UNITED STATES OF AMERICA

+ + + + +

DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

+ + + + +

ANTI-INFECTIVE DRUGS ADVISORY COMMITTEE (AIDAC)

+ + + + +

Meeting

+ + + + +

Tuesday,

February 19, 2002

+ + + + +

The Advisory Committee was called to order at 8:00 a.m., in the Conference Room of the Holiday Inn, Two Montgomery Village Avenue, Gaithersburg, Maryland, by Dr. L. Barth Reller, Chairman, presiding.

PRESENT:

DR. L. BARTH RELLER	Chairman
DR. VINCENT T. ANDRIOLE	IDSA Representative
DR. GORDON L. ARCHER	Member
DR. DAVID M. BELL	Consultant
DR. JOHN E. BENNETT	Consultant
DR. P. JOAN CHESNEY	Consultant
DR. CHRISTY CHUANG-STEIN	PhRMA Representative
DR. ALAN S. CROSS	Member
DR. STEVEN EBERT	Consumer Representative
DR. JOHN E. EDWARDS, JR.	IDSA Representative
DR. ROBERT J. FINK	Consultant
DR. THOMAS R. FLEMING	Consultant
DR. MARY GLODE	Consultant

PRESENT: (CONT.)

DR. RICHARD GORMAN	Consultant
DR. CATHERINE HARDALO	PhRMA Representative
DR. JAMES E. LEGGETT, JR.	Member
DR. CELIA MAXWELL	Consultant
DR. GEORGE H. MCCRACKEN, JR.	Guest
DR. JOSHUA P. METLAY	Guest
DR. ROBERT M. NELSON	Consultant
DR. JUDITH R. O'FALLON	Member
DR. JAN E. PATTERSON	Consultant
DR. JULIO A. RAMIREZ	Member
DR. COLEMAN ROTSTEIN	Guest
DR. DAVID SHLAES	PhRMA Representative
DR. CIRO SUMAYA	Consultant
DR. GEORGE H. TALBOT	IDSA Representative
DR. FRANCIS TALLY	Industry Representative
DR. DENNIS D. WALLACE	IDSA Representative
DR. JANET WITTES	Consultant
DR. LIANNG YUH	PhRMA Representative
DR. TARA P. TURNER	Executive Secretary

C-O-N-T-E-N-T-S

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

(8:11 a.m.)

CHAIRMAN RELLER: Good morning. I'm Barth Reller, in the Division of Infectious Disease, Professor of Medicine and Pathology at Duke University Medical Center, and Director of Clinical Microbiology. I would like to welcome you to this morning's and this afternoon's Anti-Infective Advisory Committee of the U.S. FDA.

We will begin this morning's meeting with a conflict of interest statement read by our Executive Secretary, Tara P. Turner. Before that, however, I would like to introduce or have the other panel members introduce themselves.

We will start at the right and continue around, but in addition to that, there are three members of the Pediatric Subcommittee for Anti-Infective Agents, and after Dr. Glode, if those three members who are not sitted at the table would please come up to a microphone and introduce themselves. We will start with Dr. Goldberger.

DR. GOLDBERGER: I am Mark Goldberger, from the Office of Drug Evaluation IV, FDA.

DR. ALBRECHT: Renata Albrecht, Acting Director, Division of Special Pathogen and Immunologic

Drug Products, FDA.

DR. SORETH: Good morning. I am Janice Soreth, the Division Director for Anti-Infectives at FDA.

DR. LEGGETT: Good morning. Jim Leggett, Infectious Diseases, in Portland, Oregon.

DR. SUMAYA: Ciro Sumaya, Dean, School of World Public Health, Texas A&M University System Health Science Center.

DR. GLODE: Mimi Glode, Pediatric Infectious Disease, University of Colorado Medical Center.

DR. O'FALLON: Judith O'Fallon, Cancer Center Statistics, Mayo Clinic, Rochester, Minnesota.

DR. ARCHER: Gordon Archer, Infectious Diseases, Adult Infectious Diseases, Virginia Commonwealth University, in Richmond, Virginia.

DR. RAMIREZ: Julio Ramirez, Division of Infectious Diseases, University of Louisville, Kentucky.

DR. TURNER: Tara Turner, Executive Secretary for the Committee.

CHAIRMAN RELLER: And could we have the other three members of the Pediatric Subcommittee come up to a microphone and introduce themselves, please.

DR. FINK: Bob Fink, Pediatric Pulmonology, Children's Hospital, in Washington, D.C.

DR. NELSON: Robert Nelson, Pediatric Critical Care, Children's Hospital, Philadelphia.

CHAIRMAN RELLER: Thank you very much, and we look forward to your participation in today's discussions.

DR. EBERT: Steven Ebert, Infectious Diseases Pharmacist, Meriter Hospital, and Clinical Professor, University of Wisconsin, Madison.

DR. BELL: David Bell, Assistant to the Director for Antimicrobial Resistance, National Center for Infectious Diseases, at CDC in Atlanta.

DR. CROSS: Alan Cross, Division of Infectious Diseases, University of Maryland at Baltimore.

DR. PATTERSON: Jan Patterson, Infectious Diseases University of Texas Health Science Center, San Antonio.

DR. CHESNEY: Joan Chesney, Pediatric Infectious Disease, at the University of Tennessee, Health Science Center, in Memphis.

DR. BENNETT: Jack Bennett, NIH, Bethesda, Maryland.

DR. FLEMING: Thomas Fleming, Department

of Biostatistics, University of Washington.

DR. WITTES: Janet Wittes, Statistician, Statistics Collaborative, D.C.

CHAIRMAN RELLER: Thank you. Dr. Turner.

DR. TURNER: Thank you. The Food and Drug Administration has prepared general matters waivers for the following special Government employees: Julio Ramirez, Steven Ebert, John Bennett, Jan Patterson, Celia Maxwell, Ciro Sumaya, L. Barth Reller, Alan Cross, Gordon Archer, James Leggett, Jr., Joan Chesney, Celia Christie-Samuels, Janet Wittes, Robert Fink, Richard Gorman, Thomas Fleming, Robert Nelson, and Kathryn Edwards, who are attending today's Anti-Infective Drugs Advisory Committee Meeting on the proposed approach for selection of delta in non-inferiority equivalence clinical trials.

And the impact of this approach on studies of anti-infective drug products, with a focus on acute exacerbation of chronic bronchitis and hospital-acquired-pneumonia being held by the Center for Drug Evaluation and Research.

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

Unlike issues before a committee in which a particular product is discussed, issues of broader applicability, such as the topic of today's meeting, involve many industrial sponsors and academic institutions.

The committee members have been screened for their financial interests as they may apply to the general topic at hand. However, because general topics impact on so many institutions, it is not prudent to recite all potential conflicts as they apply to each member.

The FDA acknowledges that there may be potential conflicts of interest, but because of the general nature of the discussion before the committee, these potential conflicts re mitigated.

With respect to FDA's invited guests, there are reported interests which we believe should be made public to allow the participants to objectively evaluate their comments.

Dr. George McCracken, Junior., is a researcher with Bristol Myers Squibb and Abbott Laboratories. In addition, he lectures for GlasxoSmithKline and serves as a scientific advisor for GlasxoSmithKline, Abbott, Bristo Myers Squibb, Aventis Pharmaceuticals, Bayer, and Johnson & Johnson.

Dr. Joshua Metlay lectures and is a scientific advisor for Aventis.

Dr. Coleman Rotstein serves as a researcher and has contracts and grant from Pfizer, Merck, ICOS, Schering, Wyeth, and Fujisawa. In addition, Dr. Rotstein consults for Merck, Schering Pfizer, and Pharmacia. He also lectures for Pharmacia, Pfizer, Bayer, Merck, and Fujisawa.

In addition, we would like to note for the record that Drs. Catherine Hardalo, David Shlaes, Lianng Yuh, and Christy Chuang-Stein from PhRMA, Dr. Francis Tally from Cubist Pharmaceuticals, and Drs. Vincent Andriole, George Talbot, Dennis Wallace, Louis Rice, and John Edwards, Jr., from IDSA, are participating in this meeting as industry representatives, acting on behalf of regulated industry.

As such, these participants have not been screened for any conflicts of interest. And I have two announcements. I just want to remind the participants that when you want to speak into the microphone, please pull the microphone towards you, and press the button until the light turns on red. And to be sure to turn it off when you finish speaking.

Also, if you wish to enter a statement for the record, comments on this meeting topic may be submitted to Docket Number 98D-0548, Development of Antimicrobial Drug Products, and there is a handout that has been distributed at the front table. Thank you.

DR. ANDRIOLE: Barth, I have a comment to make about Ms. Turner's introduction of the four of us. We are here to represent the Infectious Diseases Society of America and not any industry.

CHAIRMAN RELLER: Thank you, Dr. Andriole, and actually this is a great segue to asking you and others from IDSA at the invited guest table to introduce themselves. Could you start?

DR. EDWARDS: I am Jack Edwards, from Harbor UCLA Infectious Diseases.

DR. WALLACE: I am Dennis Wallace, and I am from Rho, Incorporated, in Chapel Hill.

DR. TALBOT: George Talbot, Talbot Advisors.

DR. ANDRIOLE: Vince Andriole, Yale University, and a previous member of this august and dye infective advisory committee, and previous Secretary of the Society, and President of the Society.

And a person who was involved in the guideline preparations in 1988 to 1990, and the four of us are here to represent the Infectious Disease Society of America.

CHAIRMAN RELLER: And also at the table on the far left, Dr. McCracken.

DR. MCCRACKEN: George McCracken, University of Texas, Southwestern Medical Center, Pediatric and Infectious Disease, but also a member of IDSA.

CHAIRMAN RELLER: Thank you. Tara. We have the facing table on the far right, and would Dr. Tally begin, and then we will move to his right. Dr. Tally.

DR. TALLY: Thank you, Barth. I am Frank Tally, from Cubist Pharmaseuticals, where I am the Chief Scientific Officer.

DR. SHLAES: I am David Shlaes, and I am here to today representing PhRMA, part of the PhRMA group. I run the infectious disease discovery research group, in the therapeutic area, for Wyeth-Ayerst.

DR. CHUANG-STEIN: I am Christy Chuang-Stein, Statistician, from Pharmacia Corporation, here representing Pharmacia as well.

DR. METLAY: Josh Metlay, from the University of Pennsylvania, from the Departments of Medicine and Epidemiology.

CHAIRMAN RELLER: And lastly, Dr. Fleming. He was here earlier, and Dr. Temple just joined us at the table.

DR. TEMPLE: Bob Temple, Associate Director for medical policy at FDA.

CHAIRMAN RELLER: Thank you. We will begin the presentation with opening comments from Dr. Mark Goldberger, who is the acting director of the Office of Drug Evaluation for the FDA. Mark.

DR. GOLDBERGER: We would like to extend our welcome to the advisory committee members, guests, consultants, and everyone else in the audience who is here attending what has been a reasonably highly anticipated event, I think.

Our goal in having this meeting, which we regard as the start of a process, is ultimately to ensure that we have antimicrobial therapy that is in fact adequate to meet the broad range of therapeutic challenges that we face, challenges that range from routine infections, to very difficult to treat infections illnesses, and of course some of the challenges having only been heightened by some of the

recent events in our country.

To accomplish this, we obviously need to consider approaches to facilitate the development of new antimicrobials, as well as to consider ways to preserve the usefulness of those products that are already available.

We regard this as the beginning of a process. We are having today's meeting, tomorrow's meeting dealing on issues related to the development of antimicrobials for resistant indications.

As Dr. Turner noted, we have established a docket, which I think will be open for the next four months or so to ensure that we get comments and participation from the broadest range of individuals and organizations who are involved or interested in the process of antimicrobial drug development and infectious disease.

We will be presenting some questions for discussion today, but again these questions are really for discussion, and we will not be asking for any formal vote on them, nor do we anticipate reaching any decisions as the results solely of the discussions today.

I think that we certainly recognize the issues that and that it is important to consider the

resources required to perform clinical trials, as well as the types of information that we would like to be able to get from such studies, and at times it appears as though these two things, these two issues, have a certain tension between them.

And that is sort of the subject of much of our discussion and I think some of the questions that we will be asking this afternoon. We certainly believe, and the FDA has long used this approach, that the quantity and strength of evidence should take into account the seriousness of the disease, and the availability of alternative therapy, and again we think that the questions we are posing, as well as the substance of much of the discussion today, will focus on issues like that as well.

And I would like to thank everybody. We are looking forward to a very interesting discussion today.

CHAIRMAN RELLER: Thank you, Mark. Our next speaker will be Dr. Renata Albrecht, who is the Acting Director, Division of Special Pathogen and Immunologic Drug Products as FDA.

And she will speak to the "Historic Perspective, Selection, and Implications of Delta." Dr. Albrecht.

DR. ALBRECHT: Thank you, Dr. Reller, and good morning everyone. I would like to add my words of welcome to Dr. Reller and Members of the Committee, guests, and colleagues.

My task this morning will be to give you a brief historical perspective on the selection of Delta, and to talk about the implications of Delta on clinical trials and patient care. Next slide, please.

Many of you may recall that originally this meeting was scheduled for September 13th of last year, and in fact the meeting had been planned for the better part of the year, but needed to be postponed because of national events on September 11th of 2001.

Discussion of Delta and related issues, however, continued in the intervening 6 months, and resulted in two letters being sent to clinical infectious disease, which have been added to the background material for this talk or for this meeting.

Members of the Office of Drug Evaluation IV had the occasion to have discussions with individuals from academia and representatives from industry, and as a result of these discussions, we have expanded the agenda to include presentations by these groups.

I would like to speak on a few broad

areas. One is the historical perspective on the role that Delta has played in regulatory decision making, and the procedures used to select Delta as outlined in the 1992 points to consider document.

Then I would like to speak about the impact of delta on clinical trials, and finally the consequence that Delta has on patient care. Next slide.

During most of today's presentations there will be detailed discussions on the definition of Delta, as well as the scientific and clinical issues important in the process of selecting Deltas. So I will not cover these in my presentation.

Instead, I wish to address the question of why is Delta important, and what role has Delta played in the regulatory decision process. Next.

In general, the regulatory decision about a particular product for a particular indication has been that; if the Delta of the trial is met, the indication is approved, and if the Delta is not met, the indication is not approved.

There have been rare exceptions to this pattern. For some drugs and indications, the Delta was met, but the indication was not approved due to concerns about the drug safety. And in some examples,

the Delta was not met.

Yet, the indication was approved due to an overall risk benefit evaluation of the product, and in those cases the results of the trials were reflected in the product labeling. Next.

Thus, one may conclude that Delta has been one of multiple important factors considered in making a regulatory decision. Next slide.

So how do we select Delta? The selection of Delta has been guided by the 1992 points to consider document, entitled, "Clinical Development and Labeling of Anti-Infective Drug Products. This document is available on the FDA Guidance Website, and is also included in the background material. Next slide.

The 1992 points to consider document suggested that the 95 percent confidence interval approach may be used, and recommended that Delta be based on the observed success rate. So as shown in the green rectangles to the left, for a 90 percent success rate, the recommended Delta is 10 percent.

For an 80 percent success rate, it is 15 percent. And for 70 percent the Delta is 20 percent. And as seen in the rectangles on the right, the corresponding sample size is 142 patients, 112

patients, and 83 patients per arm, respectively. Next slide.

The points to consider document also stated that the design and conduct of clinical trials was influenced by factors such as incidents of infection, natural history of infection, realistic numbers of patients available for study, cure rates of other; that is, control drugs.

In addition, one has to take into consideration properties of the test drug, such as pharmacokinetic and pharmacodynamic properties; in vitro microbiology data, information from already approved indications, and safety and efficacy data on other drugs within the drug class. Next.

The document also advised that demonstrating effectiveness is one part of the burden of proof, and that a risk benefit profile for the drug must be established.

The document also stated that there are situations where the morbidity and mortality of the illness under evaluation will dictate that an absolute difference in success rates will be clinically unacceptable. Next.

However, over the years the step functions specified in the points to consider document

persisted, while the other elements were lost, much like the body of the Nike of Samothrace remains, while her head does not.

Therefore, the agency held an advisory committee meeting in 1998, during which a draft general statistical guidance was presented, and then in February of 2001, the agency published a disclaimer to the points to consider document, stating that the sliding scale method for determination of Delta was no longer used.

Both of these events will be further discussed by Drs. Lin and Brittain during their presentation. So in 2001 then, the agency started putting together motions and plans for this advisory committee meeting to allow for a public discussion of the selection and determination of Delta as Dr. Goldberger stated in his introductory remarks. Next.

As we hear the presentations on the statistical and clinical issues for selecting Deltas, it is important to keep in mind the impact these decisions will have on clinical trials.

This is the same slide that I showed earlier about the 95 percent confidence interval approach suggested in the 1992 points to consider document.

This approach is familiar to industry, and suggests the sample size of around a hundred to 150 patients per Arm for most clinical trials. However, what if an alternative Delta is selected. Next.

For the sake of illustration, and also in the interest of time, I am going to focus on the impact of selecting Deltas that are the same as or smaller than suggested in the 1992 points to consider document.

So if one were to say that a Delta of 10 percent should be used for all studies, meaning that the test drug could be no more than 10 percent worse compared to the control drug, the same size for the study with a 90 percent success rate remains at a 142 patients per arm.

However, for a drug with an 80 percent success rate, the sample size would double from 112 to 252 patients per Arm; and for a 70 percent success rate, it would increase four-found, from about 83 to approximately 330 patients per Arm. Next slide.

And if one were to take an even more conservative approach and select a Delta of 5 percent for a trial, the sample size would increase four-fold from 142 to 565 patients per Arm, with a success rate of 90 percent.

For an 80 percent success rate, the sample size would go from 112 to 1,005, which is a nine-fold increase; and finally if the study has a success rate of 70 percent, the sample size would increase approximately 16-fold from 83 patients to 1,319 patients per Arm. Next, please.

So the clinical trial implications of Delta are the following. For a given Delta, the lower the success rate, the larger the sample size. And for a given success rate, the smaller the Delta, the larger the sample size. Next slide.

This relationship is nicely illustrated and summarized in this graph, and I would like to thank Drs. Lin and Brittain for making this slide for us. In this graph the X-axis represents the success rate, and the Y-axis the sample size, and the different colored bars represent Deltas.

And as one can see, going from a Delta of 20 percent, the light blue, to 15 percent, green, and 10 percent, a darker blue; and 5 percent, the yellow, and the sample size goes up.

And the same pattern is seen as one goes from a success rate of 80 percent, 70, 60, 50 percent, and the sample size for all of the Deltas do go up. So as we can see from these numbers, the demands on

clinical trials and the impact of these has impact on a variety of groups and stakeholders.

For industry and investigators, there is a time commitment and a cost commitment of doing clinical trials. And the larger the trials, the more time and resources they will take. Clinical trials impact physicians, health care providers, and pharmacists who rely on the availability of information from such studies to guide their knowledge of drugs and use of drugs in patient care.

And clinical trials impact patients. They impact patients as participants in clinical trials. The larger the study, the more patients need to participate. And they impact patients as recipients of drug therapy.

Clinical trials and predefined Deltas determine the extent of information that is available when making these treatment decisions for patients. So in conducting a clinical trial, if one accepts a Delta of 15 percent instead of a Delta of 10 percent as evidence of non-inferiority, the consequence may be that the drug may be potentially 5 percent less effective than a drug that would have been approved with a 10 percent Delta.

And which also means that an extra 5,000

patients may potentially fail therapy for each hundred-thousand patients treated. Next slide.

But things are never one-sided. What if the Delta selected for a trial is small, so small as to be unrealistic, and then no clinical trial is conducted.

Then in fact no clinical data are available to guide patient treatment. And even under the 1992 points to consider approach, some diseases were rarely studied, including endocarditis, osteomyelitis, and meningitis.

So, in summary, the selection of Delta impacts not just clinical trials and all parties involved in clinical trials, but impacts patients who then use the agents approved on the basis of these studies.

Selection of Delta raises a number of issues and questions, and we would like the committee and our guests to provide us with comments on these issues. As Dr. Goldberger said, we are not asking for any votes on any of these topics today.

And in addition as Tara Turner said, we are making available Docket 98D-0548 for those groups and persons who wish to provide us with written comments.

After this meeting, we do plan on reviewing these comments, and plan at least one follow-up advisory committee meeting, and plan to summarize the advice in updated guidance documents. Thank you.

Thank you, Dr. Albrecht. We will next hear from Dr. Robert Temple, who is the Associate Director for Medical Policy, Center for Drug Evaluation and Research, at FDA.

Dr. Temple will speak to us about Active Control Non-Inferiority Studies: Theory, Assay Sensitivity, Choice of Margin. Dr. Temple.

DR. TEMPLE: Well, good morning. It is a pleasure to be here to talk about one of my favorite subjects, which is active control trials and how to interpret them.

We have as an agency been interested in this in a very long time. I have been writing about it since the early '80s, and we have hinted in regulation since 1985 that equivalence trials presents special problems, and have written various guidances for years about how to analyze such trials.

Susan Ellenberg and I wrote an article in the Annals of Internal Medicine in September of 2000 that discusses the theory of all of this. But

probably the most prominent document that we have participated in is an International Conference on Harmonization document called, "E-10, Choice of Control Group and Related Issues in Clinical Trials," that was issued in 1997, I guess.

Just in case anybody doesn't know, a little bit about what the ICH is, because this represented a remarkable degree of international harmony. It is the International Conference on Harmonization.

And three regions -- the U.S., Europe, and Japan, made an effort to harmonize the technical requirements for the marketing of drugs. Not the approval decisions, but the technical requirements, where disharmonies appeared to be unnecessary.

They focused on what they called quality, which means manufacturing control, and safety, which means pharm/tox, and efficacy, which means human efficacy and safety.

And produced a series of mutually agreed upon guidelines. The participants in this organization are the three regulatory authorities and their respective manufacture organizations, such as PhRMA for the U.S.

The organization develops guidance

documents in the three scientific areas, and these are then adopted more or less uniformly in the three regions, and sometimes as guidance in the U.S.

Sometimes you need to change your regulations, and in the final stage, the guidances are controlled by the three regulatory bodies, and not the pharmaceutical organizations.

And as I said, there is no attempt to make the decisions to the same. The ICH E-10 document, which can be found on our website and the other parties, is called "Choice of Control Group and Related Issues in Clinical Trials."

And it is actually a general discussion of all kinds of control groups, including historical controls, which it doesn't like very much. It discusses the ethics of placebos,and a wide range of other matters.

But it devotes particular attention to the use of active control equivalents, sometimes called non-inferiority designs. Not to dehumanize them as has been alleged, and to say that they can't be used, but to describe their logic and their inferential difficulties, and to emphasize the need for evidence of assay sensitivity, which I will describe in a moment.

Much of what follows is considered an ICH E-10, but that document discusses the issue of margin and the distinction between M1 and M2, which is here called Delta-1 and Delta-2. amd we actually tried to call it that in the document, but it was considered too statistical.

Anyway, it discusses that rather minimally, and so this meeting and others like it are an important next step in all of this. When it comes to demonstrating efficacy, there are two quite distinct approaches.

One is to show a difference between two treatments in a randomized trial, or for that matter in an historical controlled trial. That shows the superiority of the test drug to whatever the control is -- placebo, active drug, or a lower dose of the same drug -- and that demonstrates a drug effect if you show such a difference.

The second approach is to show that the new therapy isn't worse or isn't much worse than some of therapy. Showing similarity to a known effective therapy, and that is an inactive control, and attributing the efficacy of the active control to the new drug, and that in-turn demonstrates drug effect.

There is nothing wrong with that logic,

but it poses certain problems, at least in some cases. A non-inferiority trial, which is really what equivalence trials are, shows that the new drug is not worse than the control by some defined amount.

That amount being the margin, M or Delta, and that amount can be no larger than the effect that the active control would have had in the study. If you can't rule out a difference that large, then you have not shown that the new drug has any effect at all.

And I just want to emphasize that I didn't change all of my slides. M and Delta are interchangeable terms. We are not so far from a time when the naive approach in active control trials was in fact used, and in fact one can discover such a naive use in the recent New England Journal of Medicine article, comparing coumadin and aspirin in prevention of stroke.

The idea is that you compare the new and control drug, and if there is no significant difference, then you declare the new and old drugs equivalent, and the new drug is effective.

The problem with that is that increase in variance all by itself -- that is, making the study too small -- will lead to success. And that is now

widely understood. So what is done now is that a non-inferiority study specifies as a null-hypothesis that the new drug is inferior by some margin, M, and tests this statistically.

So if the 95 percent confidence interval upper bound for the degree of inferiority, that is, the control drug minus the test drug, is less than M, then the null hypothesis of inferiority is rejected, and if it were greater than M, then of course or then the hypothesis is not rejected.

If the confidence interval is very wide, because the sample size is too small, the study will not declare non-inferiority. So it solves the size problem. But it doesn't solve what I will describe as the assay sensitivity problem.

Any time you do an equivalence or non-inferiority trial there is a question. Did the active control drug have an effect of the size expected in the trial that you actually carried out.

That may not seem like a pertinent question in many antibiotic settings, but it is in lots of others, most symptomatic treatments. If the active drug didn't have that expected effect, then showing equivalence or non-inferiority by the expected margin -- and that is a typo there, sorry -- by the

expected effect, that's meaningless, because the equivalent or non-inferior drug could have no effect at all, and this study just is one that could not tell anything from anything.

So if no difference greater than the margin is seen, does that mean that both drugs work or that neither drug worked, and you have to know something from outside the study to answer that question.

Assay sensitivity is a property of a clinical trial, and it is the ability of the trial to distinguish effective from ineffective drugs. Assay sensitivity depends on the effect size that you need to detect. A trial may have assay sensitivity for an effect of 10, but not an effect of five.

So you really need to know what the effect of the control drug was in that study, and of course, you are not measuring it in an equivalence trial, and so you have to learn it from historical information.

So there is an unstated assumption in any non-inferiority trial, which is actually nowadays it is stated, but it used to not be stated, that the active control was effective in the particular study.

That is, that the trial had assay sensitivity, and that is not necessarily true for all

effective drugs. It is not testable in the data collected because there isn't any placebo group.

