases they possibly could have been better with some discussion with statisticians who know about design, but that doesn't mean that I don't appreciate the work that has been done so far. I think it really is a great job. But future work, I think, could be improved if we got together more. Thanks.

DR. GOTSCHLICH: I feel like a lemon that's been squeezed too many times and there's not enough juice left. I think the data and the approach on the general use prophylaxis is obviously on track.

I think very much the suggestion made to provide this in an integrated way so that it is understandable by everybody, even those that are significantly statistically challenged, but I think it's on track. I must say that in the other part I tend to remain stubborn.

I think it's two separate problems that should be dealt with separately. It's antibiotic treatment, and it is prevention of anthrax. I think you will get further by thinking of these as two separate problems. If you want an indication, a different indication, then I think one could think of one that is an accelerated immune response.

Then what you would do is simply design an immunization scheme that achieves that in terms of the protection data available from the general use prophylaxis. I think the experiments were very well done, and they were very worthwhile in exactly pointing out that it's best, in my mind, to separate the two problems. That is my opinion.

DR. FERRIERI: I'll be brief here. I want to thank all of you who organized this. It was wonderful being here and hearing the data. I think what we've heard is extremely positive and gives me a great deal of confidence in how you will move forward and we will be able to accomplish what your

original goals were.

I guess my concluding remark really is that all of you who have all the smarts, who have worked in the past, who are currently working on the problem, all the stakeholders in the room need to continue to be together talking through all the issues and sharing your information, or there will be obstacles in finalizing a lot of the goals.

One final comment on PEP. I sort of agree with Emil on that. I'm not sure I quite understand him but there was a dismissal of someone who brought up some antibiotic issues, "That's treatment. We're not going there. We are dealing with prophylaxis."

I would contend that the approach being used if someone comes in post-exposure is treatment and prevention. We need to keep that in mind as we move forward. They are inextricably linked, the issue of the antibiotics as well as augmentation of the immune response. They are not separate.

This is not just all prevention.

There is an element here of treatment so if someone comes in, they are going to say, "How are you going to treat me, Doc?" Or someone will say, "What was your treatment?" It's really PEP but linked with that is the notion of treatment that cannot be dismissed and separated.

Again, my thanks to all of you.

DR. NUZUM: Maybe I'll just comment on that. Yes, Pat, I agree completely. It kind of gets back to are you dealing with post-event or post-exposure. I always prefer to say post event because we are not going to know if exposure occurred and MDs are going to use whatever is available.

Depending on when exposure occurred, where the course is they could have the clinical disease that -- well, it wouldn't be clinical yet but bordering on clinical disease and it is treatment. I think that is all going to be teased out between the GUP model, the PEP model, and the treatment model.

I think that is all going to be

addressed. For example, I mentioned monoclonals and polyclonals for treatment. They will be looked at for PEP as well. I think that -- maybe we haven't emphasized, but all these years of vaccine model development work is moving or being transitioned into the treatment model work. It's not lost data.

DR. LYONS: Yes. I want to thank everybody, too. I just wanted to reiterate something Art said because the more I think about it, the more I think if there were any more numbers or animals that were going to be used for different things, I think trying to get to a statistical robust measure of the protective level of antibody may be one of the most important things that comes out of at least the GOP studies because I think that is going to come down to being very important for the PEP studies in the future.

If I had more animals put toward anything, I would talking to people like Don and figure out exactly what would satisfy him and Steve and satisfy them.

DR. NUZUM: Is that possible?

DR. LYONS: Not me. I don't speak Don's language I don't think. I think that is where I would focus the initial studies.

Then, secondly, I think Don raised a very good point. I think it is interesting when you listen around the room is that there are -- most people are making assumptions about the models and about what we can know and what we don't know and what we will likely never know including myself.

I don't think we have ever really listed that for anything. That is one of the reasons I think the benefit of having that meeting would be good because I think a lot of people have the same assumptions but let's get them down so that we all understand what the playing field is.

Then, just finally, you know, I really think this is ground-breaking type of work, and I appreciate NIH doing a very good job. To me vaccines and particularly for the next 20 years other than the dominant

infections we all know about for emerging infections, though, this may be developing the paradigm for how we do things in the future. I can't tell you how important it is to try and develop a flow chart that we can use. Thanks.

DR. HEWLETT: I certainly have learned a lot through this exercise, and I think just very briefly it is very clear that the vaccine works. It is very clear that above some certain TNA that that's a good indicator of protection.

It does make a difference, however, whether the antibody titers are going down or up at the time. I think that is the difference between the passive and the active.

Bruce and Drisilla both said to me the objective of this is to get scientific buy-in to the concepts that are the basis for this. I think there is no doubt about that. I believe we are working on refinements of that rather than changing the concepts.

Finally, there is not any doubt to

the fact that ultimately there has to be a leap of faith. It's a matter of how close we can get, how wide the gorge is but ultimately it has to be that leap of faith and there's not any getting around that.

DR. NUZUM: Thank you. One thing I'm thinking of, coming across the table, several of the recommendations that come out of this could be combined. The data meta-analysis, the table that summarizes all the data, that could be a meeting in itself of the biologist and the statisticians it seems to me. It might be the next meeting whenever that would be.

Are there any comments from the audience? Does anyone out there want a last word? We are just about on time. I'll just wrap up with a couple things. First of all, Sonia Gales isn't here unfortunately. I want to acknowledge her help in starting the planning for this meeting. This has been a very good meeting, and she helped lay the ground work for it.

I especially want to thank Freyja who came in later on here and really has done the heavy lifting. For those of us that had slides to present know that Freyja doesn't let you slip. If you don't do it, you keep getting reminders, but that is a good thing. It's the reason this meeting happened as well as it did.

I want to make a note about the animal studies group and acknowledge them. These studies are the result of years, literally years of, for a lot of it, weekly calls. Now we don't have them as frequent but these calls involve DOD. I'm not going to mention names because I'll forget somebody and there are too many good people. I don't want to leave them out. But, DOD, FDA, NIH, the contractors, you know, Battelle, and the vaccine manufacturers. Depending on the agenda we would have a mix and match kind of call where we may be government only or it may be just the animal contractor Battelle or it may be the vaccine manufacturer.

It's been an ongoing very active, very responsive group, and I think we wouldn't have this data without this group. Sometimes I think how could this model apply to other counter measures and other product lines. I think it is applicable, but it takes a lot of stars to align. It takes time and commitment and responsiveness and engagement from a lot of people.

Then, of course, there is business interest and you get into sensitivity of data very quickly so that really complicates things. Anyway, I think it has been a good model, and we need to think about how it can work for other products. We will certainly keep it going.

Finally, I want to thank the panelists. Their input and experience has been invaluable. I know you are all busy and it's hard to make these trips. I know you travel a lot and none of us like to travel more than we have to. We really appreciate your coming and participating.

That's all I have. Thank you everybody for coming.

(Whereupon, at 11:54 a.m. the meeting was adjourned.)