And it gives an active control study some elements of a historically controlled study. Again, I know that I am repeating myself, but superiority equals efficacy as long as the control is better than placebo, which is usually safe to assume.

And non-inferiority doesn't equal efficacy unless assay sensitivity is present. Assay sensitivity has to be deduced or assumed based on historical experience showing sensitivity to drug effects, and that means that it is usually possible to distinguish the control drug from placebo.

And then you have to do the study in a way that doesn't mess it up. If, for example, nobody took the drug, then even an effective drug would not be effective in a given trial.

And it is important to make the new trial as similar as possible with respect to patient population and end-points as the trials in which the active control was effective.

This is just an advert from three-arm trials, where there is both an active control and a placebo, which is nice if you can do it. So what one component of deciding that a drug -- that a control

drug -- that a study has assay sensitivity, is to look for historical evidence of sensitivity to drug effects.

That means that well-designed trials pretty regularly can distinguish the active drug from placebo. Sensitivity of drug effects is an abstract conclusion, and assay sensitivity is a conclusion about a particular trial that takes historical evidence of sensitivity to drug effects, and adds to it a proper study quality.

Now, many people don't appreciate this. When you raise the issue of assay sensitivity, and say, well, not every drug is effective against placebo every time, and the next question is, well, why did you approve a drug that bad.

And the answer is that is the best that we can do. There is some settings in which it is not easy to distinguish drug from placebo. Some of these situations are very well understood, and antihistamines are very hard to show a difference between drug and placebo, because the pollen blows away, and then you can't see anything.

And we know that studies of antidepressants, even the effective antidepressants that we all know and love, fail a significant fraction

of the time.

We have looked over several years, and in about almost 50 percent of well done trials, or apparently well-done trials, and they are as done as near we can tell, of effective antidepressants, can't tell drug from placebo.

And no one yet knows how to choose a population sample size or design that would alter that state and everybody would like to, because failed trials are a burden for everyone.

And just a list of situations in which studies of current drugs cannot be assumed to have sensitivity to drug effects include depression, anxiety, dementia, symptomatic congestive heart failure, seasonal allergies, GERD, which is the devil to study. Systematic GERD, I mean.

It is post-infarction beta blockade, and only about post-infraction aspirin, and only about 5 out of 35 studies have actually shown a benefit. Post-infarction aspirin, only occasional studies show an effect on survival, and the largest study ever leaned the wrong way.

That doesn't mean that the drugs that are approved aren't effective for these conditions. It means that you have a problem if you are going to do

an active control trial, because you can't be sure that the drug would have an effect in your particular study.

It is always worth remembering that even if sensitivity to drug effects does exist for a therapeutic class assay sensitivity in a particular study, can be undermined by a variety of study, conduct factors, that give you a bias towards the null.

That is, obscure true differences between treatments just to illustrate these. Poor compliance, and nobody takes the drug, and the drug can't tell drug from placebo.

Too many cross-overs, and a population for one reason or another improves very rapidly, and spontaneously. On the other hand, a population that is very resistant, and too much use of concomitant medication that treats everybody independent of the drugs so that you can't see a drug effect anymore.

Poor diagnostic criteria. You put the wrong people into the trial. Insensitive measures of a drug effect, and poor quality of measurements, mixing up the treatment. All of these things don't necessarily affect variance very much, but they might affect the treatment size.

It is also worth remembering that what you think you know about historical evidence really applies only to trials of a particular design, and different trials may or may not have that property. Changes in these can effect the size of the active control effect, and therefore one's choice of margin, or in fact even completely undermine assay sensitivity.

So the non-inferiority margin, or delta, and completely equivalent terms, is the degree of inferiority of test drug to control drug that the trial is going to exclude statistically.

In other words, if you take the 95 percent confidence interval for the difference between control and test drug, it has to be less than that margin, whatever that margin or delta is.

Obviously the margin can't be any larger than the effect the control drug would be reliably expected to have. And we will call that M-1 or Delta-1, and if M-1 is the entire effect, the control drug can be presumed to have in the study.

And if C minus T is greater than M-1, then the new drug has no effect at all. And it is always worth remembering that the choice of margin is very critical for everybody, including regulatory agencies.

And if you allege that the control drug has an effect on M-1, and find that the control minus test drug is less than M-1, then the test drug is effective, which is what we want.

But if in this trial you are wrong, and it really only had an effect size of half of M-1, then the test drug will not really have been shown to be effective, and it will only look at that way, and that is why we worry.

The margin used in a trial could be the entire effect of the control drug for many symptomatic conditions. We are content to know that the drug has any effect. But the margin chose could be smaller, and we have been calling that M-2, or delta-2, if there were a clinical need to assure preservation of more than just some of the control drug effect.

That is, preservation of some fraction of the effect of the control drug, or some absolute benefit. Choosing an M-2 smaller than the whole effect of the control may be important when the effect is clinically critical. For example, mortality.

It might then be 50 percent, which would be 25 percent, and you wouldn't want to lose more than 25 percent of the effect of the control agent, or as you will see, sometimes even less.

Just to illustrate these, these are five examples. What you see on the left axis is the difference in effect of what I have been calling C minus T, and there are five examples, 1 through 5 across the top.

And the top dotted line is M-1, and that is the whole effect of the control drug, and M-2 is some smaller effect, because you want to preserve more than just any effect. And M zero is the line of equivalence.

In example number one -- and this is the point estimate, plus a confidence interval for the difference between C minus T. The drugs look about the same, and the confidence interval is narrow, and so you have shown that the effect is at least M-2.

And if that is what you were trying to do, you are happy. In Number 2, the point estimate is somewhat adverse to the new drug, and the confidence interval includes a value larger than M-2. So you have not ruled out loss as drawn here, say 50 percent of the effect, although it does look as if it has some effect.

In M-3, the point estimate is adverse to the new drug, and now the confidence interval includes a value that is even worse than the whole effect of

the drug. So in this case, you haven't really shown that the drug does anything.

Example Number 4 shows superiority to the control drug and that is always good. And in the fifth example, it shows a point estimate that is favorable, but the study is too small or something else is wrong with it, so that you haven't excluded a loss of all of the effect of the drug.

Just briefly, this will be a point discussed later. In the past, and actually some of the original descriptions of non-inferiority studies, the margin was chosen clinically.

That is, you decide how much difference you were willing to accept, and you rule that out. Where the effect of the drug is very large, which is certainly the case in many antibiotic settings, and certain highly responsive tumors, that is okay.

You don't have to worry about losing all of the effect of the drug, because it is very easy to tell the difference between an effective drug and an ineffective drug.

If you are looking at urinary tract infections, you don't have to worry about whether your effect side is 10 percent or 20 percent. You would be able to tell an ineffective drug from an effective

drug.

So the only thing you are really interested in is how much of that effect you are willing to lose. That is, M-2 becomes the matter of interest. In oncology, for many years, we considered assurance that you hadn't lost more than 20 percent of the survival of the population, an acceptable evidence of effectiveness.

The trouble with that was that the drugs that were being used as the control drugs didn't have an effective survival that large. So that what we were ruling out in many cases didn't rule out the possibility that the drug had no effect at all.

Anyway, that is going to be an important discussion later. So I won't dwell on it now, except with this one slide. In many situations, the effect is very large, and there isn't really a problem in knowing what the historical -- in knowing that a trial has assay sensitivity.

If acute lymphocytic leukemia has a complete response rate of 80 or 90 percent, you don't have to worry about ruling out a difference of 50 percent of that, or 60 percent. You are going to worry about how much clinical difference you are willing to accept, and so you are going to worry

mostly about M-2 or Delta-2.

Similarly, for testicular cancer, acute response to bronchodilators, anesthetic effects, and even in the case of thrombolytics, a look at the available data shows that it is fairly easy to tell whether an active drug is -- to be sure that a drug is active in a particular study.

But you know more about this than I do, but that that would be equally true for urinary tract infections, meningitis, and lots of other situations. One of the things you will talk about is how much effect needs to be retained in situations where the effect size is large.

And of course it is worth remembering that the very reason that you can't do a placebo control trial is the reason for assuring that you are preserving a good part of the effect of the control agent.

So, for thrombolytics, we have said that you need to show that you are not -- that you have not lost 50 percent of the effect, and in certain cancer drugs, we have asked for retention of 50 percent of the survival effect, where that is a matter of a few months.

In adjuvant breast cancer, however, we

have asked that you preserve at least 75 percent of the effect because one does not want to lose more than 25 percent.

This is in some sense a practical question, and one doesn't actually want to lose any of the effect of the control when it has an important effect, but sample size has become rapidly out of sight when you try to do better.

Thrombolytic trials show that you preserve 50 percent of the effect in 14,000 people, if you wanted to preserve 75 percent of the effect, you would get into the 70,000 range.

Again, as you will hear, there are at least a few situations where the effects of active agents is not so large, hard to discern, and hard to demonstrate. And when that's true, then a non-inferiority design becomes a problem, and one does have to think both about M-1 and M-2, and it may be very difficult to use a non-inferiority design, and therefore placebo controls need to be considered.

One question is whether those will be ethical. So a brief word about ethics, which ICH E-10 considers at some length. That document clearly distinguishes between available drugs that prevent serious harm and those that treat symptoms.

As a general matter, where an available treatment is known to prevent serious harm, death, or irreversible morbidity in the study population, you really can't use a placebo control.

The only generally is a hedge because sometimes the drug is so toxic that people will reject it anyway. Where there is no serious harm, however, it is generally considered ethical to ask patients to participate in a placebo control trial even if they may be uncomfortable, provided the setting is non-cohesive, and that patients are fully informed about available therapies.

Of course, it is also true that whether a particular placebo control trial will be acceptable to patients and investigators is a matter of investigator, patient, and IRB judgments. So it might be ethical, but it might be that no one would be in it.

One question again, and this is just the briefest introduction, but it may be possible to design trials where it is impossible or difficult to specify M-1 that randomize patients to drug and placebo, and preferably with an active control as well, and that allow early escape for any one not doing well.

For example, failing to respond by time-X or something like that. Again, you will hear a great deal more about all of this. Thank you.

CHAIRMAN RELLER: Thank you, Dr. Temple. We will next hear from the Statistical Team Leader, Dr. Daphne Lin, and Dr. Erica Brittain, the Statistical Reviewer for FDA on Statistica Issues in Specification of Delta. Dr. Lin.

DR. LIN: Thank you, Dr. Reller. Good morning. This is a joint work with Dr. Erica Brittain. We are going to present statistical issues in specification of delta.

I am going to give the first part of the talk, and later Dr. Brittain will cover the second part. The outline of our talk. First, briefly, an introduction to non-inferiority trials, and non-authority margin; that is, delta.

Later I will give a brief introduction, a brief history, of the reaction in FDA's anti-infective drug product area. Later, Dr. Brittain will talk about the principles for determining Delta, and difficulties in practice, and alternative design, and finally a summary will be made.

If there is a new drug, how can we show the new drug has identical efficacy to the standard of

the drug? A short answer is that we can't. And the alternative availability in the clinical trial, statistically, we cannot prove the effect of treatments.

So what can we do? The short answer is that we must allow for some potential difference in efficacy, and that is Delta, the topic of today's talk.

So what is delta? ICH E-9 has a definition of delta, which is that it is the largest clinically acceptable difference, and it should be smaller than differences observed in superiority trials of active comparator.

Or Delta can be described as the largest acceptable line in efficacy between tests and the active counter drug. For example, if we tried to design a meningitis trial, then what is the largest clinically acceptable difference between tests and the active counter drug?

We can design a non-inferiority trial to answer the previous question. A non-inferiority trial is designed to ensure that the new drug is not worse than the standard drug by some margin delta.

In the anti-infective drug product area, in general, what defines treatment effect as the

absolute difference is the absolute difference of percent cure rates.

For example, if an observed success rate in control is 85 percent, and the observed success rate of test drug is 75 percent, then the point estimate of difference is 10 percentage points.

An in general a confidence interval around this estimate of treatment effect is used as the primary analysis for non-inferiority trials. So what is a confidence interval?

The 95 percent confidence interval for the difference in success rate between two drugs means we are 95 percent confident of that. Now the true difference in efficacy between these two drugs is contained in the confidence interval.

Next, let me give you two examples to illustrate and how to use the 95 percent confidence interval to interpret the result from a non-inferiority trials.

The first example is if a trial of two hundred patients per Arm, designed with a delta of 10 percent, and if the trial results shows the success rate of the test drug is 88 percent. Control drug, 90 percent, and if the point estimate of the difference is minus 2 percent, and the 95 percent confidence

interval along this point estimate is between minus 8.6 and the 4.6 percent, and in this example, since the 95 percent lower limit is no less than 10 percent -- I'm sorry.

So in this example, which can concur the test drug is non-inferior to the contour. The second example is similar design with a trial of 200 patients per Arm, with 12 percent.

However, in this example, the trial results show the success rate of test drug is 84 percent and control drug 90 percent, and the point estimate of the difference is minus 6 percent.

And the 95 percent confidence interval falls between minus 13 and the 1.1 percent. And in this example, since -- I'm sorry, I just don't know how to operate this.

So in this example, 95 percent lower limit is less than 10 percent, and so we concur that non-inferiority is not demonstrated. From these examples, we can see that the decision of non-inferiority depends not only on the success rate of test and control drugs also depends on how Delta is chosen.

There are two objectives in non-inferiority trials, and the first objective is that non-inferiority indirectly determine if the test drug

is better than placebo.

And it directly determines if the test drug is similar to the active control drug. So we need to choose delta appropriately to achieve both objectives.

Next, the history of the history of a selection in FDA's Division of Anti-Infective Drug Products area. As Dr. Albrecht mentioned in her talk, in 1992, points to consider in her document used the staff step function approach.

This slide shows the relationship between Delta and the success rate described in the points to consider document. Choice of delta only depends on the success rate. If the success rate is greater or equal to 90 percent, delta is 10 percent.

If the success rate is in the 80 percent range, delta is 15 percent. if the success rate is in the 70 percent range, delta is 20 percent. Since this is a step function which can lead to problems of interpretation, and if a few outcomes are changed, then a different standard will be used for evaluation.

For example, if the success rates is changed from 80 percent to 79 percent, and delta will be changed to 15 percent to 20 percent. Since delta in points to consider has been chosen primarily based

on success rate, it did not take into account the seriousness of disease and the consequence of treatment value.

And whether delta was small enough that a drug with no efficacy could meet the standard was not considered.

In addition, as I described previously, this step function approach has undesirable statistical properties. Another concern. If the active control arm and the delta are not appropriately chosen, then the so-called "Bio-Creep phenomena may happen.

And that is that if trials over time used progressively less effective control arms, and the delta is not appropriately chosen, then they are already in attenuation of efficacy.

For example, if Drug 1, with a success rate of 70 percent, is used as an active comparator to compare with the new test drug Number 2, with a success rate of 60 percent, and if a delta of 20 percent is used, then in this case, Drug 2 is not inferior to Drug 1.

And if later on there is another test drug, Test Drug Number 3, and if Drug Number 2 is used as an active comparator, and if a delta of 20 percent

is still being used, then we might approve a drug with a success rate of 48 percent, which is much lower than the success rate of Drug Number 1.

Another case, and the worst case scenario, how about if the placebo rate is here, and that is that we might have another drug which is not much different from the placebo.

In July of 1998, on the advice of the committee, we have discussed that, and the choice of delta should reflect many important clinical factors, such as historical cure rate with and without therapy, risk associated with treatment failure, and advantages and disadvantages of study drug.

In addition, in '98, on the advice of the committee, we also proposed that when delta is chosen for simple size computation, it should be clinically relevant, and since delta will be picked based on clinical issues, it may need to be indication specific, and they are some special situations for individual indications when delta may need to be chosen on a case by case basis.

In addition, we also encourage sponsor to discuss the choice of delta with the Medical Division during protocol development. And a sponsor should provide the rationale for selection of control arm.

The CPMP, counterpart of the FDA, published a guidance on the evaluation of anti-bacterial medicinal products in 1997. And this guidance recommended a delta of 10 percentage points for common non-serious infections.

But it needed to be smaller for very high cure rates. Also, this guidance recommended the choice of delta should be based on the clinical judgment, and it is based on a minimum clinically relevant difference, and should be justified in the protocol.

For the past two years, we have worked with sponsors on a case-by-case basis to specify delta. In February of last year, a disclaimer was added to the points to consider document, stated that the step function approach has been phased out, and the choice of delta should follow the ICH E-10 principles, and there is a need to establish standards. This is the end of my talk.

Next, Dr. Brittain will talk about a general principle for selection of delta. Thank you for your attention.

DR BRITTAIN: Okay. So, now what? Here is a road map for the rest of the talk. I am going to be talking about principles for determining delta, and

these are going to be based on the ICH E-10 principles, and then the very real difficulties in practice.

This is the hard part; how you apply this in practice, and this is where we need your advice. Then I will mention alternate designs, a summary, and I also want to say that one of my main goals here is to get across the idea that the choice of delta is not a technical matter, but actually one that potentially impacts patients.

Again, to demonstrate efficacy, the experimental drug needs to be better than placebo, and in some settings, it should have similar efficacy to the existing therapy, and so we want to choose a delta to assure that both of these goals are met.

Here is an important quote from the E-10. This design, "is appropriate and reliable only when the historical estimate of the drug effect size can be well supported by reference to results of previous studies of the control drug."

So what does this mean? We must know with good precision the magnitude of the advantage of the active control drug over placebo in the setting of the clinical trial.

Now, in practice, as Dr. Temple was

talking about, if the advantage is very large, the precision of this estimate probably won't matter. On the other hand, if it is potentially modest, the precision is critical.

And the sort of unfortunate corollary of this is the active control that is based on a single trial with borderline efficacy, we are going to have poor information about the magnitude to support a non-inferiority trial.

So here is some important principles from the E-10. First, a delta could based on both statistical reasoning and clinical judgment; and, second, it cannot be larger than the advantage of the "active drug would be reliably expected to have compared with placebo in the setting of the planned trial."

And it goes on to say that we usually choose delta to be even smaller to ensure that some clinically acceptable treatment benefit is maintained. This is a very artificial example, but I hope that it will convey some important concepts.

Say we actually knew the true success rate of the placebo was 70 percent, and the true success rate of the active control was 85 percent. So the difference between 85 and 70 is 15. So that is the

advantage of the active control over placebo.

One could choose a delta of 15 percent, but you could not use a delta larger than 15, because a drug that has no efficacy has too high a chance of being successful.

And then you might say, well, I don't want to have a drug that is down near the placebo rate. I would like to keep it up closer to that 85 percent rate. So maybe you would want to preserve half the benefit and have a delta of 7 percent.

And then somebody else might think, well, in a particular situation we don't want to lose much of the benefit of the active control, and then you would want a delta of 3 percent.

The main point here is that you can't go bigger than 15, and there might be -- there are all sorts of infinite choices of delta smaller than 15, depending on the objective.

And we have been using this approach to delta as a two-step process and have found this way of looking at it very useful.

We first determine a conservative estimate of the advantage of active control over placebo, the delta one; and this is data based. And then we select the largest clinically acceptable difference between

the active control and the experimental drug, and we call that delta two, and that is judgment based.

And then the smaller of these two values would be the delta that we would use in the non-inferiority trial. So what is this benefit of active control over placebo.

You could define that as a true success rate of the active control, minus the true success rate of the placebo in the setting of the clinical trial.

In other words, by how much is the active control better than placebo in the non-inferiority trial setting if the placebo were actually present. And again I want to emphasize that this is based on historical data.

And it is not a judgment. It is not a choice. At some level there is a right answer. We may just have trouble finding out what it is. And again it is not that critical to get it just right if the benefit is very large.

So why did I say conservative estimate? Well, E-10 says delta, quote, should reflect uncertainties in evidence on which the choice is based, and should be suitably conservative.

The problem is that if the delta is

overestimated, the chance of concluding efficacy when the new drug is no better than placebo is too high.

So if we are going to err at all, we want to err on the side of underestimating the benefit. So what this means is that we have poor historical information.

We are not going to use our best guess of the estimate. We want to use some smallest of the reasonable values. I know that I am being very vague here, partly because even in the statistical community there isn't agreement about exactly how to do that.

Okay. So what is the best information for estimating the benefit of the active control, the delta one? The best case would be if you had a whole bunch of placebo control trials, with exactly the same design that you want to use in the non-inferiority design.

We just -- I don't think there is any situation that in anti-infectives that meets that situation. Sort of halfway down this list would be if you have multiple placebo control trials, but not with the same design that you would want to use in the non-inferiority trial, and maybe not with the same design that the others have used.

And then at the bottom would be the

observational data, and antidotal data, and this obviously is not the best situation, but again if we are talking about large treatment effects, it is probably fine.

But the case that in a way is the most interesting case for anti-infectives, what if we have some placebo control data in the literature, but there is some problems with it.

The trials are old, and so antibiotic resistance that is taking place in the meantime changes in clinical care management may mean that the values in the old trials aren't that valid or relevant.

The proposed active controls may not be studied because these trials were old, and there may not be very many of them, and so we would not know if the treatment effect is consistent.

And very importantly, there are probably differences in entry criteria, assessment criteria, the timing of the assessments, and the populations. So as wonderful as these data are compared to having no information, we have to take the data with a big grain of salt.

So how do we then come up with an estimate of this delta one with this situation, and we don't

know. We are hoping that you can give us some advice. So the bottom line for estimating delta one is we want to use historic data, preferably from placebo control trials with similar designs as possible as the upcoming non-inferiority trial.

The bad news is that in anti-infectives, your historic data is often poor, and maybe poor for good reason because of ethical constraints in doing placebo control trials.

But the fact that the data is not there makes it hard for us to come up with this conservative estimate. And again the good news is the precision of this is probably irrelevant for those indications where the benefit is known to be very large.

So again let me take you back. This was a two step process, and we are just talking about step one, the determination of the estimate of the advantage of active control over placebo, delta one.

And the second step is the acceptable loss from active control delta two, and delta is the smaller of these two components. Now, the selection of delta two is going to be the primary concern for the majority of anti-infective indications probably.

I want to emphasize that unlike the delta-one, which really is pretty much a statistical

decision, this delta one, because of the clinically acceptable loss, is not. It is really a clinical judgment of the largest acceptable difference between active control and the new drug.

It is a difference that is such that it is so important clinically that it must be ruled out, or you could think of it as a borderline value between just barely acceptable and not acceptable.

So what is important to think about? Certainly the consequential treatment failure. If most err study failures are deaths or very serious morbidity, you would probably want to use a smaller delta two.

If treatment failure can be easily reversed or addressed, we could be more lenient. And then this is an important way to look at it. It is kind of obvious, but if in fact the true loss in efficacy of the new drug from the active control drug were say five percent, if a hundred-thousand patients used the new drug instead of the active control, 5,000 extra patients would have failures than if they had used the active control drug and so on.

And if the true loss is 10 percent, then there would be 10,000 extra patient failures. You could kind of go down the right side and say what is

the worst case scenario that we could accept, and then see what delta would correspond to that.

Then there is another issue to think bout with the clinically acceptable loss, and it is a little more subtle, and I kind of call it clinical trial reality.

It is clinical trials that measure the abstract concept that we might be thinking about in our minds. For example -- and this would be one example of a clinical trial reality. And for those indications where there are going to be patients in the studies who do not have disease, and where the indications are hard to diagnose the disease exactly.

Say in a case where the treatment difference among patients with a bacterial infection were 12 percent, and a case with patients without a bacterial infection is zero. So if you had a 50-50 mix in your trial, the treatment difference that you should be measuring would be six percent.

So if you had selected a delta of 10 percent, you may end up concluding the new drug is sufficiently efficacious. But notice that in the key population the patients with the bacterial infections, the treatment difference was actually greater than 10 percent.

So we have to think about -- we can't just think about the clinically acceptable loss in an abstract way. We need to know about how or what you are actually measuring in the clinical trial. And there are other factors that can dilute treatment effects as well.

So, in summary, for the selection of the clinically acceptable loss, certainly the consequence of treatment failures is primary in this consideration. And then the potentially large impact on patient care.

And then we have to be careful about these clinical trial realities, and again I want to emphasize that unlike the delta one, this component, the clinical judgment is really the primary judgment.

Now, for a long time we have been thinking about selecting for each indication its own delta, and this would provide regulatory consistency, but we want to acknowledge that even once we have finally decided what the delta should be for each indication, we are not going to be done, because we are going to have to stay vigilant because we could have the bio-creep problem that Dr. Lin mentioned.

And that if we could keep changing the active control, and that the delta may not be small

enough. And then emerging resistance on other temporal changes can diminish the efficacy of any active control.

So we are going to have to stay on top of this unfortunately. You are going to hear a lot today about consequence to sample size. When you assume that the cure rates are the same in the active control and the new drug, when you cut the delta in half, your sample size quadruples.

One other important thing to mention though is that is the new drug, if it is reasonable to assume that the new drug is slightly better than the active control, the sample size can be sharply reduced.

For example, in this particular case, say you are using 80 percent power and you were using a delta of 10 percent. If you assumed that both cure rates were 80 percent, you would have about 250 in a group.

But if you assumed that the new drug cure was just a little bit better at 82 percent, your sample size would be cut by one-third. So what is the biggest challenges? And we have plenty of challenges for you.

The biggest challenges are indications

where the treatment effect is potentially modest, but not precisely known, and on sort of the flip side, serious indications where we may be pretty comfortable that there is a large treatment effect, but there is low incidence, and so it is hard to do the kind of size studies that we might want to do.

Now, superiority designs may offer an important alternative to the non-inferiority design, particularly in the first case. They can provide stronger evidence, and it in some situations with smaller sample size.

So the question is can they be done ethically. The early escape approach that Dr. Temple mentioned is something that we have been thinking about for quite some time, and I know that it was discussed in the previous advisory committee on a titus media, and a few people brought this up as a possible situation for a titus media.

But the question is whether it is ethical, and this is applicable probably only to a handful of our indications, the less serious ones, or potentially applicable.

But these are big indications in terms of numbers of millions of prescriptions a year. So these are important indications. The two arms, experimental

versus placebo, the key element is that patients are seen several days after baseline, and at that time if a blind assessment shows no improvement, the patient is considered a failure in the analysis, and then the therapy is switched.

Now, this is ethically consistent with the way and see practice of medicine. So if you are comfortable with wait and see, you can be comfortable with this.

A variant of this would be an early escape with three arms, where you would add the active control arm, and obviously that would be the most informative design.

I just wanted to mention other superiority designs. I just want to encourage people to consider superiority designs, even though the non-inferiority design has been the mainstay in this area for so long, we think it would be important to you to open to considering other designs.

One design could be like the placebo add-on design, where the existing drug -- one arm is the existing drug, plus the new drug, versus the existing drug, plus placebo, which answers the question does the new drug have benefit in the presence of the existing therapy.

And a question would be labeling implications with that design. But the dose response design versus low dose situation, superiority to some comparator, or perhaps some combination of these.

Okay. I want to move back to summarize the selection of delta, the big picture. Again, choice of delta impacts patients. If delta is incorrectly chosen so that it is greater than the advantage of active control over placebo, patients may end up getting drugs with no benefit, while being exposed to toxicity, and there is potential for development of resistance.

And even in those situations where we are not so concerned about the placebo rate, there is still potential benefits of using smaller deltas. Potentially, more patients are cured overall and there are higher survival rates, and subtle, but important, differences are detected that might not be detected with bigger deltas.

Of course, other consequences of this, of the smaller delta, would be larger and longer studies which may impact drug development, as of course we will be hearing more about today.

And as a final slide here, as an absolute, delta must be smaller than the conservative estimate

of the advantage of the active comparator over placebo,a nd the challenge here is that we really do not have very good historic data to know what that advantage is.

And so we really need your advice about how to handle that, and then using clinical judgment, we may want to increase delta further to rule out important loss in efficacy.

And again we need your advice in determining what is an important loss in efficacy. And finally that superiority designs can play an important role in some settings. Thank you.

CHAIRMAN RELLER: Thank you, Dr. Brittain,and Dr. Lin, and to the other speakers this morning for their insightful presentations. Are there any questions from the committee specifically on the material presented thus far? Yes?

DR. FINK: I guess my question is that in terms of the issue of bio-creep, which I think is an important one, could a propagation of errors analysis be applied to this data if one could define an initial gold standard?

Propagation of error analysis is commonly used in more defined settings, such as manufacturing or in physical chemistry, but it doesn't seem like it

would be impossible to apply it potentially to biologic systems.

CHAIRMAN RELLER: Thank you, Dr. Fink. Drs. Lin or Brittain? Dr. Albrecht, any comment?

DR. ALBRECHT: In reviewing and approving of new drug products, we don't actually have gold standards that would apply in this case.

CHAIRMAN RELLER: Dr. Temple.

DR. TEMPLE: In a lot of situations, what you are looking at is hazards ratios where you are very worried that you don't know what the actual rate of the untreated condition would be.

It seems to me, but I don't really know the field very well, that in antibiotic treatment that you might set a minimum response rate that would apply to whether you count the study at all.

If you were dealing with urinary tract infections, for example, and you had a 60 percent response rate, you might say, oh, well, that is not typical, and you would throw it out, and it just would be a null study, and you would insist that it be 80 or 85, or whatever you are familiar with.

That might prevent bio-creep to a degree. I don't know how that relates to propagation of errors.

CHAIRMAN RELLER: Dr. Bennett.

DR. BENNETT: I wonder if I could ask Dr. Brittain to clarify something, and that is the early escape with three arms that she alluded to. One arm would be the control drug, the active control, and the other new drug, and do I assume the third arm would be a placebo, because if you have got an early escape clause, you wouldn't want to then go to placebo would you?

DR. BRITTAIN: This is the early escape placebo design, and what I was saying in the two arm study is that it is the new drug versus placebo. The three arm version of that would be new placebo and an active control.

And the idea being again that after maybe two days after base line, patients are determined to see whether they have improved or not. And if they are not improved, they would be put on other therapy.

In other words, no one could stay on a drug that wasn't working for them for more than two days.

DR. TEMPLE: You have to introduce a time element into those kinds of studies. It isn't total response rate, because everybody is going to respond before you are done. It is how many responded three

days or five days, or whatever, or time to response, or something like that.

CHAIRMAN RELLER: Dr. Leggett.

DR. LEGGETT: Just a historical question and a couple of things. Have we actually seen evidence of bio-creep, and have we -- and by we I mean you or the society, or the Europeans, or the Japanese, have we actually seen cases where the step function has resulted in retrospective analysis of saying, oh, I wish we hadn't done that, or is this all still hypothetical/theoretical?

CHAIRMAN RELLER: Dr. Brittain and Dr. Goldberger.

DR. BRITTAIN: I just want to add one comment. I think the worst case of bio-creep is when you can't see it, when you don't know that it is there, and that is the most insidious form of it.

CHAIRMAN RELLER: In listening to this morning's presentation, the language is remarkably similar to some of the dilemmas faced in the practice of evidence-based medicine, evidence based on regulatory process.

And the best available evidence, which may not be ideal, and then plus experience, and then after the break we will hear the experience component from

industry, and infectious disease practitioners, to blend these together to try to come to a full and complete discussion with all perspectives presented.

So that the agency and other interested groups over time can come to a reasonable approach, though not necessarily a perfect one, with a continuing evolution of the evidence on which these decisions can be based.

Dr. Soreth, you had a comment before we take our 15 minute break?

DR. SORETH: To answer further Dr. Leggett's question about whether or not we have evidence, hard evidence of bio-creek. I think there is one approval that we took a number of years ago that illustrates this.

It was a drug, Monul, used as a single dose for the treatment of cystitis in women, and there were three trials submitted in that package. Two, which compared the use of that drug, with 7 days of ciprofloxacin and 10 days of bactrim, in which the drug proved itself to be inferior to those treatments.

And a third trial in which Macrodatin or Nitrofurantoin was chosen as the comparator, and which equivalence was shown. The product label gives the results of those clinical trials, and so hopefully

one, a prescriber would understand where it fits in the spectrum of treatments for urinary tract infections.

But I think that could be -- that is an illustration of having a drug on the market that is inferior to other treatments, and equivalent to another.

CHAIRMAN RELLER: We will reconvene at 9:50. Thank you.

(Whereupon, at 9:40 a.m., a recess was taken and the meeting was resumed at 10:02 a.m.)

CHAIRMAN RELLER: We will begin the second half of this morning's presentations with a presentation on the Medical Perspective: Bacterial Meningitis, by Dr. George McCracken.

DR. MCCRACKEN: Dr. Reller, Committee Members, Ladies and Gentlemen, the title of my presentation is evaluation of antibiotic treatment of bacterial meningitis, an increasing challenge.

At the outset, I want to repeat that what was made, the comment that was made originally at the outset of the meeting that the reason for presentation -- and you can see that I am going to touch briefly on fluoroquinolone, and there is a protocol in front of the FDA for gatifloxacin therapy in meningitis.

I hope to be the principal investigator if it is approved, and thus have potential or conflict of interest with regard to that, and I am an advisor to Bristol-Meyers Squibb, and several other companies that were mentioned to help develop drugs.

I would take some issue with the comment that I speak for companies. I speak for no company. The companies provide money to institutions where I speak, but there is a difference in how that is said.

So fluoroquinolones are coming to pediatrics, whether we like it or not, and I have some reservations, but for some conditions it is critical, and bacterial meningitis is one of those.

So why fluoroquinolone therapy for bacterial meningitis? Well, increasing resistance of pneumococci is a problem worldwide and these drugs are active, at least the newer generation compounds are.

They have expanded coverage against many of the meningeal pathogens, including coliforms, and it can be used in a simplified regimen of a step-down from IV to oral in some settings, in which this would be feasible.

And it certainly penetrates well and has superior or at least comparable bactericidal activity in spinal fluid. Next slide.

Now, how do we study a drug for bacterial meningitis? The first step is in a rabbit model of meningitis, which has been used for many, many years, for more than 25 years, and we are able to apply the pharmacogenetic and pharmacodynamic principle of relevance, which for the fluoroquinolones is area under the curve, and over the MBC, and not MIC, but MBC.

We want cidal activity, and we apply this to spinal fluid, and we adjust the regimen in order to achieve a dosage that has concentrations in plasma or serum that are comparable to those in adults, and the actual amount given to the animal is irrelevant to what we use in humans.

It is only to achieve that concentration, and then we think the regimen in order to achieve the AUC over MBC, and that would be optimal. Now, we can pretty much predict what that would be when you look at dosing intervals, and half-life those, and then we can predict from that what the dosage will be in humans, in infants and children.

So I am going to show you now the next step in which we looked at one drug, which was trovafloxacin just recently published in the January of the Pediatric and Infectious Disease Journal, in

which we evaluated trovafloxacin, and compared to the comparator, which was ceftriaxone, with or without vancomycin.

The dosages was exactly what was predicted from the animal models. Now, we had chosen a 20 percent difference in proportions as the end-point which we were achieving in clinical results.

It was a multi-center trial of 30 centers, in 11 different nations, and it could not be performed in the United States because we don't see enough cases of the disease.

And we had desired to have 284 evaluable patients. We enrolled 311 patients, and the study was stopped because of the concern for liver toxicity in adults, but it was not observed in infants and children.

But because of that concern, we stopped the study at 311, and 65 percent of the patients were evaluable, which gave a total of 203 at the time of the end of therapy, 203 patients, which was underpowered then for even a 20 percent difference in proportions.

However, there is important lessons to learn from this study that apply directly to any consideration of a drug in the future. Here are some

of the demographics.

The age is comparable by 2-1/2 years, and that is about reasonable for infants and children. Symptoms. The number of days to enrollment, 3.1 and 3.2, is long, because the standard deviation, you can see, is broad.

And there were at least three institutions in the study from other countries in which the delay in diagnosis was 4 to 6 days, and the outcome in that group was clearly inferior, and that is a problem when you go outside the country, that the duration of illness is often longer.

Approximately 40 to 50 percent of patients received prior antibiotic therapy, and by definition they could receive no more than one dose. But let me remind you that one dose intermuscularly of ceftriaxone will sterilize the spinal fluid of meningococcus disease in many of the patients.

And in those that it does not sterilize it, or any drug, we know that it drops the log concentration of bacterium CSF, a study that we did in the '70s, Bill Feldman and others, that showed clearly a two log drop, even with oral ampisone, with a number of the different agents.

So if you drop the log concentration, a

drug is going to look easier because you are dealing with many 10 to the 4, or 10 to the 5, on admission with the study drug; compared to 10 to the 7, which is the average concentration in spinal fluid of bacterium.

Looking at etiologic agents, it is reasonably distributed, but let me remind you that we really want to see Strep pneumoniae. That is the most difficult to treat, and it is the one that is resistant, and we see that it is not always easy to get, and it is not going to get easier.

Meningococcus is nice to have, but anything works for that disease, and so it doesn't tell you much. If a single dose of a sulfonamide works for a bacteriologic cure, I am not going to be too interested in whether a comparator works to an experimental drug, because they all work for that. So it is a very important consideration.

Now, here is the clinical and microbiologic end-points. Now, remember we chose a 20 percent difference in proportion, and by the FDA standard of 10 percent, the trovafloxacin would have looked inferior.

Now, there are two mistakes here. This should be minus 2.9 percent, and this is minus 4.8

percent. So they are all minus here, tilting against trovafloxacin.

You can see the 95 percent confidence limits do not exceed the 20 percent, but clearly the 10 percent it does. So, does this mean that trovafloxacin in this particular study was inferior to the comparator, which was ceftriaxone, with or without vancomycin?

I don't think so, and let me explain why. First of all, look at bacteriologic success, and I ask you a simple question. What is the purpose of antibiotic therapy for bacterial meningitis? To eradicate the bacteria. It does nothing else.

So, bacteriologic eradication, 98 percent, minus than 1 percent, very tight bounds. There were eight patients who had a delay in bacteriologic eradication. And 6 of those 8 had poor outcome, totally expected.

Now, let's look at the ITT analysis. The last was for protocol. And here we encounter some problems. You can see here at the end of the therapy there was clearly a big difference. Now, why is that?

Well, if you look at the designation, clinical success, and then come down and say 13 patients were considered clinical failures. Those 13

patients were in two centers in one country outside the United States.

And 11 were in the trovafloxacin arm, and two were in the ceftriaxone arm. And nine had haemophilus meningitis. All 13 had immediate sterilization of their spinal fluid.

And 11 of the 13 had follow-up at 5 to 7 weeks, and at 6 months, were considered normal. And yet they were called clinical failures, which we had to designate. And that is because the investigator had a concept of what was expected.

It wasn't correct. Subdural effusions were called failures, and subdural effusions are part-and-parcel of meningitis and portend no poor prognosis, and have no bearing on prognosis.

So it must be very -- when you go outside the country to do these studies, it becomes very difficult. We had an oversight committee of non-investigators in the study.

We chose not to act on this because the drug was not going to be used again anyways, and so we decided to show all the data, and not eliminate those patients, but it represents an important point to consider.

This one shows the adverse event profile,

and the only significant difference was in abdominal pain, and more common in trova. I would point to the joint abnormalities which we followed.

This is at 5 to 7 weeks, but even following out to six months, there is no difference. In fact, it was a little higher in the ceftriaxone group. Next slide.

There are many restrictions on performing studies of antibiotic therapy for meningitis, and the first and most important in the United States, and in any developed country, is the development of the conjugate vaccines.

They have been a blessing. We don't see haemophilus disease in the United States. I have seen on meningitis as of Memorial Day, 1999. That was the last case.

Now we have pneumococcal vaccine, a conjugate vaccine, and it has been in the United States for two years, almost two years now. With the implementation of these vaccines throughout the world with time, we will virtually eliminate haemophilus, which we have where it is used.

And certainly it will reduce, if not eliminate, pneumococcal. Probably not eliminate. At least 50 percent of the patients are pre-treated, and

I told you what the issue is there. It drops the concentration or will sterilize if ceftriaxone is the drug administered.

The necessity to have large numbers as required by the FDA for a 10 percent difference in proportion is simply not possible. A requirement for a clinical end-point, rather than a bacteriologic end-point, I think is not reasonable any longer, particularly when you understand what the effect of antibiotics are in bacterial meningitis.

And of course we know the logistical problems performing studies anywhere, but most especially outside the United States. However, it is necessary to have study centers outside of the United States, outside of North America.

But to have those, we must do them, we must enroll them, we must conduct the study in the following ways. This is my opinion, and I feel very strongly about it. It must be FDA approved obviously.

We must have participation of U.S. centers, and most especially the principal investigator. They must have his center or her center involved. IRB approval in all centers.

Informed consent for every patient. And there must be a preliminary investigators meetings.

Everyone there to go over word by word the protocol for approval.

Now, the next two slides we can skip because they were covered beautifully before me, and probably much more authoritatively. Let's just go to the sample size estimates.

Now, we are talking about an 80 percent response rate, but let me remind you as we move outside the country, and we go to developing nations for these studies, 80 percent is not going to be the end point.

I just reviewed a study from Malawi, 582 patients with meningitis, and 40 percent response rate. Now, that is because of underlying conditions obviously, and this becomes a very important point, malnutrition, HIV, other conditions, have impact on the outcome.

So 80 percent is really a little high now, and I am going to use multi-center trials. And we knew that from the Trova study. Nevertheless, let's just take 80 percent.

And we know that the evaluation rate is actually 65 percent, and may even go lower than that because of prior treatment. It is become very common. So if you use 80 percent, 10 percent difference in

proportions is over a thousand patients.

If it is 15 percent across the board, then 65 percent evaluation, and it would be 462. If it is 20 percent, 262. So it shows you the range. I can tell you in a simple word that there will never be a meningitis study where 500 or more patients need to be enrolled. It is simply not possible. Next slide.

There is one paper looking at equivalence and randomized control trials of therapy for bacterial meningitis. It has not been published, but will be in our journal, the Pediatric Infectious Disease Journal sometime this hear by Kryson and Kemper, from the University of Michigan.

They looked at 25 trials since 1980, and all of these trials claimed equivalence among control and investigational drugs. Only two studies were designed to test true equivalence.

And 24 had sufficient sample size to exclude a 20 percent difference in case fatality rate, and three trials could exclude a 10 percent difference. Proving therapeutic equivalence will be a challenge. Next slide.

So the potential problems with enrolling centers from outside the United States, mainly in developing nations where these conjugate vaccines will

not have been instituted yet.

And even in some that were in the Trovafloxacin study that were large contributors to the trial are now using the conjugate vaccines. The problems include non-adherence to the protocol, and monitoring issues, and severity of illness.

And let me remind you that at least a third, if not more, will have underlying conditions in children, which will impact outcome.

Performing appropriate audiometric and psychometric evaluations, complete follow-up is often difficult. There is no system, and no infrastructure to be able to do that.

There will be larger percentages of meningococcus haemophilus cases, and lower pneumococcal, and of course storage of specimens. So let me again go back to what I think is the essential point here. An antibiotic has only one effect; to eradicate bacteria from the CSF, and we can very objectively measure that.

And we have found in the multiple studies that we have done that they follow the prediction from the animal models beautifully. Next slide.

This just shows a further breakdown from the trovafloxacin study that I showed you. So that in

18 to 36 hours, this was the difference trova versus ceftriaxone. Very close. This should be a minus 1.5 percent.

The bounds are very tight, and at 72 hours, even closer, very tight bounds. So this was a very objective end-point, and I think should be considered the primary endpoint in bacterial meningitis.

It is not to say that there shouldn't be a clinical harm to that as well. Now, I made a point earlier that the eight children in the trova study who had delayed sterilization, 6 of those 8 had poor outcome, death or severe sequelae.

We knew that and it is based on many studies, and this summarizes many of those studies, and shows that the positive or rather negative bacteriologic cure or positive culture at 18 to 48 hours and is on average is 8 percent, with a range of 2 to 23 percent, depending on the antibiotic.

And in a study that we looked at here, we looked at four control trials in Dallas. We had a 6.7 percent positive culture at 18 to 48 hours. These are all significantly different.

A higher rate of neurologic abnormalities at discharge, 45 versus 19 percent, and 45 percent in

those with delayed sterilization; and at follow-up, 41 versus 13 percent.

So a very big difference, and so one of the determinants of clinical outcome is bacteriologic response. So, in summary, the critical end-points for assessing bacterial meningitis, and the antibiotics for bacterial meningitis, are the following.

One, bacteriologic eradication at 18 to 30 hours. It validates the data in animal studies. Again, in my estimation, this should be the primary end point. We obviously must study tolerance and safety, and clinical outcomes should be evaluated at 6 weeks and 6 months.

The end of therapy is not very important, and 6 weeks and 6 months is by far the better end point. However, let me again point out that clinical outcome is very subjective. There are many variables, many variables that determine clinical outcome that have no bearing on which antibiotic was used.

These include duration of illness, and etiology, severity of illness at the time of admission, fluid and electrolyte balance, availability of intensive care management, underlying conditions, just to mention a few.

They are all independent of the antibiotic

given. However, the one determinant that is objective and does influence outcome is eradication of the pathogen.

My suggestion is to enroll approximately 300 patients to distinguish a 20 percent difference in proportion, and this is currently achievable using many centers outside the United States.

It will also provide enough patients to determine tolerance and safety, and of course bacteriologic success. A 10 percent difference in proportions currently, and in the future, is not feasible.

It cannot be accomplished in the type of setting in which we now have to study bacterial meningitis, because of the availability of conjugate vaccines and other factors that I have mentioned. Thanks very much for your attention.

CHAIRMAN RELLER: Thank you, Dr. McCracken. At the end of the presentations, and this afternoon, we will have ample time for questions and a thorough discussion of all of the issues presented.

Our next speaker is Dr. David Shlaes, who will give the industry presentation for PhRMA. Dr. Shlaes.

DR. SHLAES: Hi, and thank you very much.

My name is David Shlaes, and I am presenting PhRMA today. Just a little bit about me. I spent 16 years in academic medicine, working mainly on antimicrobial resistance, but also treating a fair number of patients in a Veterans Administration Medical Center with infectious diseases.

So today I am representing the Antimicrobial Working Group of the Pharmaceutical Research and Manufacturers of America. Next slide, please.

This group offers a forum for exchange of scientific information among PhRMA companies, and our deep commitment to anti-infective drug products. It provides industry's scientific perspective in response to proposed rules, draft guidances, and relevant issues affecting anti-infective drug products. Next slide.

In our working group, there have been a large number of companies involved. We have had prior meetings with the FDA and a number of teleconferences and other meetings within our Antimicrobial Working Group. Next slide.

Today I want to cover three topics, and just a little background on the antibacterial clinical trials and the selection of delta. Implications of

the delta in antimicrobial development, including a number of unintended consequences I think, some of which have already been discussed.

And then I would like to present a number of alternative proposals that one could consider going forward. Next slide.

So the key or bottom line messages that I will try and support during the talk are what in our view is the current system for designing clinical studies and registering antibacterial drugs has worked well.

In fact, we recognize that there is always room for improvement here, but in our view this system has worked well, and a lot of the considerations that you are hearing about today are mainly theoretical ones.

What you are also hearing is that a single approach for all antibacterial drugs, for all indications, is unlikely to be an optimal one because of the differences in patient populations, variability from one patient population to another, and even within the population that you are studying.

Clinical studies must be feasible as you just heard from Dr. McCracken. The sample sizes must be practical. We have to be able to get these studies

done in some reasonable period of time for a variety of reasons.

And also we need to be able to do studies that direct our attention to areas of public health need, something that we will talk about more tomorrow. Now, one of the major ways that we can address the worry about bio-creep is in choice of comparator.

And I would say in the example that Dr. Soreth cited that this may have been just a problem of choice of comparator and poor study design, rather than actual bio-creep related to statistical concerns around the delta. So PhRMA's proposals are offered in this context. Next slide.

Now, there are a few differences comparing anti-infective drugs with drugs in a lot of other therapeutic areas. First of all, in the case of anti-infectives, we can get considerable information about activity against targeted pathogens from our in vitro testing, from animal models, and from pharmacokinetics and pharmacodynamics.

And this is something that is not shared by many other therapeutic areas. We do carry out trials with rigorous design, usually using an active control.

And it is important to keep in mind that

the magnitude of efficacy observed in a given study as you have already heard varies with the severity of the pathogen, or of the infection rather, the specific pathogens that are involved, and a variety of other conditions.

And therefore within any given population there is going to be a certain variability. Next slide.

Now, the approach of the FDA throughout the '90s as you have heard is the following. Regulatory approval has been based on evidence from multiple clinical studies, typically from multiple indications. So in most cases, there are two well controlled clinical trials for each indication.

The evidence must show that the success rate of the new drug is reasonably close to the success rate of an active control statistically; that is, that the new drug is not inferior to the control drug by more than a predetermined amount.

And that is the delta essentially, and the main assessment is to compare the lower bound of a two-sided 95 percent confidence interval on the difference in success rates for the new drug, versus the active control, to a pre-specified limit, or the delta. And this was explained actually by Dr. Temple.

Next slide.

This just shows again the step functions to remind you as explained in the FDA's 1992 points to consider, which we think is still a very reasonable way to approach clinical trial design actually, where we have a sliding scale of delta, with a cure rate.

This does allow for reasonable trial sizes, varying with severity of infection and cure rate. Next slide.

One of the major merits of the step function is that it recognizes that one size does not fit all. So that there is a smaller margin when comparative success rates are higher, and therefore a higher hurdle for new treatments, compared with very effective controls.

The step function recognizes the magnitude and variability of the success rate to establish non-inferiority criteria. It recognizes the need for both statistical and clinical aspects of efficacy evaluation.

It supports study design using realistically achievable sample sizes, which I think as you have heard is a clearly important consideration.

And the approach in fact has been used

effective for a decade of drug development, and we as you heard earlier, I don't think anybody is aware of any evidence that newer agents approved to treat serious infections, especially those involving resistant pathogens, are less effective than previously approved products.

This is just a list of some effective products that have been developed and approved sine the early 1990s using this approach, and again I don't think there is evidence that this approach results in the approval of inferior products. Next slide.

Now, there are some implications of a smaller delta, and I would like to go through a few of those. Clearly, there is an increased time to drug availability.

So that if you carry out a trial, for example, in the example that Dr. McCracken mentioned, where if you carried out a meningitis trial for a 10 percent delta, even at an 80 percent power, that trial might last for 5 or 6 years, if you could do it at all.

And the question is would the comparator that you chose at the start of that trial be relevant at the end of 5 or 6 years. Is that relevant? So there was a question about the validity of a trial

being carried over a number of years, and this adds further to the inherent variability in a given infectious disease indication.

And the other problem is the increased number of investigators that are required, which gives another source of variability. So basically what you get is a smaller delta, larger sample size, increased development, time, costs, and variability.

And as Dr. McCracken also mentioned, frequently increased numbers of investigators outside the United States, because you simply cannot gather or enroll the number of patients that you need to enroll for many of these trials within the United States alone. Next slide.

And I won't go over this because Dr. McCracken covered this in great detail. Next slide, please.

So what do you gain by reducing the delta? If you have a control cure rate of 85 percent, and a new cure rate of 75 percent, you run a 90 percent powered study with 120 available patients per group; and two trials, powered at 50 percent delta; and the risk of incorrectly concluding non-inferiority is 2.7 percent.

Therefore, I think in this design there is

very little risk of approving non-inferior products. So I am not sure how much advantage you get by reducing that delta to 10 percent.

The other thing that I will point out is that a lot of the examples that have been shown today assumed an 80 percent beta power trial.

If you run an 80 percent beta power trial, at a 10 percent delta, your chance of falsely concluding inferiority is about 30 percent, and most of us in the PhRMA group wouldn't run such a trial. Next slide.

So disadvantages will require considerably larger sample sizes. It is unrealistic for some indications in patient populations, and there is a disincentive therefore to develop new antibiotics, particularly for indications with inherently low success rates.

You just heard about meningitis, but that is not the only one. There are a variety of others, where you have seen very few clinical trials in the last decade.

Endocarditis, osteomyelitis, and those are neglected areas because of already statistical design requirements. The other problem is that by increasing the trial size, you could potentially unnecessarily

expose patients to investigational treatments for longer than what might be otherwise required. Next.

An increased cost and time will further disadvantage investment in new antibiotics and company's portfolios relative to other therapeutic areas. We are already seeing this, and fewer companies will be developing new antibiotics.

Because of this, there is a risk that existing drugs will continue to be used in lieu of a constant pipeline of new drugs, and even if there is an invest so that we get new drugs that delay an availability, we will continue to put pressure on the existing drugs just because of the increased stringency of the trial requirements.

And obviously the fewer new anti-infectives will be exacerbated by the current trend in industry towards dis-investment in anti-infective R&D infrastructure.

And this all leads to public health considerations, which I think we have to keep in mind. And we must have an ability to respond to these public health conditions going forward. Next.

Just to point out that anti-bacterial drugs are already disadvantaged in the R&D portfolios of the pharmaceutical industry. The reason for that

is that the antibacterial drugs are usually intended for short duration of use for acute diseases, unlike an anti-depressant, which you take for a very long time; and an antihypertensive, which you take forever, et cetera.

The size of patient population is relatively unpredictable and can vary dramatically from year to year, depending on the indication. And as I pointed out, an economic justification within companies is stronger for the development of drugs in other therapeutic areas.

So this therapeutic area is a therapeutic area within the industry that always sits on the brink. It is always on the brink, and it doesn't take much to push it over the edge. Next.

So what PhRMA would like to suggest is a number of alternatives. One is to continue to use the step function approach until an optimal alternative is agreed upon, and we think this basically works.

As I pointed out, the comparator agent should be a consensus standard of care and this should thereby address concerns about bio-creep in our view. And for indication specific deltas, a consideration of the seriousness of the disease, the variability of the response rate, and the feasibility of conducting

the trials, must be undertaken for each indication. Next slide.

There are several options. One could conduct two independent Phase III trials with a delta of 15 or 20 percent for each trial, which essentially is included in the step function as it stands now.

There is a low risk of incorrectly including non-inferiority in this case. One could conduct two independent Phase III trials, one larger and one smaller, with a combined analysis or Meta-analysis, providing a power of 95 percent, and a combined sample size using a delta of 10 percent to assess non-inferiorities. So you could achieve an analysis in that way. Next slide.

One could analyze results of trials by comparing the lower bound of a one-sided 95 percent confidence interval on the difference in success rates for new drugs, instead of using the two-sized confidence interval, and this in fact was suggested in the ICH E9 document.

Another approach would be to use the FDA's general equivalence definition for selected indications, and I will show the nosocomial pneumonia one on the next slide.

So this is just to summarize the general

equivalence for nosocomial pneumonia, where you would use one well controlled trial and an absolute clinical success rate of new drug no more than 5 percent in absolute terms, less effective than an agreed active comparator agent.

And this requires at least 80 patients in each arm, and clearly well-defined patients, and this sample size in fact, in measure of equivalence, describes an 80 percent power design and a 20 percent delta.

This would be quite feasible, and we believe we could do these trials in nosocomial pneumonia, and they would be valid. Next slide.

Now, we agree with a lot of the previous speakers, in terms of alternate designs for diseases where there may be placebo effects, such as acute bronchitis, acute exacerbation of chronic bronchitis, acute otitis media.

People have talked about a so-called rapid cure design, where again you could do a 50 patient per arm study, and evaluation at some time point, and we chose day four to five year, but it could be 2 to 3, or whatever the time point is, to show that active treatment provides a two-fold increase in success rate, compared to placebo.

And then a no improvement would be failure, and then failures are treated with open label antibiotics. Also, a time to cure design, where a placebo controlled study is done to demonstrate a 50 percent reduction in time to symptom resolution.

Obviously, this would have to take into account the severity of infection within these specific indications somehow. But these are approaches to getting placebo designed, placebo controlled, trials, and some indications for not serious infections.

So, in summary, PhRMA recognizes the medical need for discovering development of new antibacterial drugs. I think nobody more than me. PhRMA companies' welcome and rely on informative and realistic guidances to provide the latest thinking of FDA and its advisors.

This is terribly important to us because it allows us to know the path forward in the development of new drugs. We are planning a workshop for industry, FDA, IDSA, and other stakeholders, in order to define clinical and statistical standards consistent with efficient development of safe and effective antibacterial drugs.

And we hope that this will be part of the

process of coming to consensus on how we can go forward from here. And I think that is all that I have to say. Thank you very much.

CHAIRMAN RELLER: Thank you, Dr. Shlaes. Our next speaker with an industry presentation will be Dr. Francis Tally. At the completion of Dr. Tally's presentation, and before the IDSA presentation, I would like to have questions directed at the first three speakers, if there be any, including Dr. McCracken's presentation for him. Dr. Tally.

DR. TALLY: Thank you, Dr. Reller. I would like to thank the FDA for inviting me to participate in this advisory committee meeting. What I am going to talk about to day is the biotech approach to this topic.

The difference between big Pharma and biotech is that biotech companies usually focus in one area, and doesn't have the luxury of having several of the areas to support the research structure in the development group involved.

We also have a lower threshold for getting drugs into development, but we need to have a threshold. And we have strong influences to have frequent dialogue with regulatory bodies so we can take the most focused path in achieving a registration

of our drugs, because we don't have the luxury of studying eight different indications.

What I would like to do today is give a view from our perspective. The disclaimer about companies is on every slide, and I am the chief scientific officer of Cubist Pharmaceuticals.

But like David, I had a 15 year history in the academic world, studying a number of different drugs, and like Vince Andriole, was on the committee, the ISDA-FDA Committee, back in the mid-1980s to early-1990s.

I then went into industry and first worked in big pharma, and had the pleasure of registering a large drug for resistant infections with piperacillin or tazobacam, and also doing some discovery.

And for the last seven years, I have been at a small pharmaceutical company or biotech company, and we are currently developing a drug for the treatment of serious Gram-positive infections.

The majority of antibiotics developed over the last several years, or last 40 years, have been broad spectrum drugs, and we have had a number of "me-too" drugs in the same area, which I know has brought up a problem with development.

But now we are looking at different drugs

that we have both broad spectrum and narrow spectrum, and it is going more towards the narrow spectrum. We also have oral and/or IV, and there are special problems when you have an IV only drug with the practice of medicine in the United States, and now we are seeing the same problem in Western Europe.

And finally you will see existing -- modification of existing drugs, but what the big effort now in research is to develop novel classes of drugs with novel targets.

And I will touch on that a little more tomorrow in the resistance discussion. But I am listing some of the drugs here, and a couple that have been recently approved -- quinopristin, dalfopristin, and linezolid, representing an old class streptogramins, and a new class, the oxazolidinones.

On the other drugs that have been from existing classes, Wyeth and David's shop has tigecucome. amd we have dalbavancin and oritavancin, which are analogs of glycopeptides.

And ertapenem that Merck had approved was the pharmacological advantage of an important class of drugs. The other new classes we see are daptomycin and telithromcin.

The details of some of the drugs in

development to cover both VRE and MRSA are listed on this slide. I am not going to go into the details. It is in the handout.

But what I would like to do is to look at what justifies in 2000 the development of new drugs. First, you have to have microbiological superiority. I think the days of a lot of "me-too" drugs in the same area are over.

And particularly with microbiological superiority is going through resistance, and we will talk a lot more about that tomorrow. You could look for pharmacological advantages, and clearly one a day carbapenem that Merck just got approved is an improvement in therapy patients.

And so ease of administration, and finally safety advantages are always looked for at different classes of drugs. There are a number of different drugs around, and the only reason that I put this slide up is there are some cephalosporins coming along with MRSA activity.

And so I think you will be seeing a couple of these drugs come down to see whether or not they can hold out for MRSA, because as you will see tomorrow, one of the main problems we have in the future is at MRSA.

We have heard a lot about protocol design, and I think the drug's characteristics actually dictate in protocol design. Specifically, spectrum and distribution of drug is going to dictate what clinical indications you use.

You heard about the PK/PD guides to therapy, and they are just guides, because we need also dosing studies. And a preclinical safety profile is whether or not you are going to have this drug developed for broad indications and outpatient, or a restricted drug for use in serious infections.

We have heard a lot about superiority and non-inferiority today, and I think superiority trials are very limited in anti-infectives, probably to the out-patient oral drugs that David Shlaes just talked about, and some areas.

But in sick patients in hospitals where you have a known mortality rate, superior trials using placebo are not possible. And that's why we do the non-inferiority trials for almost all of the antibiotic trials for serious infections.

And I think there are a lot of data out there in the serious infections where we can look at rates. Finally, in considering these infections, you have to consider whether the infection is a

monomicrobial or polymicrobial.

My scientific area was in the study of mixed anaerobic infections, and depending on the type of infection, it presents a number of different challenges on control agents, and covering all of the infecting flora, because if you don't cover all of the infecting flora, you will have a higher failure rate.

And this is particularly true when you are picking the comparative agents to prevent the bio-creep that we have heard about. And it really dictates the comparative agents.

If you look at the narrow selection rate, such as complicated skin and soft tissue, with Staph aureus, and Group A beta strep, are the main pathogens.

We have very selected therapy in that particular area, depending upon whether you have an MSSA, or MRSA. And so it is either an amoxicillin or vancomycin, and that is what you are limited to.

But when you go to community-acquired pneumonia, or nosocomial pneumonia, because of the diversity of pathogens that you see in this disease, you run into a much different problem.

And when you run into this problem in different countries, you are also running into

different types of patients, which we have recently seen.

Indeed, in community-acquired pneumonia, you have Gram-positives, and Gram-negatives, atypicals, intracellular, cell wall minus, and so there is a whole host of therapies that could complicate your choice of comparative agents.

It is similar in nosocomial pneumonia, but it is much more limited because of the predominance of Staph aureus and Gram negatives, and with the high mortality rate that you see in these groups of patients.

When we are looking at trial design, to prove non-inferiority, you are looking at blinding. Everybody would like the Holy Grail of randomized perspective double-blinded studies.

However, with narrow spectrum drugs, you run into problems in your comparative therapy, and in the companion therapy for the potential pathogens that are not covered by a narrow spectrum drug.

I covered that a couple of years ago in one of the ICAHC meetings. You can get around some of those by investigative blinding, and it is not quite as good as double-blinding, but still you can come up with dialogue with regulatory authorities to establish

a well controlled study.

Open label studies are reserved for end-points which are hard microbiological end-points. You keep the microbiologist blinded, but not the physician.

We have heard a tremendous amount about sample size today of the patients enrolled in your study, and it is driven by delta. I don't have any numbers in my slides. I was trusting that everybody in front of me would have beaten that to death, and I am pleased that they have.

We are looking at 95 percent confidence levels, and then project efficacy rates, and we have heard a lot about that. And finally we are looking at end-points, be it microbiological or clinical.

And we heard from Dr. McCracken about the importance of the microbiological end-point in meningitis. We have also heard about the challenges with when you have a small delta.

In challenges of selecting a delta, you can look at is it better than placebo, and that is a superiority trial. It just requires a monitoring board because if you reach the statistical significance that the drug is working better than the placebo, you should stop the trial.

Like the Pharmaceutical Manufacturer's Association opinions that David just presented, I think the seriousness of the infection affects the delta.

You can look at mild infections, severe infections, or moderate infections, or severe infections, and you want to see that the drug is equal to the standard of care, and this is the concept of bio-creed.

Outside of the people in this audience, you really have to define what bio-creep is, and I think with serious infections that you want to select the best therapy.

I am going to skip bio-creep because everybody knows what it is, and the fear is that we will approve a drug that is no better than placebo, and I think that was nicely presented by the statistical group from the FDA.

And I think that it is important -- and one of the things that has to be developed -- and I would agree with David's recommendation, is that we should try and wipe out the bio-creep that has occurred.

And I know of a couple of other bio-creeps, particularly in impetigo, and cutaneous

ulcers, where when you are measuring the effect of drugs, when you give adequate care to these diseases, it is no better than good soap and water, and good nursing care.

And so it is important to prevent the bio-creep in this particular area. Once again, I am not going to go into the 1992 recommendations. That has been beat to death this morning.

I would like though to look at the impact of a small delta as David did, and the number of patients is greatly enlarged, to the point where it drives expenses way up, and for a small farmer, raising all their money on the open market, it puts added pressure.

But that's not a reason for not having a small delta. The time to complete studies may be in years, and I think this is a major impediment that has been pointed out previously.

One, you are losing investigator interest in the study, and if it stretches out over a couple of years, and you start to get poor patient selection, you may no longer have the appropriate comparator agent.

And when you are finished, you may not have the proper study after all that time. We have

heard about enrollment outside of the United States, and we have recently experienced that in community-acquired pneumonia by getting a very different patient population in other parts of the world, and as shown from our sub-analysis.

And that is because of the size of the study, we could not hope to enroll all of the patients in the United States. And finally the costs of drug development.

It is a burden on big farmer and on biotech and specialty firms, but that is something that I think -- my fear at electronic presentations. And this is Frank Tally's opinion now in collaboration with several of my colleagues at Cubist Pharmaseuticals.

And what would be my opinion on looking at deltas? I think for oral drugs for common community diseases listed here, such as skin and soft tissue infections, sinusitis and otitis media, bronchitis, UTI, and gonorrhea, this is the area where 10 percent deltas make a lot of sense.

There is big patient populations, easily enrolled, and you can clearly define the character at stake, and it doesn't take years to do the studies, and these studies can be done in the United States.

Indeed, I would even say that in some urinary tract infection studies, and in the treatment of gonorrhea, where the cure rates are very high, even a delta at 5 percent may be acceptable in these particular areas.

For IV drugs for more serious infections though, I would agree with the recommendation that David just put forth for PhRMA. When we are looking at different -- I am jumping all around. Let me go back.

(Brief Pause.)

One of the other ways to stop bio-creep is when you select a comparative agent, and I think it is important to select the standard of care, and I think there is a lot of guidelines coming from a number of the academic societies.

And I think this is an area that should be worked on to work out the standard of therapy to prevent the bio-creep from going forward. With looking at IV drugs for serious infections, what I did was look back at 2 or 3 of the drugs that have just been approved, and looked at the cure rates in nosocomial pneumonia, hospitalized community-acquired pneumonia, intra-abdominal infections, and complicated skin and soft tissue infections.

And most of them are not in the 90 percent area. Most are in the 75 to high 80s, and I think the delta for these should be carefully selected in consultation with the regulatory bodies based on the clinical knowledge of the disease in the hard end points.

And I think the sliding scale that David talked about that was exposed and published in 1992 still fits, and that there has been very little bio-creep in the IV drugs.

And that's because IV drugs only in the United States present major problems in doing the clinical studies. And if we put very small deltas on them, we won't be able to achieve enrollment of enough patients to come to the appropriate conclusions.

And the patient population is limited, although there are large numbers of patients out there, it is difficult to get them into these studies. Here it is imperative that you select the best therapy, because in these infections, there is an attendant mortality that you can affect.

And I think that this is an area where you have to go with the current standard of care based upon the bacteria involved, the resistance rates, and proven efficacy.

We have the further problem with IV drugs of in-hospital use and home IV use, and finally a number of these patients have switched to oral step-down, and for drugs without an oral component, if you switch them to another drug, it is currently considered a failure.

Whereas, really this has been a switch to oral therapy because of a clinical response. And I think this is an area which has to be worked on also in the development of drugs going forward.

Finally, we heard from Dr. McCracken about the problems with doing studies for meningitis. These are hard end-points when we look at meningitis. People die from this, particularly with strep pneumo.

We have been looking at endocarditis because of the characteristics of our drug, and we have been working closely with the FDA, and I think we have come up with an approach to this, because there has not been an endocarditis approval since the mid-1980s.

And a couple of companies have tried to study this area, but have been unsuccessful. And this is an area of unmet medical need. Why? Because when you look at endocarditis, there has been a change. Staph auerus is now a major problem with endocarditis,

and this is because of our sicker patients in hospital, and the higher incidents of endocarditis in hospitalized patients.

And with mortalities of 24 to 40 percent in Staph aureus and endocarditis, there is a major unmet medical need in this particular area. And so I think getting the widest delta in order to study this is appropriate, because like meningitis, you have a hard end-point due to bacteremia.

And there are a bunch of other confounding factors that go into this, but the hard end-point of clearing the bacteremia, because if you don't, you have the hard end-point that the patient has failed.

And so in conclusion, I think community-based common infections are where the most bio-creep has occurred. Therefore, small deltas are appropriate and the best comparative agents should be selected.

For intravenous therapy, and serious infections, the main problem is the clinical development, and where the physician should select the best therapy.

And in human studies committees, and the FDA, and the physicians themselves, will ensure that you select the best comparative agent. Thus, I don't think that bio-creep comes in in 2000 and into this

particular area.

The delta should be based on the statistical considerations that we heard, and clinical considerations in a comparative therapy should represent that standard of care.

And finally severe infections require the widest deltas, and it is fortunate in those that we have higher microbiological end-points, and the incidence of infection; that is, the patient population to do these studies is very low, and if you put a small delta in this particular area, it will continue to be an unmet medical need.

Finally, I think one of the things that I have been trying to bring about is it really takes a closer interaction between industry and FDA to come up with the appropriate design of the clinical studies for new agents, and I think we will hear more about this tomorrow when we are talking about the evaluation to drugs for resistant organisms. Thank you.

CHAIRMAN RELLER: Thank you, Dr. Tally. Questions for the first three speakers in this session? Yes, Dr. Goldberger.

DR. GOLDBERGER: Given that Dr. McCracken was kind enough to come all the way here for just essentially one day, we would be remiss if we didn't

make sure that we got the maximum use from his advice.

I first wanted to ask just a couple of basic questions. You were talking about important issues on severity of illness, patient's underlying status, et cetera, as being important components of outcome in meningitis, and not impacted, for instance, by antimicrobial therapy.

Is it fair then to conclude from your comments that you don't believe there are drug disease interactions with regards to treatment of bacterial meningitis? That all of the information basically is simply captured by what happens in the spinal fluid a X-hours?

DR. MCCRACKEN: Well, it is hard to be a hundred percent about anything when you deal with a complicated disease. but certain features of patients with meningitis that have clear impact are irrelevant to the antibiotic, and duration of illness, before the doctor ever sees them and they are enrolled.

The severity of the disease at the time of enrollment can be a one hour illness with meningococcemia shock and meningitis, and the antibiotic is -- the only effect it is going to have is on that bacterium.

Underlying HIV and underlying

malnutrition, availability of intensive care management, all of these things are really peripheral to the central issue of whether an antibiotic is effective or not.

Now it is not to say that an antibiotic doesn't have interaction, and of course there are people who are interested in the possibilities of the anti-inflammatory aspects of the drugs, et cetera.

But at this point, I think the clearest and most objective end-point is bacteriologic cure in the spinal fluid. And we know that is one of the variables, and probably the only variable, that an antibiotic has clear impact on. It eradicates that bacterium.

And in fact I feel so strongly about that, that I think you could use a delta 5 percent for that, and if a comparator is inferior, and is less than 5 percent on the 95 percent confidence interval, I don't think that drug should be considered. I think it should be very narrow, but the clinical one is much more difficult.

DR. GOLDBERGER: Well, our concern might be to use an example. If you had an infection with haemophilus influenzae in a person with bronchitis, and assuming you felt that the patient needed to be

treated, you might be comfortable with using a macrolide antimicrobial.

If you had established haemophilus influenzae pneumonia, you might very well want to be looking at a different class fluosoquinolone third-generation cephalosporin.

I just wanted to get your feel whether issues like that exist within the area of bacterial meningitis from your perception.

DR. MCCRACKEN: Yes. I would not consider the use of a bacteriostatic agent. You want cidal activity, and so although the general concept, and beautifully illuminated by Bill Craig, is the AUC over MIC for consideration of fluoroquinolones for systemic infection.

I won't accept MIC. It has to be MBC. I want cidal activity. So as that goes in classes of antibiotics, there would be some that I would consider clearly inferior and should not be studied. Within the classes, it would depend on which the agent is.

But as long as it has two characteristics -- well, three, but two characteristics from a meningitis standpoint. One, it penetrates well. It maybe has lipophilic activities, much like the lipophilicity, like the fluroquinolones.

And so it gets into the spinal fluid, and two, it has demonstrated cidal activity; first in the animals and then in the human. Of course, there are other features; safety and tolerance, and all of those.

But other than those two, which you can clearly demonstrate before you even get to a patient, I don't think the class matters as long as it is cidal.

DR. GOLDBERGER: I just wanted to make an observation. You were kind enough to go through some of the trovafloxacin data in some detail, and we are sort of forced to be in the position regrettably of having to be at times skeptical when we look at information.

But looking at that data, the kind of questions that probably would come up if someone here were reviewing that, for instance, to get that indication for trovafloxacin, were the proportion of the retreated patients in the trovafloxacin arm was noticeably higher.

The proportion of pneumococcal infections in the trovafloxacin arm was notably lower. You correctly brought up this issue of the early failures,

and how it didn't seem as though that was related to microbiology.

Yet, the kind of thing that would always bother us is that there were 13 early failures, and 11 were in the trovafloxacin arm, and only two in the comparator.

And as you can imagine, when we look at data, we are forced to just look at that and wonder, well, why did it turn out that way. And I was wondering if you had any observations about that, and also just to give you our perspective.

And although we agree with you, that big trials are a big problem. These are the kinds of problems that come up when you have smaller amounts of data.

DR. MCCRACKEN: I think those are very justified concerns. Indeed, the smaller number of pneumococci is worrisome, because that is the one pathogen that you would like to have for bacterial meningitis.

I mean, meningococcus, when I reviewed the data from Malawi for a paper in the Lancet, the case fatality rate for meningococcus meningitis was 4 percent. The case fatality for haemophilus was 30 percent, and 35 percent for strep pneumo.

Well, 30 and 35 percent for those two organisms, in the United States, it is 4 percent for haemophilus, and 8 percent for pneumococci, and yet you see the huge difference.

And so one agent, given as a pre-dose, or prior therapy, can have a huge impact on meningococcus. So I tend to discount that and look more to the other two agents, and most especially pneumococcus. So there was that issue.

This early failure thing gets down to one issue. I mean, I hate to mention it, but it was a bias of the investigator. He did not like fluoroquinolone, and he should never have been allowed in that study.

He did not come to the investigators meeting, and that is the issue. And that's why I pointed out that it is unacceptable, totally unacceptable to do a study now where an investigator is not part of the original description and review of the protocol.

And if that investigator feels that the protocol is not suitable for his or her institution, fine, it shouldn't be in it. But that wasn't what happened there, and so we had to go back and look at that, and see why was there a failure.

And he just had a bias towards the other drug to compare. It is unfortunate, but fortunately trovafloxacin is never going to be used for bacterial meningitis. So it wasn't an issue.

It would have been an issue had it been -- I would have made a big issue of this, and probably appealed to the FDA if it had ever come to them for this. It was purely an error in that regard.

CHAIRMAN RELLER: I would like to ask the same question and comments from Drs. Tally, McCracken, and Dr. Shlaes. In your presentations, there was a recurring themes that for some of the most serious infections, where the numbers of plausible patients enrolled would be the smallest, such as infective endocarditis, meningitis, the deltas should be larger.

But paradoxically those infections also, at least some of them, have the most objective end-points. Where on the other hand, Dr. McCracken has emphasized that deltas could be very small, 5 percent or less.

And then the analogy to a not so serious infection, where in fact there are specific threshold criteria for even considering the efficacy of the drug, and specifically gonococcal infections, where the eradication rate must be 95 percent, or any other

considerations in approval of the compound are not considered.

So my question is this. Should we consider different deltas for clinical end-points and bacteriologic end-points with specific infections? And also pathogen specific.

So, for example, with meningitis, that if there were approval, there would have to be X-number of patients with pneumococcal infection, and they would have to have a 95 percent or delta 5 percent eradication of the organism by specific methods at particular points after initiation of therapy.

And that other considerations of second end-points for clinical outcomes at 6 weeks, 6 months, follow-up blood cultures at X-number of months with endocarditis, might have different criteria.

Because it seemed to me that one of the driving issues for considering wider deltas was not a clinical reason, but rather a practical reason having to do with economics and number of enrollable patients.

So how does one bring those clinical necessities, objective possibilities of really tight criteria for efficacy microbiologically into consideration with the realities of the numbers and

the economics? Drs. Tally, McCracken, and Shlaes, comments on those possibilities?

DR. TALLY: That's why I put the paradox on the severe infections, because if you look in the response from the FDA in the beginning of this material that was handed out, that is the paradox.

You want more surety in the most severe infections. But the fact is that when you put that tight clinical delta, you are increasing the size where you never are going to have that study to be even -- to measure anything.

So what are some of the alternatives? And one of the reasons that has been pointed out is that if there is a hard microbiological end-point, and I think we should talk about your proposal with that microbiological endpoint, because it is going to be clear early on that if somebody doesn't clear their bacteremia by the fifth or sixth day with endocarditis, I mean, that is a clear failure.

And as you move along -- and it may come down to a smaller delta with that clear number of patients. It is in designing a study to say that you have to enroll 600 patients in a study, I think you probably with these various serious illnesses, with the hard base line, that you can do lower numbers, and

draw valid conclusions from those lower numbers.

And based upon everything that Mark was just saying, taking everything into consideration, and the different pathogens, and the predicted outcome in those.

But a priority to say that you have to do a 700 patient study in endocarditis, you are never going to see that based on that small delta. So I do think you open it up for different approaches. David, do you want to comment?

DR. SHLAES: Yes. I think the comments that we made were based on the current clinical outcome at trial design. Clearly, if you have microbiological end-points, and one of the points that we are going to make tomorrow, and which we will start make tomorrow, is that it is about time for us to be using surrogate end-points in trials of anti-bacterials, one of which could be bacterial eradication.

And it is something that has been done in the anti-viral group for a very long time already, and so I don't see any reason why we can't do it. I think you could have smaller trials with hard end-points using microbiological end-points for certain infections.

I think though that your suggestion of having different deltas for each specific pathogen within an indication is going to get down to being difficult to get the appropriate number of patients for those cuts.

So you probably will have to take microbiological end-points in all-comers for a number of those infections. The other limiting factor would be, for example, an osteomyelitis, to getting follow-up cultures will be technically an issue.

And having enough centers in the case of otitis media that could do tabs to support all of the development that might be going on might also be an issue.

But I agree with the idea. I think we all agree with the idea that microbiological end-points is a very good way of going forward, and it is long overdue.

CHAIRMAN RELLER: Dr. McCracken.

DR. MCCRACKEN: Well, it is an interesting question, Barth. I hadn't really thought of you quite the way that you put it. But I can tell you that in 30 years, seeing I don't know how many hundreds of cases of meningococcal meningitis, I have only seen delayed sterilization once, and that was because the

wrong drug and the wrong dosage was used.

So there you could have a delta of one percent. I mean, that is a rule. You get bacteriologic cure. With pneumococcitis, and haemophilus, it is not a rule.

The studies in the late '70s and early '80s by Ken Altland showed about an 18 percent to 20 percent delayed sterilization at 18 to 24 hours. But by 36 hours, it was a hundred percent.

So it depends on when that end-point is taken, and I would definitely never go out beyond 36 to 48 hours. I think the end-point, if it is taken at 18 to 24 hours, can be a little broader. Maybe 5 to 7 percent.

But if it is taken at 30 to 36 hours, then it should be very tight, because by that time you have cure. I am talking about meningitis only, and I am not addressing issues of endocarditis or other diseases.

CHAIRMAN RELLER: Dr. Wittes.

DR. WITTES: Yes, I have a question about when you develop a new drug, do you in fact expect that it is no better in terms of cure than what is on the table?

And the reason that I am asking this is

that as Dr. Brittain pointed out in her presentation, the sample size is really driven by the assumption that the underlying rates are identical. And that's what makes the sample sizes really high.

But if in development you have seen better bacteriologic end-points, and you really believe that the efficacy, the clinical efficacy, is even slightly better than the comparator, then the sample size goes way down.

So my question is that in development are you aiming for improvement that you can't see, or are you aiming at equality?

DR. SHLAES: Okay. So I think, at least from our point of view, and I have a few colleagues who will chime in, I hope, when you look at the variability in the population within and an indication is such that it is very hard to prove a superiority in terms of clinical end-points, such as a cure.

And if you look at other end-points, such as time to cure, you might be able to do superiority trials, or there may be other end-points that may be more applicable to a superiority study.

But if you look at the usual clinical end-points, superiority is difficult to show. The other issue is that if you actually run the numbers on a

superiority trial, taking all-comers in a clinical study for clinical end-points, they are actually not all that much difference.

Again, at the 90 percent power. So if you do a 2 percent superiority study and a 90 percent power, and you account for non-evaluable patients, the numbers actually get to be just as large as they would be in the current step function study.

So I am not sure that in terms of patient numbers that there is an advantage there anyway.

DR. WITTES: But you are answering a different question. Can I clarify the question?

DR. HARDALO: Maybe I could also add something in. When we develop a drug, we really believe based on our animal data, and our lab data, that it is better than what exists.

However, real life often times gets in the way of proving that. And as Dr. McCracken said, and as I am sure as Dr. Talbot has experienced, that in diseases where there is a significant mortality rate, like VRE infections, or bacterial meningitis and immunocompromised hosts, or I can name a whole list of infections, including endocarditis with Staph aureus, and hospital-acquired pneumonia.

The inflammatory sequelae caused by the

bacteria is responsible for the vast majority of the morbidity and mortality that ensues. Therefore, I take a clinical only based end-point is going to be very difficult for you to prove that significant differences between the treatment groups exists.

And there is no way that one could do a placebo controlled trial, and because there is inherent variability in the patient populations, you will enroll, least of which is the standard of care in the center that you are having in your study.

And it can present significant issues for trial design, and it is not always something that you can take care of in a prospective stratified, randomized, clinical trial.

DR. FLEMING: Can I make a suggestion in the interest of time? I think Dr. Wittes is raising a very key point. I am going to be discussing this in some detail in my presentation, and maybe we can return to it after that if there are still remaining issues?

DR. WITTES: Sure. I just wanted to make it clear that you both answered a question different from the one that I have asked.

DR. FLEMING: Yes.

CHAIRMAN RELLER: Thank you, Dr. Fleming,

and that is the approach that we will take. We heard from Dr. Hardalo, and we had earlier hands up. Dr. Maxwell, do you have a question, and then Dr. Bell, and then Dr. O'Fallon, and then we will get on to the next presentation.

DR. MAXWELL: Yes. The question is for Dr. McCracken, just to clarify for me. Would bacteriologic outcomes, and let's say in the case of haemophilus meningitis, be the same in a child that had the vaccine, and one that didn't? Should it have the same exact measure?

DR. MCCRACKEN: Well, one would hope that the child who received the vaccine wouldn't develop the disease. With haemophilus, they wouldn't, most likely. With pneumococcus, we are seeing a couple of failures, and their disease looks identical to those who had gotten no vaccine.

And the reason is that the spinal fluid is a sequestered or privileged site, where there is no native immune function. Antibody compliment white cells are not present.

So the organism, once it gains footing there, can multiply without any control from immune function until late in the course. So if it develops, which is less likely in the vaccinated child, it

probably would have a similar course.

And this is not true necessarily in the systemic disease, but for meningitis, I think it is, yes.

DR. BELL: I wonder if the speakers could comment on how these issues apply to the development of drugs for resistant infections in particular. Are those study designs considered to be superiority trials, in the sense that the new drug has to be better than the drug for which the drugs are now becoming resistant?

Do they also have to meet non-inferiority criteria in the treatment of sensitive infections? What are the implications of some of what you have been discussing specifically for resistance? How do you address that issue?

DR. SHLAES: Actually, I think we are going to have a whole day on resistance tomorrow. Can I hold -- are you going to be here tomorrow? Can I hold you off until tomorrow?

CHAIRMAN RELLER: We will do that tomorrow. Dr. O'Fallon.

DR. O'FALLON: I have a couple of questions. I am trying to understand the thinking process that has been processed in the documents that

we have seen from industry in our packet.

The first one is I was a little surprised, or there was some support that has been voiced for the delta procedure. Now, I am a little bit puzzled as to why that is considered a good idea to be able to when you have a very successful comparator, that you would want to spent a lot of patients to try to prove a very small difference.

Whereas, you want to spend far fewer patients when the successful rate is down closer to 50 percent. You know, 70 percent, 65 percent, and that sort of thing. You are willing to spend half as many patients to try to prove what you would call efficacy as being non-inferior to the other thing.

Why are you not asking instead to just hold a sample size constant for your study, and then take the delta that comes out of that?

CHAIRMAN RELLER: Dr. McCracken.

DR. MCCRACKEN: I don't know exactly how that applies to what I am -- well, I don't know what you are leading to with regard to --

DR. O'FALLON: I don't think you spoke in favor of the delta method, and some of the others did.

DR. MCCRACKEN: Oh, I am not against the delta method. I just want a broader limit. I think

it is wonderful, and I just propose that it be a 20 percent difference in proportions for clinical outcome and a much narrower one for bacteriologic outcome.

DR. O'FALLON: But the delta is defined to be a step function where you spent fewer and fewer patients in order to establish a bigger delta. Why do you go with fewer patients around when there is a lower success rate? What is considered to be, or why is that a good idea? It is not obvious to me.

DR. MCCRACKEN: Well, I don't know if it is a good idea or not, but unfortunately what you are faced with, with bacterial meningitis, when you leave the United States and go to developing nations is a very good outcome.

That is to say that the clinical outcome there is probably in the rage of 60 to 70 percent success, and maybe not even that high. Therefore, it is easier to do a study because you might be able to show a difference with the smaller numbers.

But my point was only that using a 10 percent difference in proportions for a disease in the United States, or even throughout, we can't get a thousand patients. We just cannot do that any longer.

We need -- we -- I -- it is not me, but to do a study, and for me to be a principal investigator

of that study, I can't -- 10 years is too long.

DR. O'FALLON: I understand that part.

DR. MCCRACKEN: I may not be here in 10 years.

DR. O'FALLON: But why not go for a, sat, set number of patients; that you are going to serve a minimum sample size, and then take whatever the delta is that you can buy with that. Spend fewer and fewer patients, the harder it is to distinguish the differences.

DR. MCCRACKEN: Well, I guess my -- and probably statisticians can answer this far more competently than I could, but I am afraid that if I used -- whatever that defined number of patients would be, I am afraid that you might be surprised by the outcome.

It could by chance be that you have a much better outcome in the countries that were selected, and therefore, it is in the 80 to 85 percent range, and small numbers would give you inferior data, and you couldn't tell the difference.

So, therefore, you would shoot yourself in the foot by preselecting without knowing exactly where you stand. And that would worry me.

CHAIRMAN RELLER: We need to get on to the

same presentation, and Dr. Albrecht had a comment that she wished to make.

DR. ALBRECHT: Actually, I wanted to just follow up on the microbiological discussion that we had earlier. Dr. McCracken, you indicated during your presentation of the trovafloxacin data that the patients that were pretreated even with a single dose, you could often see up to a two-fold reduction in the colony count when the patients were entered.

So I just wanted to use that as an opportunity to ask whether we might consider if we are going to hear suggestions about microbiology a quantitative approach to microbiology.

And I just wanted to mention that we use that in the evaluation of urinary tract infection agents currently, but not in other sites, and in meningitis, a sterile site, I would appreciate comments on that.

But also then in the afternoon as we hear other presentations, I would like to raise that same issue relative to sites that are not normally sterile.

DR. MCCRACKEN: I mentioned that there can be up to a two or even larger log count drop in the pre-treated, and that was based on data int he '70s by Bill Feldman, in which ceftriaxone was not one of the

agents used.

It was mainly ampicillin and other drugs, and amoxicillin, which had an impact. Ceftriaxone might even have a greater impact. The problem with doing quantification of bacteria in CSF is that it is not as simple a thing to do.

The investigator who did that study was up all night. He came in whenever a patient came in, and that is a tough chore. You could put it in the refrigerator. It is doable, but it is very difficult, particularly when you get outside the country to actually do quantification.

You do get a rough estimate of bacteria by just looking at the stains smear, knowing that the break point, and seeing bacteria per field, is about 10 to the 5.

So if you see multiple organisms, which we have a child in the hospital now, probably has 10 to the 8, or 10 to the 9, and we know the outcome there is very poor.

CHAIRMAN RELLER: It is time to hear from the Infectious Diseases Society of America, a group that is very much involved, both in the development and carrying out of clinical trials, as well as importantly in the use of these agents in clinical

practice. Dr. Andriole, your team.

DR. ANDRIOLE: Thank you, Barth. As I pointed out earlier this morning, we are here to represent the Infectious Disease Society of America, and my colleagues, Jack Edwards, and Dennis Wallace, and George Talbot.

As you know the society has now more than 7,000 members, and it was founded 40 years ago. I was one of the founding fathers. No comments, please. And the member really cover all of the areas of infectious disease.

And without being arrogant, they are people who have contributed their life to studying particular issues, and I know that you recognize this. Seven of you on this committee are members of the society.

And so the agency has to recognize this, and one, as a past president. In addition, we have some very excellent members from the pharmaceutical industry who are members of this society.

And that we would like to help the agency accomplish the goals that it has set out to do. My involvement with the agency, as secretary of the Infectious Disease Society -- and Lillian will remember this if she is -- yes, she is right here.

Awe were very concerned about clinical investigation, and the guidelines that had been written in 1977 were pretty much outdated. And so in the mid-1980s, or actually the late-1980s, the Society and the Agency put together a task force to redo the guidelines.

The late Tom Beam was our liaison, with Matt Lufkin, and Lillian, and Dr. Peck. And we volunteered -- all of the members of the society volunteered to come down and to write guidelines.

We were given two years to do it, and in two years, and this is the flow sheet -- this is a classic paper -- we wrote 13 guidelines. And they were finished in 1990, June 24th.

Now, that is a decade ago, and I think they have served us well for the majority of that decade. I have also been co-author of one of those guidelines, and I was a member of this committee for 3 years, and paid my dues, and did all of that.

And I have been doing clinical research in Phase III and IV trials for 43 years. But now that I have joined the more mature population, I don't do that any more.

How I wound up being the spokesperson for this meeting is not clear to me, and I just have drawn

the short straw, and once I was told by the council that I was going to be the speaker, reminded me of a story.

When I was teaching in Kenya, on the edge of the Serengeti, I had wanted to go down and visit a village of Masai warriors. So my wife and I went down there, and I was talking to the chief, and my wife was playing with the children and talking to the women.

And the chief looked very sad, and I said to him what is the matter, and he said, well, I just lost one of my best warriors. I said, oh, that's too bad. What happened?

Well, he said that he was running across the Serengeti to come back to the village, and he came around a clump of trees and there was a lion. And he looked to his left and he looked to his right, and he looked behind him, and it was clear. There was no escape.

So he dropped to his knees and clasped his hands, and started to pray. And after five minutes passed, nothing happened. And so he looked up and there was the lion on his knees with his hands clasped.

And the warrior said to the lion why are you doing what I am doing, and the lion said to the

warrior, I don't know what you are doing, but I am saying grace. Well, that's how I feel right now.

(Laughter.)

DR. ANDRIOLE: First, I will be very brief, Barth, because somebody asked me before the meeting started what are you going to say. I said, well, I don't know what I am going to say until I hear what everybody else has to say.

And I don't have any slides, and so you are just going to have to pay attention to me, or fantasize, or whatever you want to do. But the fact of the matter is that everybody has touched on all of the issues that I have been instructed to tell you from the Infectious Disease Society of America.

I want to make a couple of points clear. One, as an organization, we have no vested interest in this agency, or in the pharmaceutical industry. I am here as a representative of the society for two major reasons. One, we want to be able to treat our patients with the best medical care.

And without the continued development of anti-infective agents, forget it. We will be out of business. We want to help people, and we know that the agency doesn't want to embarrass itself by preventing the development of new agents. That would

be a tragedy.

Number 2, you can sit here and talk all you want all day long; the industry and the agency, who does the work? We do. We are the clinical investigators.

So we beg you, we know that he current guidelines should be updated in different ways, but we are a little concerned about the criteria, because if you set the bark too high, you can't do the work.

And George McCracken said that very clearly, as have others, by discussing the mathematical approach to clinical investigation. We would like to the agency to adopt a scientifically and statistically appropriate, but also a clinical practical approach, to determining efficacy.

I don't care whether you want to call it a delta, or a mega, or a zero, or whatever. But that is what we would like to see. That you have when you review these NDAs that come into you in trucks, and electronically now, that you have a reasonable chance of evaluating this data to determine whether we are going to get to use it in our patients.

Now, is it -- do we really need to focus on a delta? Is that going to be the end point for clinical investigation? I mean, you just raised that

question. Is that he end all and the be all of what we should be doing? I don't think so, and neither does the Society.

We think that you have to evaluate, one, the frequency of the disease. If the disease is very frequent, make the delta whatever you want. Number 2, if the patients are not available in order to study thousands of them, then you have to come up with a different plan. You really do.

Otherwise, there is not going to be any more anti-infective research for the kinds of diseases that we need to treat. Well, how can we do that? We are not going to settle that today, but some of the suggestions have been already nicely stated by our colleagues who have already presented.

And some of the suggestions, and the details of all of this, the nuts and bolts in working it out can be done later. But we need to know what surrogate end-points we should be using based on the type of infection that we are treating. We have to really look at that.

And what are surrogate end-points? Well, George pointed out that clearance of the bacteria from the cerebral spinal fluid in meningitis. Others have asked the question can we do quantitative microbiology

and other infectious diseases.

That's a hard thing to do from a practical point of view, and there are other ways that you can use surrogate end-points; rapidity to cure is one that people are now looking at. Those are just some of the examples.

The second thing is that animal models of disease have been the bridge between Phase II studies and Phase III studies for years. And many of us have spent our lives developing animal models, which the agency has used in hits deliberation before a Phase III protocol is designed.

Pharmacokinetics and pharmacodynamics are extremely important. I am now speaking for the society, and they really feel that that kind of data is very helpful in determining whether a particular Phase III study is likely to work.

And finally the level of anti-microbial resistance in your ability to determine what the comparative agent is going to be. In patients who have very serious illness, we have to lower the bar. We really do.

An example -- Frank gave examples of this, and David gave examples of this, and these are very important things in our view. We wanted to compliment

the agency actually on the paper that you wrote on resistant pathogens.

We all went through that in great detail, and we thought that was really good, and we hoped that it could be refined just a little bit more. But the final message from the Society is the Infectious Disease Society of America is here to help you. That is the message that we want to leave you with.

We want to help in any possible way. We are prepared to volunteer any member of the Society. You tell us what you want us to do, and we will make a list of people that you can call on to help you solve some of these problems.

We have done this in the past, and Lillian knows that, and we worked very hard for two years to get done what had to be done, and we are prepared to do that now.

We will update your guidelines, and we will help you work out a delta. I don't think that can be accomplished in a big meeting like this. So we are suggesting that maybe the agency might want to consider a task force to meet with representatives from the Infectious Disease Society of America, with representatives of PhRMA.

After all, they are integral players in

this, and with representatives from the agency, to try to fix the issues that have been raised so clearly today. We have many qualified members who are really willing to volunteer their time, just like they did 12 years ago.

And that is probably the most important message that I have, Barth, from the Society. Any questions that you have, and I don't know, one, Barth wants to have the questions.

I have three distinguished colleagues who will be very happy to answer them, and I am very happy to have escaped the lion. Thank you.

CHAIRMAN RELLER: I think it would be actually a good time for questions for Dr. Andriole and other members of the IDSA. Not that they aren't also included on our advisory committee as Vince pointed out. Questions? Yes.

DR. NELSON: I would be interested in some comments on the surrogate end-point issue, and in particular whether one can extrapolate microbiological end-points from meningitis, which I thought was well argued based on clinical data, to other infectious diseases.

Working in an ICU and seeing the result of

host response, I would have to be convinced that there is no drug disease interaction that would have to be considered in some of these other conditions.

There was clinical data to support that use of surrogate end-point meningitis, but does that data exist in a lot of these other conditions?

DR. ANDRIOLE: Well, that is one of the issues that needs to be hammered out, and that is a very important question. Microbiologic endpoints in the intensive care unit in patients with a hospital-acquired pneumonia, forget it.

You can't even get the pathogen to begin with. You don't know what you are treating. But there are other surrogate markers that can be looked at, such as APACHE scores, temperature response, radiologic clearance, improvement, oxygen saturation.

Now, you can say, well, that might happen anyway, but it doesn't. That is a disease with a high mortality and you know that. But this is what we need to do to sit down and talk about what are the surrogate endpoints for each type of disease that are acceptable, and will provide information to help with the agency decide on efficacy. But I don't have any specific criteria.

DR. MCCRACKEN: I can give one. Acute

otitis media. The data are quite clear now that a double-tap study giving bacteriologic endpoints correlates beautifully with clinical outcome.

Now, studies are not easy, particularly in the United States, but that is a very good example of bacteriologic eradication in clinical cure.

CHAIRMAN RELLER: It was Dr. Nelson who imposed that question to Drs. Andriole and McCracken. Other comments from the IDSA in response to this query, or other questions? Yes.

DR. EDWARDS: Just to cite another example of consideration is the resolution of candidemia, which is in a problematical area for studying of the antifungals.

It is a complex issue again, but the surrogate endpoint of just the resolution of the candidemia is a factor to consider.

CHAIRMAN RELLER: That was Dr. Edwards. One of the constraints with the less commonly encountered, and often requiring many patients enrolled from outside of the United States, specifically meningitis, is there any room for looking at it from the direction of what are practical numbers of patients, and then what criteria experienced individuals would be comfortable with that would

demonstrate reasonable efficacy.

For example, the concept of if you had X-hundreds of patients, to demonstrate efficacy, you would need these etiologies, these deltas as regards eradication of organism at 24 and 48 hours, or 24 this delta, and 48 this delta.

And this latitude of clinical assessments out at six weeks or six months. Basically, not starting with a delta in one or the other areas, but starting with this is the maximum number of patients that are possible, and then how much information?

I mean, basically, it is issues of numbers versus quality of information in smaller numbers of patients. Dr. McCracken, any thoughts on that approach?

DR. MCCRACKEN: Well, I think it is an interesting approach. When I sort of threw out those numbers of up to 24 hours, or 24 to 36 hours, I really wasn't proposing those.

And I would really have to think about that in terms of numbers, because it gets a little tricky, particularly as you get the pneumococcal disease.

I think that approach is a very reasonable one, and I would echo Vince's comments that surrogate

markers become more and more critical as we try to evaluate diseases that are becoming less and less common.

I would think five years from now that there will be no meningitis studies in any developed nation with the prospect of a meningococcal vaccine, and already there are conjugate meningococcal and haemophilus vaccines, and that disease will be in small numbers.

And one could argue then immediately, well, why even worry about it. Well, it doesn't mean that it disappears. And it is in other countries, and resistance, and we all know when you disappear, or when one pathogen disappears, something pops up sometimes in its place.

So they are necessary. But your approach, Barth, I think, is an appropriate one, but I am not willing to give numbers yet because I really have not given it enough thought.

CHAIRMAN RELLER: This is only a concept to increase the repertoire of things that could be considered. It looks like it is time to hear from Dr. Thomas Fleming from the University of Washington on issues regarding choice of the margin in non-inferiority trials. Dr. Fleming.

DR. FLEMING: Thank you, Barth. Well, as Vince has pointed out, there has already been -- much has been said, and what I would like to try to do is highlight and amplify several of the key issues that are important in the choice of the margin. Next slide.

I think it is important when we are thinking about choice of margins to keep in mind as has been stated today there really is a dual goal here in non-inferiority trials.

First, to enable a direct evaluation as to whether or not the benefit to risk profile of the experimental therapy truly is adequate relative to the benefit to risk profile of the active comparator.

And also to contribute evidence to evaluating whether or not the experimental truly is superior to the placebo. Well, what I would like to do, and it is going to be kind of a quick overview, because a number of these issues have been covered, looking at factors that influence the choice of margin.

I will be talking about issues of clinical relevance, as well as active control effects, and I will be briefly talking about some issues that impact the interpretation of non-inferiority trial results.

Next slide.

So if we look first at issues of clinical relevance, and in choosing the margin, it is very important to consider the clinical relevance of the primary end point.

If it is a morbidity, major morbidity or mortality end point, even most changes in efficacy can have considerable clinical importance. At the same time, it is important to consider when thinking about the experimental against the active comparator, do we expect an alteration and hopefully an improvement may be in the safety or tolerance profile, and convenience of the administration, or other issues such as resistance or drug interactions.

If in fact there are important improvements in these areas to be expected by the experimental, that should in fact be factored in, in the choice of the margin, and it could influence choice of margin. Next slide.

The ICH guidelines also point out that factors relevant or related to the active control effect should influence the choice of margin. And essentially they are arguing that ideally we want well designed superiority trials to clearly establish the efficacy of the active comparator.

And that ideally, and this assay sensitivity issue that Dr. Temple referred to, we would like those estimates to be reliably predictive of what the estimates or what the actual efficacy of the active comparator would be in the non-inferiority trial. So, the next slide.

I would like to on this slide illustrate then three factors related to the active control effect that really should be influential in our choice of the margin.

First of all, ideally we would like to be doing active comparator trials in settings where the active comparator is very effective with a precisely estimated level of efficacy.

So, for example, to illustrate. Suppose that a placebo has a 45 percent cure rate, and the active comparator increases that to an 80 percent cure rate. And this is estimated to within plus or minus 10 percent.

So, for plotting here along this X-axis down at the bottom, the cure rate on placebo relative to active comparator, then the placebo is 35 percent less effective, with estimates consistent to as much as 25 percent less effective.

Now, Dr. Temple has pointed out, as has

Dr. Brittain, that in some settings that you might set the margin when you are choosing the margin to be specific to preserving a fraction of the effect. Let's say it is half of the effect.

If we use this 25 percent estimate, and we choose half of the effect, we might choose the margin to be 12-1/2 percent. Using this then in the non-inferiority trial, if the experimental or the estimate of the experimental efficacy is favorable relative to the active comparator, such that the lower limit rules out this margin, this is a positive result.

Now, this margin is greater than 10 percent, and part of what justifies this is we are dealing with an active comparator that is highly effective.

And if in fact it could be clinically argued that losing this much efficacy would be acceptable, then one would have a margin of this size. You might note that when I derive this margin that I used the 25 percent rather than the 35 percent estimate as a rationale for that caution.

And part of it is this assay sensitivity issue. Is the estimate of the active comparator obtained from these historical or placebo controlled studies relevant to the actual efficacy of the active

comparator in the non-inferiority trial.

So specifically suppose in these historical control trials we were looking at patients that were at lower risk than the patients that would be looked at in the non-inferior trial.

It might be that the active comparator is more effective in lower risk patients than in the higher risk patients in the non-inferiority trial. And there may be other differences as well in the non-inferiority trial from the active comparator trials.

Why are these issues important? Well, it may be that the active comparator provided a very big effect in the historical trials, but in the non-inferiority trial, its effect might be much more modest.

To position the placebo in green here might be much closer to zero, compromising then the ability or the integrity of using a margin of 12-1/2 percent.

In this setting it may be that using the margin of 12-1/2 percent not only assures us that we are maintaining half of the effect, but we may not even be able to conclude that we are maintaining any of the effect.

Other issues also relate to being cautious

when doing non-inferiority trials, and that is the quality of the design and conduct of a non-inferiority trial also raises factors that influence the interpretation, particularly in non-inferiority trials.

As the ICH Guideline E-9 indicates, many flaws in design or conduct of the trial will tend to bias results toward a conclusion of equivalence, such as eligibility criteria violations, non-compliance, loss to follow-up.

Why is that especially important here? Well, these types of biases in a superiority trial lead to an increased risk of false negative conclusions. They lead to an increased risk of false positive conclusions though in a non-inferiority trial.

I might focus for a moment on this issue of loss to follow-up. Next slide. And it is not uncommon in antibiotic non-inferiority trials for valuable datasets to involve maybe only 75 percent to 50 percent of the overall randomized ITT dataset.

If one is in fact excluding patients because of the absence of the targeted pathogen, then that probably just leads to an increase in variability.

But if we are much more seriously, and if we are excluding from the ITT, and if we are including in the invaluable, but excluding patients who are not assessed due to termination of treatment for reasons such as adverse clinical events, perceived drug ineffectiveness, or because patients took prohibitive concomitant meds, this is at risk of being what we would call informative censoring.

And it can substantially increase the bias, and hence in non-inferiority trials, these issues arise and should lead to greater caution in choices of margins, and in particular in interpretation of results in such studies. Next slide.

I would like to touch on an issue that was motivated by a question from Dr. Wittes, and that is on the issue of sample sizes, what we have heard a lot of discussion about is that non-inferiority trials, if we use scientifically rigorous margins, will always require very large sample sizes. Fact or myth? Next slide.

To address this, let's look at an active control antibiotic that has an 80 percent cure rate, and what I am plotting here along this X-axis is the experimental, minus the active control cure rate.

So, let's suppose that the experimental improves this cure rate by 10 percent. Then the experimental will have a 50 percent relative reduction in non-cure rates, reducing the non-cure rate from 20 to 10.

On the other hand, suppose the experimental has a 10 percent or 15 percent lower cure rate than the active comparator. One would then have a 50 to a 75 percent relative increase in the non-cure rate, issues that would generally would be viewed to be of concern.

Well, let's look at in the setting of doing superiority trials and non-inferiority trials when one has an 80 percent cure rate. Next slide.

Well, in this setting, I am again along this X-axis, and I am plotting the experimental, minus the active, control cure rate. And in a superiority trial one is trying to rule out the no-hypothesis of equality.

Let's suppose that the experimental arm truly provides a 12 percent improvement over active control in the cure rate. One can then obtain 90 percent power to rule out equality if one has about 340 evaluable patients in the pool sample.

A reasonable or acceptable sample size

generally, and yet one is having to presume a very substantial effect of the experimental. So, an alternative to this approach would be scenario two. Next slide.

And that would be a non-inferiority design, where one assumes a non-inferiority margin, and where one is essentially trying to rule out that the experimental arm has a 15 percent lower cure rate than the active comparator.

And in this setting, if the experimental truly is the same as the active comparator in the cure rate, then one would have 90 percent probability or power to rule out this margin with the sample size of about 300 patients.

A concern that often arises in this setting those is what if the experimental is 10 percent worse in cure rate, which is a relative 50 percent increase in non-cure.

One has almost a 20 percent chance of achieving a false positive conclusion. Next slide. And as a result, most rigorous non-inferiority margins of 10 percent have been advocated, and in that setting with a 10 percent margin, if the experimental truly is the same as the active comparator, one can have 90 percent power to rule this margin out.

But as has been noted, a substantially increased sample size is the price. Well, as Dr. Wittes was really getting at in her question, the issue is that in the superiority trial, we were having to presume a 12 percent improvement in cure rate in order to have good power.

Whereas, if that might not be highly plausible, what if it is highly plausible that the experimental is moderately better than the active comparator.

Wouldn't then we be able to rule out this rigorous margin with reasonable sample sizes, and the answer is yes, and that is scenario number four. Let's suppose in fact that the experimental is only 3 percent better than the active comparator and cure rates.

Then one would be able to achieve 90 percent power then to rule out this more rigorous non-inferiority margin with sample sizes that are in fact not a lot larger than what would have been required in the scenarios one and two.

It is important to recognize when one is looking at scenario number four these numbers in green. Essentially what these represent are what is the estimated success rate on the experimental, in

terms of cure rate, relative to the active comparator.

And in the superiority trial, one would have to estimate that the experimental arm provides a 7.3 percent increase in cure rate relative to the active comparator for this study to be positive.

Whereas, in scenario number four, a result would be positive if the experimental arm has a cure rate that is even two percent less than the active comparator, or a relative 10 percent increase in non-cure would still give a positive result.

It is interesting to compare that to the lenient criterion that you would have in scenario number two for non-inferiority, and in this setting one would achieve positivity even if you had a 6 percent lower cure rate, or a 30 percent relative increase in non-cure, would still yield a positive result.

And it is in these settings where positive results are a conclusion, even when you have a meaningful reduction in the post estimate that lead to concerns about bio-creep. Next slide.

We have heard about bio-creep and the fact that it can arise in repeated non-inferiority trials. Is this a hypothetical that we would have repeated non-inferiority trials?

Well, to give an illustration from last October, the Anti-Viral Drugs Advisory Committee was asked to consider voriconazole as an empiric anti-fungal therapy, and the data that was provided, and the basis for this, was in essence from three generations of studies.

The first generation were control trials of Amphotericin B. The second generation was looking at the liposomal version of Amphotericin B against Amphotericin B.

And then the third generation was looking at voriconazole against the liposomal version. Now, what were some of the complexities that this advisory committee had to face?

The first is that there were control trials of the efficacy of amphotericin B, and the Pizzo study and EORTC studies, did yield evidence that amphotericin B yielded a reduced breakthrough infection rate.

However, the studies were very small, not reliable, and so there is considerable variability or uncertainty in what the level of efficacy would be. Also, this study was done in patients from 15 to 20 years ago.

So there are lots of uncertainties about

the relevance of these data, interpretability of these data, in the context of present day studies.

The second generation study, and pardon the typo here, was done by the Mycosis Study Group, an important study looking at ambisome against amphotericin B.

One issue that was very relevant is that the definition of the end-point in this second generation study was somewhat different than the third generation study, so that ambisome had a very different response rate, a much lower success rate in the third generation study, rather than the second generation study.

The success rate was essentially a composite end point looking at persistent fever, death, and breakthrough fungal infections. Furthermore, it this third generation study, voriconazole was estimated to have a 6 percent lower success rate, with a lower level of the confidence interval of minus 12 percent.

And guided by the proposed use of a margin of minus 10 percent, and many other considerations, the Anti-Viral Advisory Committee voted unanimously against approval of voriconazole in the setting of empiric anti-fungal therapy.

It is interesting to speculate what decisions would have been if more lenient margins of minus 15 percent had been used, and it is also interesting to speculate that if voriconazole became a standard therapy in use, and there was now a fourth generation study looking at a new empiric anti-fungal therapy, what would be the choice of margin that you would use when comparing against voriconazole that would provide a reliable estimate of efficacy or sense of efficacy of that fourth generation agent. Next slide.

In closing, just to highlight a couple of the key conclusions. Non-inferiority trials that use scientifically rigorous margins do not necessarily require very large sample sizes, particularly as we were hearing before if we are developing new agents that we are hoping are better, but aren't so confident that they are so much better that we could provide superiority with high power, but are just modestly better.

If they are just modestly better, we can rule out that they are meaningfully worse without having an inordinately large sample size. And finally as ICH E-10 indicated, the determination of the margin in a non-inferiority trial needs to be based on a wide

array of issues, issues that relate to clinical judgment.

What is the clinical importance of losing a given level of efficacy. That is one key issue, and another key issue is do we expect major important tangible benefits to patients, in terms of safety, tolerability, convenience of administration, resistance, drug interactions, et cetera, that would allow us to give up some margin or some level of efficacy on the primary end-point.

In addition, there are important statistical issues. What is in fact a reliable estimate of the efficacy of the active comparator. If the active comparator is highly effective, with precisely estimated efficacy, where we have assay sensitivity, where we can believe that that estimate of efficacy in the historical trials reliably predict what the efficacy would be in the non-inferiority trials, then we would be able to with confidence have larger margins.

However, as the ICH guideline indicates, to the extent there are uncertainties in these issues, that should influence the size of margin that we are willing to use.

Finally, the question or finally the

comment here is the choice of margins should be suitably conservative. It is certainly the case that we would want to have efficient and timely development of new agents.

But to follow this concept of being conservative, the question arises isn't public health best served by using approval standards that do reliably rule out experimental therapies that do have an inferior benefit to risk profile relative to standard of care. Thanks.

CHAIRMAN RELLER: Questions for Dr. Fleming? Jim.

DR. LEGGETT: In terms of the practicality, from the PhRMA and the other speakers, they talked about the impracticality of having a smaller delta. What about the factors of having a practicality for an agency such as the FDA when you want to factor in the other things that you talked about?

How do you make the hurdle the same for Drug A, Drug B, Drug C, that come into these same designated indications? If Drug A is a much better tolerant, and Drug B you can give once a year, and Drug C -- well, how can you bring those in so that there is one hurdle?

DR. FLEMING: You mean so there is one hurdle for all agents in a class, or for agents across classes?

DR. LEGGETT: How do you determine when a particular drug company wants to present something to the FDA about what kind of numbers they should go for?

DR. FLEMING: Right. Well, what I am arguing here is that there are a myriad of issues that need to be considered, and the actual choice of a margin really should be specific to a given agent and a given indication.

And the ideal time for this is in the planning process for the trial, as opposed to after data are available in the trial. Clearly there is a requirement here for both clinical and statistical judgment, and that clinical judgment I believe needs to take into account the trade-off's between what are the negatives for allowing a loss of a certain level in the primary end point, the primary efficacy end point.

And weighed against what are the perceived or expected benefits that the experimental therapy is going to provide. And if that experimental therapy is providing significant improvements in safety, tolerability, resistance to drug interactions, et

cetera, one, I believe should have a willingness to allow a somewhat larger margin.

If on the other hand we are looking at a new agent that is not anticipated to be any different, then I am arguing that if in fact the efficacy of that is thought to be modestly better, then you can have a rigorous lower limit, or a lower margin, and have very reasonable sample sizes.

On the other hand, if it isn't any better, then admittedly there would be either the need for a larger sample size, or a risk of a false negative conclusion if the new agent truly isn't any better and doesn't provide any tangible benefits relative to standard of care.

CHAIRMAN RELLER: Dr. Bell.

DR. BELL: I am wondering if somebody from the FDA could answer how much leeway does the agency have, either legally or practically, to set different deltas for different -- for the myriad of different considerations, including different drugs for different -- I mean, how uniform do they have to be?

DR. GOLDBERGER: Actually, our last question this afternoon deals with some of these issues about the factors that ought to be taken into account beyond simply delta in making regulatory

decisions.

But to answer your question, products are supposed to be substantial evidence of safety and efficacy. There is in fact a lot of flexibility that can be applied.

I think one of the things that you have heard this morning, and that you will hear again this afternoon, is we have to be satisfied that the drug is more effective than placebo or no treatment would be in that situation.

I mean, that is sort of the minimum standard. Beyond that, there is just a lot of flexibility. It would depend if this is the tenth drug for an indication, and it doesn't appear to be any different, in terms of tolerability, activity, pharmokinetics, et cetera.

And there is not a whole lot of reason to necessarily be that flexible. If on the other hand -- and we have done this in the past, the drug may in fact be less effective than comparator.

And the example that comes to mind is in trials for pneumocystis, where we have in the past approved drugs that were less effective on a mortality end-point than the comparator, because the drugs offered the opportunity to treat patients who could

not otherwise be treated by the comparator, which was trimethethum sulfur.

So that represents a lot of the flexibility, and that we can actually approve a drug that may be worse than comparator, with of course including information in labeling to the point where we would expect a reasonably tight delta in a situation where there might be 10 other drugs, and in fact this drug offers no advantage.

CHAIRMAN RELLER: Dr. Temple.

DR. TEMPLE: The people who wrote the Food, Drug, and Cosmetic Act, made it very clear that they were not trying to impose a relative effectiveness standard.

So for symptomatic treatments, we are interested in whether the drug works at all. It can be less effective than available therapy as long as it is effective.

But when lack of efficacy has important consequences, safety consequences, then the implications are somewhat different. And the very reason that you can't do placebo controlled trials in some pneumonia is the reason why you are not willing to accept too much less effectiveness.

And so there is a complex of judgments

made about how much evidence you need. It is worth remembering that when you have a delta, what you are excluding out is the lower bound of a 95 percent confidence interval.

The exclusion of 10 percent, it doesn't mean that you are likely or it is likely that the drug is 10 percent worse. It is more -- I mean, in fact, the point estimates in general would be right on top of each other.

Which means that it is most likely they are fairly close, and the question then becomes how much risk are we willing to accept that the drug is a little bit worse, and as Tom was saying, and that Mark said, you accept more risk if there is some comparative benefit; greater ease of use, less of an important side effect, and those things.

But in general -- and actually this was all described in a Presidential Proclamation about 3 years ago that I have been trying to find. But what it said was that relative efficacy is not what we do unless lack of efficacy represents a safety consequence.

And then we consider it, and we ask sophisticated advisory committees for help in thinking those questions through.

DR. FLEMING: But just to follow up on what Dr. Temple just said, we talk a lot about margins. They are very important issues. But it is important to understand for any given margin what does this really mean, the point estimate has to be in order for you to satisfy the criterion of non-inferiority.

And where I worry is when we are choosing margins so large that the point estimate can be substantially less or substantially negative, substantially less favorable for the experimental, versus the active comparator, and still be viewed to be a positive result.

That's the setting that leads to this risk of bio-creep.

CHAIRMAN RELLER: Dr. Bennett.

DR. BENNETT: Could I ask Dr. Temple about the power function in selecting or estimating sample size? I think I heard Dr. Shlaes said that the examples that the FDA was giving, you are using a power of .8, but that PhRMA would find that unacceptable because of the possibility of accepting too many ineffective drugs.

Is it true in your experience that PhRMA generally insists on a power of .9 in estimating

sample size?

DR. TEMPLE: Well, Tom has probably helped a lot more companies figuring out what power they should use than we have.

My experience is that in many settings -- for example, in different show and trials, that companies often do use a power of something like 80 percent.

And perhaps because they are going to do multiple trials and figure that it will work out all right. But nobody wants to have a substantial chance of losing.

So I think a tendency towards getting the best power you can manage is certainly there. What I would say we find more -- and this again applies mostly to different show and trials, is an estimate of the effect size that is optimistic.

So if you estimate that you are going to have 50 percent effect on something, well, then your power looks terrific, even in a modest sized study. And where failures occur is where people have been over-optimistic, and not realistic, and haven't done a large enough trial.

In the setting or in these settings, the fear would be that you are going to come out a little

bit worse for your point estimate, and therefore, will not be able to exclude the margin that you are talking about. And I would think companies would worry about that.

CHAIRMAN RELLER: Dr. Shlaes.

DR. SHLAES: Just to clarify. I think what I said was that if you do an 80 percent power at a 10 percent delta, and that sort of study, then you have a 32 percent chance of falsely concluding inferiority based on these set point considerations.

I think that is what I was trying to say, and so that most companies wouldn't do a 10 percent delta trial powered at 80 percent.

In the old step function, obviously many trials were done at 20 percent, or 15 percent deltas, and then you can tolerate a risk of an 80 percent power because your chance of falsely concluding inferiority is lower.

CHAIRMAN RELLER: Dr. Glode.

DR. GLODE: I was just going to mention that I brought with me to this meeting, because I thought it was very informative and Dr. Fleming just mentioned it, which is the article published in the January 24th New England Journal of Medicine, on voriconazole compared to ambisome.

And where in the discussion it mentions exactly the conclusion that you mentioned, that it fails the test of non-inferiority. However, in the abstract of the article and in the conclusion that is never mentioned, but rather that it is a suitable alternative to amphotericin B preparation.

Now, there is a lot in this article to explain that conclusion, but it still brings up the complexity of selecting the appropriate end point. Anyway, that is a good example.

CHAIRMAN RELLER: Thank you. It is time for lunch. Let's reconvene promptly at 1:15, and not one o'clock. We will pick up the time probably during the public hearing.

A reminder. There are 30 seats set aside in the restaurant reserved for committee members to enable people to get back at 1:15. And also the discussions about the issues that we have addressed should be kept in the public arena here and not outside of this public arena. Thank you.

(Whereupon, at 12:22 p.m., a luncheon recess was taken.)

A-F-T-E-R-N-O-O-N S-E-S-S-I-O-N

(1:24 p.m.)

CHAIRMAN RELLER: I would like to open this afternoon's component of our Advisory Committee Meeting and ask for the Open Public hearing. We have one scheduled speaker, Dr. Kem Phillips, from Advanced Biologics. Dr. Phillips.

DR. PHILLIPS: I am Kem Phillips from Advanced Biologics. We, meaning myself and Dr. Michael Corrado, submitted a paper to the committee, and we thought this was going to be a kind of stealth paper that would go under everybody else's radar right into their laps.

But apparently if you do this, it has to get presented, and so to save time from actually having to read this thing to you, I will give a brief presentation. I am just hoping that the lion isn't looking for desert here.

Our paper was titled, "Should the Non-Inferiority Margin Vary With the Comparator Rate." There were a lot of good presentations this morning on the clinical issues involved in this issue.

And some of the things that came up were that you would have a difficult time establishing a comparator rate, because for one thing, you might have

an increase in resistance.

You might have difficult indications, or you might have new designs. For example, one design for a drug that only targets GRAM positive organisms. So all of these lead to an inability to predict response rates.

In some cases, you might have a good rate, a well-established rate , and you wouldn't have a problem. But if you can't, you have a difficulty, and for us statisticians, the question is how to set the sample size.

Drs. Lin, Brittain, and Fleming discussed statistics earlier today, and did an excellent job, and I don't have anything to add to what they have said about a fixed delta method.

But how are you going to set that delta when you can't predict the success rates? And as they have said several times, if you have a 10 percent delta and a 70 percent underlying rate, you need 330 patients.

And if it is a 90 percent underlying rate, then you need 142. So that is a big disparity. The points to consider had one main feature that has been discussed a little bit, and that is that based on observed rates.

You would set the delta to be 10, 15, or 20 percent. Now, one of the things that I don't think did get discussed is this issue of the observed rates. Any many of us would interpret that as meaning if you observe in your trial, say, an 85 percent rate, then you would in the better of the two arms, then you would use a 15 percent delta and so forth.

That leads to sort of an odd test, and among other people, Rohmel, in a '98 Statistics in Medicine paper, outlined some of the problems with that procedure.

The main thing that comes up is this. We have seen before where we have this discontinuities at 80 percent and 90 percent. So, for example, if you observe a 91 percent success rate in your trial, and maybe you wished it was an 89 percent so you could use the 15 percent delta, and various other things happened.

So Rohmel says -- and it discusses a little bit about the possibility of adapting delta to the observed rates, and he says that there were two criteria.

One, there should be good reasons, clinically and statistically, for the non-inferiority margin should vary with the response rate of the

standard drug, or the better of the two.

And, number two, the boundary curve of the equivalence margin should be smooth. The standard approach takes a null hypothesis that the test rate be at least the comparator rate, minus delta, and T is greater than C minus delta.

And in that case, we get these various characteristics that we have seen. And you will notice that C minus delta is a linear function of the comparator rates, C. So why not think of it as being a more general linear function, A times C, plus B.

And if you do that, you can actually establish a valid test, and it doesn't have these problems that you have with the points to consider procedure.

You could even fit that linear function to the points to consider deltas, and get some approximates very clearly, but it still has good statistical properties.

Another thing you can get out of this test is by setting these parameters A and B appropriately, and you can get something that satisfies something you might call the Lewis criteria.

Rohmel quotes J.A. Lewis as saying that you might adopt the equivalence margin in such a way

that the response rate of the better of the two agents that the power of the study remains constant over a wide range of potential response rates, and is thus independent of the later observed response rates.

And you can set these parameters of this more general test to be able to do that. So this again is a valid statistical test, and it approximates the points to consider or some other set of criterion that you might like.

But one main problem with it that came up, and I believe that Dr. Fleming mentioned briefly this morning, is that at least if you look at the ITT population, if you get worse success rates, and perhaps intentionally, because you are getting bigger deltas with lower success rates, you might actually increase your probability of showing equivalence bogusly.

But in the evaluable population, you are probably throwing those cases out anyway. So that probably isn't so much of a problem. So, anyway, that is all that we wanted to say, that we believe that it might be a good idea to be able to adapt delta to the comparative rates, and that we do have a valid statistical test for doing that.

CHAIRMAN RELLER: Are there any questions

for Dr. Phillips or comments on this approach?

(No audible response.)

CHAIRMAN RELLER: Were there other persons who wish to present at the open public hearing? If not, we will move to the FDA's presentations. First, Dr. John Powers, who is a Medical Officer with the Division of Special Pathogen and Immunologic Drug Products at FDA, who will present a medical perspective on hospital-acquired pneumonia and meningitis. John.

DR. POWERS: Okay. We're on. Thank you, Dr. Reller. This afternoon, we would like to give two presentations, the first of which will be mine, looking at two serious diseases with high mortality rates, and that is acute bacterial meningitis and hospital-acquired pneumonia.

And then after my talk, Dr. Susan Thompson will present some similar information on a less severe disease, acute bacterial exacerbations of chronic bronchitis.

And our goal with these two talks is actually to try to give you a framework to hang some of these principles on that we have talked about earlier this morning.

So what I would like to do first off is to

reiterate what the definition of delta is, and its various components, and then talk about the impact of deltas in the clinical setting, and what it means to patients.

And then we will go through the selection of delta, or some of the issues in the selection of delta, looking at the two components that were explained this morning, the delta one, or the historical sensitivity to drug effects in acute bacterial meningitis and hospital-acquired pneumonia.

And we will look at that by examining some information from the pre-antibiotic era, and from the antibiotic era, to try to get a feel for what is the magnitude of the benefit for antibiotic therapy in these two indications.

And also talk about what are some of the confounders in determining the efficacy of control regimens in these particular diseases. Then we will talk about the issues of delta two, or that judgment related issue of acceptable loss in these two diseases, by focusing on what are the consequences of less effective therapy in these two diseases.

And then finally finish up with some of the practical issues in selecting deltas. It is important I think to start with an idea of what is the

purpose of a clinical trial in the first place.

And a clinical trial is supposed to distinguish the effects of a drug from other influences, such as spontaneous change in the course of the disease, placebo effect, or biased observations.

One could ask the question, well, why can't clinicians just do this on their own once the drug gets into common usage. And it actually can be quite difficult for clinicians to make judgments on the efficacy and safety of a drug outside of the setting of a clinical trial, and there are several reasons for this.

In a disease that has a high spontaneous cure rate, if a patient receives antibiotic X or Y, they may get better anyway, regardless of which drug they get, and it may actually be impossible to discern an ineffective therapy given that most patients will resolve spontaneously.

Also in diseases that are more serious, and that have high mortality rates, at least in today's realm, most of those people have serious underlying diseases which can be a confounding factor.

So if a patient dies on therapy, is that because of their underlying disease, or was it because

of progression of that infectious disease, and that can be quite difficult to tell, even with autopsy data that can sometimes be hard to tell what the patient died from.

And finally it can also be very difficult to tell what the safety of a drug is compared to another drug just in the clinical realm. If you give your patient a particular drug, and they get a rash, that is pretty clear.

But the real question is how does that compare to another therapy, and what is the rate of rash in a controlled regime, and it is really hard to do that outside of the setting of a clinical trial.

And just to add a point. This morning we were talking about antibiotics and their ability to eradicate bacteria. Some would also argue that antibiotics also have other effects.

And as Dr. McCracken mentioned, some antibiotics have anti-inflammatory effects, or sometimes they go in the opposite direction. And there is actually some in vitro data with amphotericin B that says that if you incubate amphotericin B with white cells, that it releases massive amounts of tumor necrosis factor.

Whether this has an impact on clinical

outcomes or not really isn't clear, and hasn't been studied. The other reason for clinical trials is that sometimes we see a result that just wouldn't be intuitive based on what we would think going into the trial.

And probably one of the best examples of this is clarithromycin studied in the treatment of disseminated microbacterium avian disease in AIDS patients. And in that trial, there were three doses tested; a low, an intermediate, and a high dose.

And in that trial the low dose had no effect on eradication of MAC. The moderate dose did have an effect, and actually the mortality was higher in the high dose than it was in the moderate dose.

And one would never have guessed that going into the trial based on the pre-clinical data. So sometimes we see results from clinical trials that we just wouldn't predict from some of the preclinical information.

And in non-inferiority trials -- and again, Dr. Fleming said this as well -- we are attempting to prove that the test drug is not inferior to the control drug by some margin, and we can't prove that two drugs are absolutely statistically identical in efficacy.

So we need some way to estimate the variability around the difference between those two treatments. And the way we do this is again looking at the non-inferiority margin or delta, which we are defining as the maximum degree of inferiority of the test drug, compared to the control drug the trial attempts to exclude statistically.

And again this is specified prior to initiation of the trial. Once the trial is over, we calculate the difference in the point estimates of the efficacy of the test agent, minus the control agent, and again I am using the convention that Drs. Brittain, Lin, and Fleming used.

Dr. Temple used the opposite of this, but I am using the test agent, minus the control agent. And here on this slide, we can see just as an example, I am showing that the point estimate of the difference of the test minus the control agent is minus 8 percent.

We then calculate 95 percent confidence intervals around the difference in the point estimate, which gives us some idea of the variability around this estimate.

And then we compare the lower bound of the 95 percent confidence interval to this pre-specified

non-inferiority margin, which in this example is minus 15 percent.

So again just to reiterate what you heard this morning, since we are all sleepy after lunch, delta-1 is a conservative estimate of the advantage of active control over placebo that is based on data.

Delta-2 is the largest clinically acceptable difference between the active control and the experimental drug, which is based on judgment. And again that judgment is in-turn based on what are the consequences to patrons of treatment failure.

So overall selecting a delta for the clinical trial, if the delta-1 is very large, or in other words, is there is a huge benefit of drug treatment over placebo, then what really matters is selecting the delta based on the delta-2.

So if we then go on to talk about delta-1, which is historically-based data, we can ask the question do we really know what we think we know about the historical information.

And again the important point to remember here is that it is not whether an antibiotic actually helps patients or not. It is what is the magnitude of that benefit, and when one actually goes through the literature, trying to tack a number on to this, it can

be actually quite a daunting task, I can tell you, having spent hours in the library looking this stuff up.

So one of the problems is that for some diseases that we deal with, there is no data from the pre-antibiotic era. These are really diseases of modern medical care in some cases.

The second thing is that there has been changes in the resistance patterns of the common organisms causing these diseases, and also the epidemiology of the disease itself.

Thirdly, there can be differing response rates in various sub-populations with the disease. Fourthly, there can be changes in the practice of medicine, or supportive care, of patients with that disease.

And then also there can be problems in defining patients who actually have bacterial infections, versus either non-bacterial causes of the same kind of infection, or non-infectious causes that may mimic that disease.

And finally a point that was brought up several times this morning, is that sometimes we use different definitions of success and failure in our current trials, compared to the end point in pre-

antibiotic trials were, which is mostly mortality for the main part.

The delta-2 is the judgment based acceptable loss relative to current therapy. In an ideal world, one could make the assumption that for more severe diseases one would like to see a smaller delta, because the consequence of treatment failure in those severe diseases could be increased morbidity and mortality to patients.

On the other hand, in less severe diseases, one would be tempted to accept a larger delta because even though there may be greater loss relative to current therapy, that may not translate into mortality for patients, although it may translate into more morbidity and discomfort for patients.

But unfortunately we don't live in an ideal world, and there are practicalities of performing clinical trials that we need to take into account when forming our judgments about what is an acceptable loss.

And this is what we are going to do for you this afternoon hopefully, is that we are going to take these three diseases, and try to go through them, and show you some of the information that you can hang this around.

The first that we will talk about is acute bacterial meningitis. Well, the delta-1 for acute bacterial meningitis, the magnitude of advantage over placebo is well known in acute bacterial meningitis.

There is data from the pre-antibiotic era, and it is a very large benefit. Therefore, the decision should be based on that acceptable loss, and taking into account the difficulty in doing trials, as well as the fact that we may increase mortality by accepting drugs that are less effective.

The second indication that we will talk about is hospital-acquired pneumonia. And actually this is a disease more of the modern era, where the magnitude of the advantage over placebo is not as clear, and when you actually try to hang a number on this, it becomes quite difficult.

And then again you are still left with that decision on what is an acceptable loss. And then finally after me, Dr. Thompson will go over acute bacterial exacerbations of chronic bronchitis, where the advantage over placebo is unclear, and may in fact be quite small.

Or it may be different, depending upon which subpopulation you are dealing with, and the decision on acceptable loss here is not as critical,

again because we are not dealing with high mortality rates.

So let's start off looking at these components of delta for meningitis and hospital-acquired pneumonia, and I have divided this up by asking several important questions for each of the delta-1 and the delta-2 components.

For delta-1, one can ask the important question of what is the magnitude of benefit of any antibiotic therapy over placebo. The second question is, is the benefit of antimicrobial therapy in current trials measured in the same way as in the original trials showing that benefit.

And the third question is, is the magnitude of benefit of therapy over placebo, or the delta-1, large enough that it should not effect the selection of the overall delta for the clinical trial.

In other words, we can skip the delta-1 altogether and make a decision on the delta for the trial based on delta-2. The important question for delta-2 is what is an acceptable loss of efficacy compared to accepted therapy in a serious disease, and there are two sides to this coin.

The first is the scientific considerations of what happens to patients who fail treatment in

various patient subsets with meningitis or hospital acquired-pneumonia.

And then what you heard a lot about this morning are the practical considerations of the effects of changing the delta on sample size as the efficacy rate changes.

Well, let's look at acute bacterial meningitis first, and try to figure out some information about delta one, or the historical sensitivity to drug effects in this disease.

Clearly, acute bacterial meningitis was highly lethal in the pre-antibiotic era. The most common organism before antibiotics was actually meningococcal disease, which occurred in large outbreaks.

And the overall mortality in these outbreaks was somewhere between 70 and 90 percent without specific therapy, and there are articles about the 1905-1906 meningococcal outbreak in New York City, which clearly defined this number for us.

The other interesting point is that those outbreaks occurred in mostly previously healthy young people, who were in crowded conditions, and who then went on to get ill. So they did not have underlying serious diseases.

When Flexner first studied anti-meningococcal serum in this paper published in 1913, it decreased the mortality in meningococcal meningitis from 70 percent to 30 percent. So, clearly a very large mortality benefit, even with meningococcal serum.

And then finally Schwenker published his paper in 1937, which gave sulfanilamide, given both subcutaneously and intrathecally to 11 patients, and this reduced the mortality to 10 percent.

And in this series, he treated 11 patients, and 9 of those 11 patients survived. One of the patients who did die actually had bacterial eradication from his spinal fluid, but went on to pass away anyway.

What are some of the problems with this historical data? Well, we use different end points in current clinical trials, and although mortality is one of the end points that we still look at, we can argue that sometimes that is not that high, and doesn't drive the overall end points.

For instance, in the trovafloxacin study that was published in Pediatric and Infectious Diseases that Dr. McCracken talked about this morning, the mortality in each group was 2 percent and 3

percent, and clearly different than what we saw in the pre-antibiotic era.

So some of the end points that we look at here, in addition to mortality, are developmental, neurologic, and audiologic sequelae. It is hard to get a handle on what the effect of antibiotics is on these, because if patients didn't get treated, they die. So it is hard to tell.

There is also different epidemiology today than we saw in the past, and today pneumococcal meningitis is the most common form of bacterial meningitis in the United States, and that is even different from 10 years ago in this country.

And finally there are different populations. In this study that was published a few years ago in the New England Journal of Medicine, it compared the epidemiology of acute bacterial meningitis in 1995, to the epidemiology in 1986, and showed that in 1986 that the average age of a meningitis patient in the U.S. was 15 months.

And the average age of a meningitis patient in 1995 was 25 years, a huge difference in the epidemiology, even over a short span of time. Now, let's switch gears, and try to look at the historical data for hospital acquired pneumonia.

It is a much more difficult task, because the clinical entity of hospital acquired pneumonia was not described in the pre-antibiotic era. If we tried to look at some of the organisms implicated in hospital-acquired pneumonia, even though they aren't acquired in the hospital in this pre-antibiotic data, we can see that in the influenza outbreak in 1918, there were a number of cases of post-influenza Staph aureus pneumonia.

And in one report, there were only two spontaneous cures out of 151 cases on a military base with Staph aureus pneumonia. So, clearly a highly lethal disease.

There were very few reports in the pre-antibiotic area of Gram-negative pneumoniaes, and again part of the problem with these reports though is how certain are we of the microbiologic diagnosis in these case reports.

So really there is no way to compare antibiotic therapy to placebo for hospital acquired pneumonia, because these studies just don't exist. So what we are left doing is trying to extrapolate data from the antibiotic era to see if we can find what the placebo rate would be.

Well, one way to try to do this is to

compare patients that get appropriate antibiotic therapy to inappropriate antibiotic therapy, and I am going to contrast these two studies to show you how difficult a task this actually can be.

If we look at this study by Celis that was published in Chest, they looked at all-cause mortality in patients that received appropriate antibiotics, versus those who received inappropriate antibiotics.

In this trial, appropriate antibiotics were defined as an organism that was sensitive to the antibiotics that the patient received. And again obviously you can't randomize patients to get inappropriate therapy, and so this is an observational study.

The all-cause mortality rate in patients that received inappropriate therapy was 91.6 percent, and the all-cause mortality in patients that received appropriate therapy was 30.5 percent. So a 60 percent difference between appropriate and inappropriate therapy.

There is a lot of problems with this data, however. The first is that obviously it is an observational study, and the second is that the number of patients that received inappropriate therapy was very small in this particular trial.

So if we attempt to look at another study that was done almost 10 years later, published by Alvarez and Lerman in Intensive Care Medicine in 1996. These people looked at this question in a slightly different way, but it tremendously changes the numbers.

They again looked at inappropriate versus appropriate antibiotics, but this time they defined inappropriate therapy as lack of clinical improvement, or an organism that was not sensitive to the antibiotic that the patient received.

So there was more than one way to define appropriate, versus inappropriate. They also looked at attributable mortality. In other words, assuming that the patient died, they died of pneumonia.

Now, how one determines this isn't clear from this paper, and it is not clear in any case how one would decide what the patient died of. So in this case, they looked at the attributable mortality to hospital-acquired pneumonia.

And comparing appropriate to inappropriate therapy. If the patients received appropriate therapy, the mortality rate was 16.2 percent, and if they received inappropriate therapy, the mortality rate was 24.7 percent.

So only about an 8-1/2 percent difference here. Now, again, there are differences in the populations between these two studies. The Celis study enrolled only mechanically ventilated patients in the ICU.

The Alvarez and Lerma study enrolled patients in the ICU, 60 percent of whom were on mechanical ventilation, but the other 40 percent were not. This is the kind of data that you have to deal with when you are trying to decide what is the effect of antibiotics.

And this is as good as it gets. So it is very difficult to find out. Again, there is also problems with this historical data. There is a great difficulty in the clinical diagnosis of hospital-acquired pneumonia, and several studies that look at this show that clinicians are only correct in their diagnosis of hospital-acquired pneumonia, at least based on autopsy studies, about 50 percent of the time.

The problem with this is that patients get enrolled in these studies that don't have the disease. So you can't expect the antibiotics to have an effect on someone that doesn't have an infection.

Also, there has been a change in

nosocomial organisms over time, with a shift from GRAM-positive organisms back in the 1950s, with the introduction of positive pressure ventilation, to GRAM-negatives and back to GRAM-positives again today.

There is also very different outcomes in various patient populations. The mortality rate in mechanically ventilated patients is much higher than that in, say, ward patients or ICU patients who are not ventilated.

And again there is the problem of how do we attribute the death to pneumonia versus all-cause mortality, and even at autopsy, it can be difficult to discern this information.

And then finally we use clinical end-points other than mortality in our current clinical trials; things such as normalization of the white blood cell count, and resolution of a chest radiograph, or resolution of fever.

So if we then go back to our original questions, and again shifting gears back again to acute bacterial meningitis, let's see if we can answer some of these questions.

For delta-1 for acute bacterial meningitis, what is the magnitude of benefit of antibiotic therapy over placebo. Well, it appears

that this is pretty clear, and it is as large as 60 to 80 percent mortality benefit.

But the magnitude of benefit on clinical parameters, such as auditory, hearing, neurologic, developmental losses, is not as clear. Is the benefit of antimicrobial therapy in current trials measured in the same way as in the original trials?

Well, yes, and no. We still use mortality as an end-point, but we do use the other end-points of auditory and neurologic developmental losses as well.

And, thirdly, is the magnitude of benefit of therapy over placebo large enough that it should not affect the selection of the overall delta for a trial. And the answer here appears to be yes, because again the magnitude of the benefit is so large that you can select the delta based on the considerations about clinical loss.

How about for hospital-acquired pneumonia if we attempt to answer these same three questions. What is the magnitude of benefit of antibiotic therapy over placebo? Much harder to answer than for bacterial meningitis.

And based on the two trials that I have presented to you, the benefit can be anywhere from 8-1/2 percent to 60 percent, depending upon how, and

in whom this benefit is measured.

And it is very unclear what the benefit of antibiotics is on a resolution of clinical parameters, such as fever, white count, and chest radiograph. The second question is the benefit of antimicrobial therapy in current trials measured in the same way as in the original trials showing benefit? Again, the answer is yes and no.

We still look at mortality, but again we are looking at the resolution of those clinical parameters, as well as part of the primary end points. And then finally is the magnitude of benefit of therapy over placebo large enough that it should not effect the selection of the overall delta for the trial.

Well, this is one of the things that we want the Committee's help on today. Given the problems in looking at this trials, how is one to decide what the acceptable loss is given some of the practical considerations as well.

The other point that I want to make about hospital-acquired pneumonia referable to some of the discussions that went on this morning, is that there is a clear difference about what the bacteriology means in a disease like acute bacterial meningitis,

versus hospital-acquired pneumonia.

And we talked a little bit this morning about using so-called hard end points of the microbiology of some of these diseases. Well, that may be appropriate for acute bacterial meningitis, where you have sterile body fluids, such as cerebral spinal fluid, where you can measure an effect of the antibiotic.

That becomes very problematic for hospital-acquired pneumonia, and in fact a number of the other respiratory indications, where the organism that you isolate in the sputum may have absolutely nothing to do with the patient's clinical course.

And the flip side of that is that you can find organisms in the patient's blood stream when their sputum sterile. So the microbiology in a disease like hospital-acquired pneumonia becomes very difficult to interpret.

And we would like to hear what the committee has to say about that as well. Finally, for delta-2, we need to talk about both the scientific and the practical considerations of selecting delta-2. Well, again this is based on the consequences to patients of treatment failure.

In meningitis, there is a clear

consequence of treatment failure, and that is death. So there is a clear mortality benefit of antibiotic therapy, and the morbidity here is developmental, neurologic, and audiologic sequelae.

And again it is unclear what the magnitude of benefit of antibiotics for those things actually is. For hospital-acquired pneumonia, well, while there may be a mortality difference as one of the consequences of failure, although again the magnitude of that benefit varies depending upon how and in whom that is measured.

And also there can be a morbidity increase, and clearly there are studies which show that patients who do not get treated appropriately for hospital-acquired pneumonia have an increased cost of their hospital stay, and an increased duration of their hospital stay as well.

But again although we have that economic information, there really is a lack of information on the effect on the rate of clinical resolution of things like the white count fever and chest radiograph.

So finally, and you have heard a lot about this this morning, and so I won't spend much time talking about it, are the practical issues involved in

selecting delta.

And the effect of the success rate on delta you have heard a lot about this morning. But there is also something that goes into this beyond just sheer economics, and that is how many patients actually have the disease.

So we need to look at the epidemiology of the disease, the limitations of the inclusion and exclusion criteria of a trial, and the inability of patients to continue on randomized therapy in studies of very severe diseases, where patients may not make it to the end of treatment.

You have seen this slide a couple of times today, and I am not going to go through it in detail, and I will just show you that what I really want to point out is that you can see the relationship between delta and success rate is not linear.

As you tighten the delta the number of patients required in a trial goes up rather steeply. So let's talk about he epidemiology of the diseases and what we know.

And you heard a little bit about this from Dr. McCracken this morning, and again this is based on this information obtained from 248 cases of meningitis acquired by the CDC and published in this New England

Journal paper in 1997 from data from 1995.

Well, what we used to see in 1986 was that haemophilus influenzae was the number one cause of bacterial meningitis, and it occurred in children at an average age of 15 months.

What we see now is that streptococcus pneumoniae is the most common organism at one 1.1 cases per hundred-thousand patients, and haemophilus influenzae has dropped all the way down into a tie for fourth place with listerial meningitis.

Why is this important? This is important because the case fatality rates are obviously going to influence the cure rate in the disease, and this varies by organism.

Haemophilus influenzae has a lower case fatality rate than disease caused by streptococcus pneumoniae. If one were to do a trial in the United States today, you would most likely get more streptococcus pneumoniae isolates, but that would also mean that the mortality would be higher.

So if you compared a trial done today with a trial done in the 1980s, the overall cure rate may be lower now because you are having more strep pneumo cases than you did haemophilus influenzae.

This paper also estimated the number of

cases in the United States in 1986 and 1995 of acute bacterial meningitis. And it was estimated that there were about 13,000 cases in 1986, and now we are down to less than 6,000 cases in 1995.

And Dr. McCracken mentioned this morning that another organism may come along to replace this, and this study actually looked at the difference here, and it really is due to the huge drop in haemophilus influenzae Type B disease, and it has not been replaced by something else, at least not to this point.

So we have a shrinking number of cases in this country as well. Switching gears once again back to hospital-acquired pneumonia. Well, just like everything else with this disease, it is unclear what the epidemiology of this disease is. It is not a reportable illness.

The National Nosocomial Infection Surveillance data estimates that there is about 250,000 cases per year in the United States, but this uses a clinical definition of hospital acquired pneumonia.

And even though hospital acquired pneumonia may account for one percent of all patients entering the hospital, and it is the second most

common nosocomial infection after urinary tract infections, and the most common infection in the ICU, it still ends up being relatively uncommon compared to some other diseases.

And again these may not be entirely accurate, because I pulled these from a number of different sources. But I just wanted to put these as a framework for you to see how things fall out.

Acute otitis media, 26 million cases a year; acute sinusitis, 23 million; and then tonsillitis/pharyngitis, 21 million; community-acquired pneumonia, about 4 million; and then we drop off down here to 250,000 cases of hospital-acquired pneumonia; 10,000 cases of acute bacterial meningitis; and somewhere less than that for acute bacterial endocarditis.

So still these things are relatively uncommon compared to some of the other ones. Getting back to that point about using bacteriologic end points. Again, it depends upon what indication you are talking about.

It may work for acute otitis media, and won't work for acute sinusitis, because we don't get puncture studies most of the time, although we do on occasion.

It won't work for community-acquired pneumonia, and it won't work for hospital-acquired pneumonia. But it may work for acute bacterial meningitis.

So it depends upon the indication whether bacteriology is helpful to us or not. So some other practical points. The success rate in recent hospital acquired pneumonia trials with piperacillin, tazobactam, linezolid, ciprofloxacin, or trovafloxacin, have all been in the 50 to 70 percent range.

If one uses a smaller delta for those trials, the delta used in those trials was 20 percent by the way. But if one would use a smaller delta than that, the sample size would go up.

However, the downside of accepting a larger delta is that theoretically a new drug could then be as much as 20 percent less effective than the comparator. And if we are talking about a drug that already starts off with a 50 percent cure rate, we are down to possibly accepting a drug with a 30 percent cure rate.

The other problem is that almost half of the patients don't complete the trials, and you need to take that into account when looking at the sample

size.

So if we just look again at the left side of this graph, which you have seen many times, if we go from a 20 percent delta, we go from a trial that needs 99 patients per arm -- and again this is assuming 80 percent power.

But if we tighten it all the way down to a 5 percent delta, we are talking about fifteen hundred patients per arm, or 3,000 patients in the study. But that is before you figure out that half of those people drop out of the trial. So you are talking about 6,000 patients per study here.

So then some of the things that we need to take into account for delta-2 to answer that question of what is an acceptable loss of efficacy compared to accepted therapy in a serious disease.

Well, the serious nature of meningitis and hospital-acquired pneumonia would seem to call for a selection of small deltas. However, as we have seen, smaller deltas would result in a larger sample size of the trials, and one of the things that we would ask the committee about today is whether this is practical given what we know.

But we need to balance this risk of accepting drugs, which may be 20 percent less

effective than currently approved therapy. And again if we are talking about a 50 or 60 percent cure rate, 20 percent less than that is a 30 or 40 percent cure rate.

So the dilemma that we are left with here today is to balance this risk to patients of accepting larger deltas, especially in more severe diseases, versus those realities of performing clinical trials.

At this point, I will turn it over to Dr. Susan Thompson, and she will talk to you about acute bacterial exacerbations of chronic bronchitis.

DR. THOMPSON: Good afternoon. I am going to be speaking with you today about the selection of delta in clinical trials of antimicrobial therapy for the indication of acute exacerbation of chronic bronchitis.

The outline of what we are going to be talking about today is given here. First of all, we will give a definition of the scope of the problem, and discuss the selection of deltas specifically for AECB trials.

Then we will spend most of our time reviewing the trials available in the literature which our placebo controlled for the indication of AECB, and discuss some of the confounding issues and

interpretation of those trials.

And we will give you some conclusions and list for you what we feel are unresolved issues, and alternatives for future AECB trials. There are approximately 12 million cases of chronic bronchitis per year in the United States.

And it is the most common category of chronic obstructive pulmonary disease. Most cases of chronic bronchitis are due to tobacco use, and most studies put it in the range of 85 to 90 percent. A few cases are due to environmental pollutants, or such genetic factors as alpha-1 antitrypsin deficiency.

It is important to recall that AECB is a distinct clinical entity from acute bronchitis. Acute bronchitis is usually defined as sputum production in the absence of underlying lung disease, and the vast majority of these cases have viral etiology as the cause.

The Division of Anti-Infectives no longer recognizes acute bronchitis as an indication for which new drugs can apply. Acute exacerbation of chronic bronchitis accounts for 5 to 10 percent of all antibiotic prescriptions in the United States.

Currently, 17 antibiotics, plus or minus one, carry the indication of acute exacerbation of

chronic bronchitis and are labeled, and were approved via non-inferiority trials.

Some of the older antibiotics carry broader indications which were granted at those times, including either upper or lower respiratory tract infections.

I have borrowed this slide from the CDC basically to just give you an idea of the proportion which bronchitis represents in outpatient antimicrobial therapy usage in the United States.

This slide is from 1992, although i suspect that the proportions have not changed. Bronchitis, as you can see, represents 16.3 million courses of antibiotics in the year of 1992, a significant proportion.

It is important to note this slide was presented in the context of a discussion of the antimicrobial resistance, and clearly some of those prescriptions that were written for bronchitis, as well as some of these other diagnoses which are given for outpatient or for respiratory infections, are given sometimes for indications which don't require antibiotics.

Moving then into a definition of acute exacerbation of bronchitis, a fairly standard

definition of chronic bronchitis itself is cough and sputum production on most days for greater or equal to three months in two consecutive years.

And acute exacerbation of chronic bronchitis is some combination of worsening dyspnea, increased sputum volume, and/or increase in sputum purulence.

The etiology is most commonly nontypable H. flu, which usually encompasses 50 to 60 percent of the isolates in most studies. M. catarrhalis is 15 to 20 percent, and Strep pneumo is 15 to 20 percent. The smaller number of atypicals has been found in various studies.

Moving then specifically to the issue of selection of delta for clinical trials, I will reiterate what you have heard many times today already.

Delta-1 is the smallest effect size, if any, that active drugs would be reliably expected to have compared with placebo, and we will spend the majority of our time on that for this indication.

Delta-2 is the largest clinically acceptable lots in efficacy between the experimental drugs and the active drugs, with the smaller of these two values representing delta.

For acute exacerbation of chronic bronchitis then, specifically the determination of delta-1 represents the estimation of the benefit, if any, of active control over placebo.

The determination of delta-2 for AECB is in a sense relatively less pressing, in that AECB has a very low mortality and morbidity, and for this indication than, delta-2 is relatively large and certainly greater than 20 percent.

Thus, for AECB, the smaller of the two values, delta-1 would represent the delta for the studies. Actually, I should have entitled this slide "Previous FDA Guidance for AECB."

The points to consider you are probably all aware of. From 1990, two recommended trials for AECB, or one if the drug was submitted for CAP or HAP. The organisms we have already mentioned.

And 10 to 20 percent was the usual delta for AECB trials based on the efficacy rates which were usually found. The approach then to determine delta-1 for AECB is essentially to review the results of the placebo controlled trials that are available to us from the literature in an attempt to determine delta-1.

The two points that I think are important

to remember during our subsequent discussion is that, first of all, in the past 40 years, less than eleven hundred patients have been enrolled in randomized placebo controlled trials of the antibiotic treatment of AECBs, and none of those trials were of identical design.

The second point that I want you to remember is actually a list of caveats that many of these trials share. First of all is the uncertainty in the definition of acute exacerbation. The second and very important caveat is the lack of consistent and a reproducible rating system for severity of the presentation of disease.

Third is a lack of standard outcome measures, and you will see quickly that this becomes a problem in interpretation of these trials. And lastly, and probably least important, is the role for non-physiologic outcomes.

I've chose to discuss in detail this trial, which was published in the Annals of Internal Medicine from the University of Manitoba in Winnipeg. It is probably the most widely quoted placebo control trial of AECB in the literature.

These authors looked at 362 exacerbations in 173 patients with AECB. These patients were

randomized to receive either a placebo or antibiotics. The antibiotics could be any one of Bactrim, amoxicillin, or doxycycline, depending on the investigator's discretion.

Patients could be treated also for a subsequent exacerbation, in which case they received the opposite treatment, placebo or antibiotics. Success in this trial was defined as symptom resolution within 21 days, and of note most of these patients had -- excuse me, all of them had a low FEV-1.

These authors did use a severity scale in this trial, and it has been referred to as the Winnipeg criteria. Type-1 are the most severely affected patients, and are patients who presented with cough, increased sputum production, and purulence.

Type-2 patients would have 2 or 3 of these symptoms, and Type-3, only one, with one of the listed, fairly non-specific, indicators of infection. This chart basically goes through the results of the trial, and I will walk you through it.

On the left side of the slide are placebo results, and on the right are antibiotic results, and the results are given in terms of either success or deterioration.

The numbers are given as percentages, with the absolute numbers in parentheses. I will direct your first to the overall results of the study, which demonstrated that 55 percent of patients who received placebo had a successful outcome, and 68 percent of those who had antibiotics had a successful outcome.

The results were more impressive when it was divided by the severity of the infection. You will recall that Type-1 were those more severely infected, and in this case 43 percent who received placebo were successfully treated, versus almost 63 percent who received the antibiotics.

The other thing that I wanted to point out to you on this slide was that the deteriorations tracked in the direction that you might expect. Again, those who were more severely infected at presentation had a higher deterioration rate when they received placebo than when they received antibiotics.

The conclusions then that these authors reached from the study were that antibiotic treatment provided no benefits to Type-3, which were the least severely affected, and could probably be justified in Type-2, and demonstrated the greatest benefit in those with the most severe exacerbations.

They also noted that a higher success rate

in the antibiotic treated groups may be less important than the clinical deterioration. They found in their study that subgroups of individual symptoms were no more predicted about the outcome than were the group that constituted their severity scale.

The caveats specific to this particular study were first of all that no microbiology was done. All of the antibiotics used were assumed to be equally effective. It was of course conducted in the pre-resistance era.

Steroid use was not controlled, and there were relatively small numbers of patients in the study. Moving on then to I think another fairly well known study, a meta-analysis conducted by SAINT and colleagues, which was published in JAMA in 1995.

This study was a meta-analysis of nine placebo controlled trials of antibiotics in AECB. And it is important to recognize that these nine trials that were included were actually out of 230 studies screened, and that only those nine studies met their criteria.

That criteria that they used was that the study should be randomized, and there should be a diagnosis of chronic bronchitis, and AECB, and at least a five day duration of follow-up, and data

sufficient to calculate an outcome size.

Now, what they ended up doing, because there were different outcome criteria used in the different studies, was to calculate what they called an effect size, which is a unitless measure of efficacy.

The results were that when the trials were combined, they yielded an overall effect size, which was indicative of a small, but statistically significant effect, favoring antibiotics over placebo.

It is important to note, however, that the breakdown of the nine trials was as follows, which were that 3 of 9 sort of statistically significant benefit of the antibiotics; and 3 of 9 showed a trend favoring antibiotics; and 3 of 9 showed no difference from placebo.

Because the authors realized that the effect size would be a fairly confusing phenomena, they also looked at the most commonly reported outcome measure, which was the Peak Expiratory Flow Rate, and that was reported by six of this nine trials.

When they looked at those trials, they found that 2 of 6 showed a trend or significant improvement in Peak Expiratory Flow Rate favoring the antibiotics, and the others obviously did not.

The conclusion that these authors reached were that antibiotics yield a small, but statistically significant, improvement compared with placebo that may be clinically significant, especially in patients with low baseline flow rates.

The caveats in this particular meta-analysis was what we have already mentioned. That there were a variety of outcome measures used. In addition to Peak Expiratory Flow Rate, the duration of the exacerbation, the PaO2, symptom scores, or overall severity scores, determined by a physician, were all used variously in these studies.

This placebo control trial by Allegra, et al, was one of the ones that was not included in the same meta-analysis because at the time their original results were published in Italian.

However, they published a more recent analysis that described their entire results, and I wanted to present that to you today as another example of placebo control trials.

This particular trial looked at the amoxicillin/clavulinic acid versus placebo, both given in a five day course. And patients were greater than 40 years old had cough and sputum production, an FEV1 of less than 80 percent predicted and no patient

received steroids.

Of 761 patients screened, there were 369 exacerbations included in this trial, and the failure rate was given here, which was 49.7 with placebo, and 13.6 also received antibiotics.

The retrospective review, which constituted the second paper, showed that those folks who presented with low FEV-1, did worse with placebo. And they concluded that those with severe function impairment, and higher number of exacerbations, derived the greatest benefit.

I would like to present to you here not a placebo control trial, but actually an evidence-based clinical practice guideline put out by ACP and ASIM, and ACCP jointly.

What these authors did -- and it was published in Annals of Internal Medicine in 2001, was to review not only therapeutic interventions, but also modalities of diagnostic testing for utility.

In the review, the antibiotic treatment of AECB, they included 11 randomized placebo controlled trials. These included the nine that we have already mentioned that were included in the SAINT meta-analysis, as well as two that had been published subsequently.

In the review of these papers, these authors concluded that antibiotics are beneficial in the treatment of patients with AECB. Patients with more severe exacerbations are more likely to benefit from antibiotics.

I wanted to very briefly mention the placebo control trial that involved antibiotic treatment of patients with AECB. This was published in the Lancet in 2001, and involved a randomized placebo controlled trial of ofloxaci, and 400 milligrams a day, versus a placebo for 10 days.

These 90 patients were sort of a unique group, in that they did have AECB, but these are patients who presented severely ill enough to imminently require mechanical ventilation. The authors fairly rigorously excluded pneumonia, and they were allowed to receive aminophylline, but not steroids.

Given the extreme presentation of the patients, we see extreme results. The mortality actually was 22 percent in patients who received placebo, and 4 percent in those who received ofloxacin, and the secondary end point that was looked at was the requirement for more antibiotics and which also showed the same trend.

In addition, these folks had a decreased duration of ventilation, and hospital stay in the ofloxacin group. I would point out that again these patients were severely ill, and really what we are seeing here is most likely a prevention of hospital-acquired pneumonia, rather than treatment of AECB per se.

What I would like to present here is actually again not a placebo controlled trial, but a review of the same. The results that you will see here are from an AHRQ evidence report or technology assessment.

This particular document was prepared by the Duke University Evidence-Based Practice Center. The procedure for these documents is that the EPCs systematically review the relevant science-based literature on their assigned topics, and conduct additional analyses when appropriate.

When this group of investigators examined 11 placebo controlled trials versus antibiotic treatment, they included the 9 that we have discussed, and the two subsequent trials that were in the Bach study, but not in the meta-analysis.

I wanted to very briefly mention one of those two additional trials here, because I think it

illustrates one of the points that we are discussing. This as conducted by Sachs, et al, and was published in 1995.

And 71 outpatients who had TMP/SMX and increasing AECB were treated with either trimethrin sulfa, amoxicillin, or placebo. All of these patients received steroids.

There were no differences observed in the recovery rates, changes in symptoms, or peak expiratory flow rate, temperature, or sputum. And the caveats to interpretation of this study include the fact that the roll of corticosteroids anti-inflammatory effect is undefined.

These patients did have relatively high peak expiratory flow rates, and a low proportion of patients with purulent sputum, implying that there were perhaps not as ill as some patients in other studies had been.

The conclusions that the AHRQ documents reached was as follows. Randomized control trials of the antibiotic treatment of acute exacerbation of chronic bronchitis show overall evidence of a relatively small benefit in pulmonary function.

These trials suggest that patients with more evidence of bacterial infection, sputum

purulents, and more severe illness, worse peak expiratory flow rate, benefit most from antibiotics. However, this has not been conclusively demonstrated.

Likewise, the hypothesed interaction between corticosteroids and antibiotic use cannot be addressed by existing trial data. That concludes the review of what is available to us in the literature regarding the results of placebo controlled trials and the treatment of AECB.

I would like to reiterate what I think are some of the confounding issues in trying to reach a definitive conclusion in that determination of delta-1. First, there is the fact that concurrent effective therapies or other eogenous factors may diminish treatment group differences.

And clearly you have seen in some of the studies that systemic corticosteroids are one of those factors, as well as inhaled, short-acting beta agonists and bronchodilators, and oxygen therapy.

All of those have been shown in independent studies to have a treatment effect in AECB, and of course cigarette smoking also is going to have that same effect.

A very important point is the difficulty in defining appropriate patient populations for study.

First is the issue which has been referred to in other contexts of looking at bacteriologic end points.

Clearly in AECBs that is not possible because of the issue of sputum colonization with pathogens in the COPD. In addition, there has always been in various studies the question of the unclear role of viruses, atypical pathogens, environmental exposure, as well as non-infectious problems in the causation of AECB.

A very significant problem that remains to be addressed is the fact that severity criteria for this disease have yet to be validated. The assumption that the AECB severity can be judged by some combination of presenting clinical features is intuitive, but is yet to be confirmed by clinical studies.

Just as an example to show potentially how different populations of AECB can be constituted, what you see here are representations of the study that I mentioned to you from Winnipeg, as well as some data that was extracted from an NDA, which came to us recently.

What I wanted to point out was two things. First of all, obviously these three criteria -- the

FD-1, the sputum volume, as well as severity symptoms, which can be used or have attempted to be used to some degree of prognostic prediction, were given here in this study, but were not available to us for the NBA review.

As well, I wanted to point out that the patients here were significantly younger, and a much lower percent of smokers, either current or past, which may well affect the results given that the patient populations would be significantly different.

And I just wanted to very briefly mention the old versus new antibiotics, and specifically we all know that resistance is increasing, and that includes the pathogens that are presumed to be operant in AECB, and most of the studies that we have reviewed were conducted before the emergence of respiratory pathogens that are resistant to multiple antibiotics.

And having said that, however, I think it is important to know that there has been no randomized control trial which have showed the superiority of newer broad spectrum antibiotics in this disease entity, and there is no data to suggest increased failures with the increase in antibiotic resistance.

Having gone through this review of the studies then can we determine delta-1, which is sort

of what we started out with in the beginning. What we would like to be able to do ideally would be to perform a meta-analysis of the available literature, and then calculate delta.

The problems that we see in this approach are, first of all, that the patient population in placebo controlled trials that are available to us for review was not uniform.

Secondly, and probably one of the most important things, is that the studies that were available used very different designs, and very different end points, none of which were ideal.

The studies clearly had different outcomes, and some have shown a treatment effect and some did not, and most of these studies were not recent.

In conclusion then, in terms of the selection of delta, the performance of a meta-analysis, with subsequent selection of delta, would not yield a meaningful value due to the differences in study design, including heterogeneous patient populations, and diverse end points.

A review of placebo controlled trials of antibiotic treatment of AECB does not allow a definitive estimation of the benefit of active control

over placebo.

Patients with more severe -- with a question as to what that definition should be, a more severe illness, may benefit most from antibiotics, but this has not been conclusively demonstrated, nor have validated severity criteria been demonstrated.

What then are some options for what future trials should represent. Well, first of all, of course, would be non-inferiority trials in all patients, which is the current practice. But I hope that I have presented you data that convinces you that it is difficult to choose an appropriate delta.

Secondly, it would be placebo controlled trials with an early escape option in all patients with AECB, or placebo controlled trials only in patients who are perceived to be at low risk.

For instance, mild to moderate Groups 2 and 3, and of course another possibility would be to do placebo controlled trials in patients who have very severe presentation.

Another option would be non-inferiority trials in severely ill-only AECB patients, with the possibility of controlling for smoking and other concurrent therapies, and understanding that we need to have a reliable and reproducible definition of

severe AECB.

You have already heard about the possibility of three Arm studies involving a placebo, the new drug, and/or the old drug. And this would certainly be an option here.

Unresolved issues in AECB. First of all, are placebo controlled trials with an early escape option acceptable in AECB studies, and a corollary of that is should only patients with less severe disease be enrolled in these trials.

Secondly, if non-inferiority trials are conducted in AECB, what should the delta be? And lastly should future AECB trials include only patients with severe AECB. Thank you for your attention.

CHAIRMAN RELLER: Are there any questions for Drs. Powers and Thompson? Yes?

DR. ROTSTEIN: I would like Dr. Powers to comment on hospital-acquired pneumonia and the use of the clinical pneumonia severity index score that people have used?

There is a modified pneumonia severity index score that people have used as criteria for entry into nosocomial pneumonia trials, and also to gauge improvement. Could you comment on that? You didn't comment on that.

And also the use of quantification, particularly endotracheal aspirates, looking at greater than 10 to the 5th organisms per Ml.

DR. POWERS: Let me take your second question first. It becomes very problematic to validate the use of BALs or bronchoscopic techniques. There was a study by Fagan that actually looks at people that had purulent sputum, abnormal chest radiograph, and greater than 10 to the 3rd organisms.

Versus those who had purulent sputum, abnormal chest radiograph, and negative cultures done by that method. And the mortality rate was 26 percent in both groups.

And so does that mean that there is no difference between those groups or does it mean that the sensitivity of those bronchoscopic techniques is not very good?

Considering that those bronchoscopic techniques are not compared to any gold standard, that becomes very problematic, trying to tell what those mean.

When I looked over the four new drug applications for trovafloxacin and piperacillin, and tazobactam, ciprofloxacin, and linezolid, I did not see a use of that score that you are referring to, to

try to determine.

So the question I was asked or is posing here is that those may be useful. I am not aware of them, and I really can't comment.

DR. ROTSTEIN: One of the problems with those trials is they use a conglomeration of patients, a smorgasbord. The trovalfoxacin study excluded ventilator-associated pneumonia patients. So you could only be ventilated 48 hours or less.

I was one of the investigators in that trial, and I was one of the investigators in the linezolid trial as well, and that included ventilator-associated pneumonia patients. It was different.

But all the other ones have been mild-to-moderate hospital-acquired pneumonia, and that is why we have been unsuccessful in doing these trials. The money is really ventilator-associated pneumonia patients.

DR. POWERS: The question that comes up though is whether a company would want to study hospital acquired pneumonia in non-ventilated patients, and what kind of advice would we give to those people, and I will let the committee address that one as well.

CHAIRMAN RELLER: Dr. Archer.

DR. ARCHER: From a statistically challenged person, namely me, I have a question. Can you stratify in a trial like an AECB trial, where there clearly are different groups, can you stratify the patients going into the trial and assign a different delta to different strata within the same study, or is that a no-no? I guess that would be to the second person who presented the AECB.

CHAIRMAN RELLER: Dr. Thompson.

DR. THOMPSON: I'm probably more statistically challenged actually, but I guess the answer to that is -- and I am going to start and let you guys work on this.

But clearly there are subgroups within AECB that respond differently to bronchitis, and so whether it is a practical matter to assign a different delta to different populations, I think that would be problematic from a study design standpoint.

And from a clinical standpoint, I would say that we have yet to precisely identify them. So I think that would be the problems that I see theoretically if you could get around all of those issues, perhaps.

But thus far there is not a set of validated severity criteria that predict outcome. I

would say no. And I think the other interesting thing that needs to be further studied, and that I didn't present, is that there is a suggestion in several studies that the best predictor of prognosis is actually not the current presentation, but rather history of cardiopulmonary disease, as well as how many exacerbations they have had in the past.

And so it may well be that looking at those factors might be more predictive, but I know that your question is really delta, and I don't think that is practical, and I will let my statistical colleague address that.

CHAIRMAN RELLER: Dr. Temple, and Dr. Fleming, if you have comments on this.

DR. TEMPLE: Well, this is a complete cop-out, but you could certainly do an all-comers trial and stratify the population by the severity, and have different criteria for success in each of the strata.

It would really be multiple trials, but in a single environment. You might even have a superiority hypothesis in one, and a non-inferiority hypothesis in the other, but it really wouldn't be one trial.

Tom will have to tell you how you could do that in a single end-point or not.

DR. FLEMING: After the break maybe?

DR. BRITTAIN: You might want to use or you might want to base your delta on what proportion of people you have in your trial in the three groups, and you could think about it that way, and that would be one overall analysis.

But if you wanted to do it within each category, then you would need a sample size, and you would need a big sample size in that case.

CHAIRMAN RELLER: I think it is time for our afternoon break, and we will reconvene at 2:45, 15 minutes.

(Whereupon, at 2:34 p.m., the conference was recessed and resumed at 2:53 p.m.)

CHAIRMAN RELLER: Before Dr. Goldberger gives the charge to the Committee for discussion of the questions, we want to have transitional comments in response to the last query before the break having to do with stratification of patients in studies of acute exacerbation of chronic bronchitis, and what the appropriate statistical analyses would be, and Dr. Thomas Fleming has some comments to make on that query.

DR. FLEMING: Just very briefly. The question was asked if it would be at least possible to

entertain having a different margin in various strata or subgroups.

Thinking about it for a little bit, my sense is, yes, it is. Whether I would suggest that it is wise or not is an entirely separate issue. But if we used, for example, the setting of acute exacerbation of chronic bronchitis that we were just talking about, and if in fact, just to simplify this discussion, one took it as reasonably established that in less serious disease there is no effective antibiotics on the end-points of interest, and in more serious disease there is a 20 percent improvement, then in less serious disease you might have wanted to do a superiority trial using a margin of zero.

And in more serious disease, you would have allowed some margin. Let's say it is in fact the fullest margin that you might allow, which is a full 20 percent. Then essentially one could aggregate the data from those two strata, essentially in essence looking at the parameter of how much better are you than placebo.

So in the stratum of less serious disease, you are just taking the estimated difference between the experimental and the active comparator. Whereas, in the more serious disease, you are taking that

difference.

But then you are adding back what you think the effect is against placebo. You are rewarding an extra 20 percent in the stratum of more serious disease, thereby doing an overall stratified analysis that gives you a global estimate of how much you are better than placebo.

So that is just one of, and I just wanted to raise the fact that you could conceptually do it, and there are probably other ways to do it, too. The advisability of doing that is an entirely separate issue, because you are really mixing apples and oranges here a bit.

And you are taking a superiority component and you are taking a non-inferiority component, and you are imputing the full 20 percent estimated benefit that you think the active comparator antibiotic has in the more serious disease stratum, and that may or may not be the right thing to do.

But it is at least conceptually possible statistically to work out something that would essentially allow a different margin essentially in different strata.

CHAIRMAN RELLER: Thank you. Dr. Goldberger.

DR. GOLDBERGER: I actually almost started to go into the questions, and so I will actually try to keep my comments brief. We have heard a lot of presentations this morning.

We heard presentations from FDA staff on sort of backgrounds for evolution of delta, and some of the current concerns and issues from an FDA perspective.

Certainly from our perspective on one hand, while we recognize that there are real issues in some of these indications, and the ability to do clinical trials, and we also hope that we made the point that talking about delta is not just a discussion of some arcane statistical issue.

It in fact does have relevance to actual patient care and patient outcome. We heard a lot of prospectus from industry, IDSA, and academia. I think industry certainly indicated a strong desire to work in the development of new antimicrobial products.

But I think they tried to make the case that there are some real economic realities that they have to live with, and in fact in other presentations industry has been even more specific about what some of those constraints are.

And that they would like to see some

approaches that would allow them to operate within those constraints. Take the Infectious Disease Society.

They certainly showed a strong willingness to help in any way that it could with this process, and also I think expressed certainly a desire to provide as much expertise as they certainly could.

I think the Infectious Disease Society clearly is interested in their continuing to be an active pipeline of new antimicrobial agents. I am sure, although it didn't come out perhaps as strongly in their comments, they are also interested in ensuring that antimicrobial products that are out there, as well as new ones, are used in a manner that sort of preserves their useful life as long as is possible.

We also then heard in the afternoon some specific examples to help focus the discussion, dealing with several different indications, and looking at how much data we actually have in terms of thinking about delta-1 and delta-2, keeping in mind that the delta-2 is ultimately a clinical judgment.

One of the areas that we certainly heard a lot about is the issue of bacterial meningitis, and it is a very good example of some of the difficulties in

approaching this whole area.

And that is that on one hand it is beyond any question that the benefit of antimicrobial therapy is enormous. On the other hand, recognizing the severity of failure, which can range from death to at least a variety of developmental delays, hearing loss, et cetera, we would like our new antimicrobials to work as close as possible, at least to the same degree, if in fact not better, than what is already out there.

Yet at the same time, we recognize that to do clinical trials like that probably has sample sizes that are almost prohibitive. Therefore, there was some discussion about what would be the usefulness of focusing more on PK/PD, animal models, and microbiologic end points, as opposed to clinical success end points.

This is clearly an area that needs further discussion. I think one of the issues that perhaps was not entirely resolved was whether or not the bacteriologic end point really captures all the information that we need to see to be satisfied that the drug will be effective clinically.

Well, we have some questions which we will get to in a second, and that we obviously would like

some discussion on. We want to point out first that these questions are meant sort of to introduce discussion, depending upon the available time.

Certainly we would welcome other comments, areas of interest that the committee would like to talk about based on personal experience, and/or what has been presented today.

One issue in fact that would be nice to hear some discussion about goes back to something that I just mentioned a moment ago.

Both in the meningitis discussion and in some discussions at the break, I did hear the comment that from an antibiotic perspective, we really should be looking at what the drug does bacteriologically, rather than clinical outcomes.

And the question is how much weight should we put on this approach, particularly in more severe disease. On one hand, obviously a major role of antibiotics is of course to effect a bacteriologic cure.

On the other hand, if we don't get the requisite patient response, what are we supposed to do with that type of situation. And if there is time, we would welcome some comments about that. Leo, could you put up the first question.

The first area that we want to ask your opinion about is using AECB as an example, please discuss some of the different clinical trial design options in infections where the magnitude of the benefit of antimicrobial therapy over placebo remains uncertain.

And we have several different options here, and some placebo controlled trials, and three arm trials, dose response trials, and as time permits, you might want to expand this discussion to some to some other areas, i.e., otitis media and sinusitis, where there have been issues at times about the overall benefit of antimicrobial therapy.

From our perspective, beyond getting some input about trial design, we are obviously interested in ensuring that our approach appears to be most appropriate, and whether that means the same approach we have been using, or some modifications, we would like to get the best possible data that we can.

We also would like to think that given the relatively limited amount of data there is about the benefit of antimicrobial therapy in this indication, some of the clinical trials that might be used to seek approval might also provide some additional information on who the patients are, and who really

benefit from therapy.

Because realistically there is a lot of antimicrobial therapy used in bronchitis, and I think there is little question that the use of antimicrobial therapy, in addition to some degree of patient benefit, probably carries with it some development of antimicrobial resistance.

The question is are we getting the best trade-off right now. And if you could go to the second question, Leo.

And this is please discuss the implication of choice of deltas in clinical trials for serious infections. Please consider in your discussion the efficacy of a new drug compared to available therapy for the indication e.g. HAP and meningitis.

And basically the issues are smaller deltas and the effect on sample size of clinical trials, particularly when the infection is rare, and/or the success rate is low.

And larger deltas and the impact on patient care if potentially less efficacious drugs are approved.

And a simpler way I think of sort of summing this up is that there is no such thing as a free lunch. Either you spend the resources to be able

to do larger trials that give you more precise data, or there will be on one hand some limitations on what you know about the drugs.

On the other hand, if the cost is too high, the trials will never get done, and I think that this is an area that we would like to hear all your comments about.

It is a very difficult area, and it is a problem for us, and clearly a problem for industry, and whatever advice you can provide would be extremely useful.

And finally the third question. Please discuss what other factors, characteristics, of a drug product other than primary confidence interval results could be included in a risk benefit analysis supporting an FDA regulatory decision.

And certainly to be included in this can be safety considerations, PK/PD, availability of alternative therapies, other factors as you think appropriate.

Traditionally, we have been more flexible in situations where therapeutic options are limited, and where the disease is severe, and the alternatives may not be ideal, at least for some group of patients.

We would clearly think that this should

continue to be the approach in the future, and in fact I suspect there will be considerably more discussion about this tomorrow when we talk about the development of drugs for resistant indications.

Nonetheless, even though we believe we have some appreciation of the factors that are important in these decisions, we think it would be useful to hear some additional comment from the committee about factors that they would consider important with the degree of specifics that people feel comfortable providing. Thank you.

CHAIRMAN RELLER: Let's come back to question one. Discussion from the Committee, and by the Committee, I would include the extended Committee, those invited from IDSA, PhRMA, industry, and Members at all of the tables, including the proximal ones. Jim.

DR. LEGGETT: I forget I was on the end again once again, and so I might as well start. I spent my time during the break trying to think about this.

And regarding Issue Number 1, I think my overall bottom line is I would favor anything but what we are doing now, in terms of non-inferiority, among those three items.

I think in a trial ongoing with AECB, it is going to be hard to restrict the categories since we don't have any validated severity criteria. And I think the other thing about going forward and trying to include everybody is the closer we can make the Phase III trial to what is going to be generalized to outpatient use in the future, the more likely we are going to get some data that will help us.

And I think we also know in that regard that there is widespread antibiotic use as was just mentioned, even with acute bronchitis, and the people that are going to be using this are pulmonologists, general practitioners, and anybody but ID folks.

I think we definitely have going forward in these trials, we definitely have to account for steroid use. And if memory serves me well, in that Anthonisen trial, they went back and you could look at the steroid use, and that is what correlated with improvement in all three of the subtypes.

I think we could consider monitoring for deterioration as a primary target end point, rather than, quote, success/failure. I don't think we should use a microbiologic end point in AECB because the prevalence of the, quote, pathogen recovery from the sputum is the same, or even greater, when there is no

exacerbation, than when there are exacerbations.

And the density, in terms of CFU per Ml in the sputum is no different in exacerbations or non-exacerbations. And to the extent that acute otitis media and sinusitis are not diagnosed by puncture, and so we don't have, quote, hard data, I think they need to be treated the same as acute exacerbations of chronic bronchitis due to the similar colonization problems and the similar pathogens.

And with the same similar high placebo success rate.

CHAIRMAN RELLER: Dr. Cross.

DR. CROSS: Well, I would agree that in a situation like bronchitis, where we have a punitive infection in a non-sterile site, I think that having a bacteriologic cure would be extremely difficult.

And I think based on the evidence presented, that it seems certainly reasonable that a placebo in a controlled trial still ought to be the norm from the point of view that it is a less severe type of infection.

We have the alternative of having the early escape, which if properly designed would allow us to identify those patients who are at the highest risk who may benefit, as perhaps was indicated in the

Canadian study.

So I think that kind of design would allow us to at least for the next study perhaps perspectively identify criteria for folks who don't do well under the typical placebo controlled trials.

So I think that certainly given the natural history of that process, I think we wouldn't be doing the patients any undue harm, but still have the safety valve to ensure that all patients are safely treated.

CHAIRMAN RELLER: Dr. Archer.

DR. ARCHER: I think with reference to AECB, the patients that I see on the wards, I think one could establish criteria for the very non-severely ill, versus those that are very severely ill, and either stratify a study or divide them into two different groups.

On the one hand, I think most of the antibiotic use is really in the not very severely ill patients, and that is probably where most of the antibiotic resistance is generated as well.

Whereas, studies may overpresent the more severely ill patients. So therefore I think it is important to differentiate those groups, and doing a study may actually help define how you can separate

those two groups out.

And I would favor doing placebo control with the not severely ill, and non-severely, non-placebo control with some estimation of delta in the more severely ill.

And I think it is important in the severely ill patients to include all current types of therapy that are used for these patients who are deteriorating in their pulmonary function, to include inhale steroids, systemic steroids, all the nebulizer treatment, maximum therapy in that group.

Plus, antibiotics of different groups, because that is what is done, and I think sometimes that it is difficult to differentiate. One could maybe even argue in some of those groups that that placebo control is appropriate with everything else that is being done, but I leave that to the pulmonologists.

As far as other types of infections, I don't see much acute otitis. I really can't comment on that, but I think that sinusitis is difficult to define, and it seems like more microbiological data should be generated, in terms of punctures.

Or possibly doing CT scans to try to define who does and doesn't have sinusitis as a

criteria for study entry, because I think there is also a lot of inappropriate use of antibiotics for poorly defined sinusitis, and a lot of antibiotic resistance being generated in that as well. Let me see. I guess those are major comments.

CHAIRMAN RELLER: Dr. Ebert.

DR. EBERT: Well, it appears that there are a variety of things that are going to impact the size of the patient population in these studies, one of which is the prevalence of the disease, and secondly, the impact of therapy on outcome.

And I think an acute exacerbation of chronic bronchitis, both of these speak towards the use of a large-scale study. It should be an adequate patient population, and also because we are not really clear on the impact of outcomes, a larger population should help us in that way.

I think if we want to go back to the basics, it would be to do a very large scale study, and try to validate subsets of patients who do in fact respond, and who do not.

If that in fact does not work, or if that is not the tract that we want to take, certainly we have talked in this committee about enriching patient populations, or selecting out specific criteria for

entrance into the study to ensure that the populations that we are treating are going to be at greater likelihood of response.

I also agree that a microbiologic response is not likely to be a good end point for this particular disease, which really leads us into the clinical response, and the question I have there is really again the issue of the timeliness of the assessment.

And I don't recall hearing any discussion of the time frame at which we are assessing clinical response, and certainly with other disease states we have talked about assessing patients at 28 days from the beginning of enrollment in a study.

And we have argued that that may in fact be too long of a time. So it may be that we need to look more closely at end-of-treatment as an assessment, rather than some time point in the distant future.

CHAIRMAN RELLER: Dr. Ramirez had a question, and then Dr. Patterson.

DR. RAMIREZ: Just a comment. Just to add a new factor to the complexity of the problem, is that even though these factors are not well-defined in the literature, when all different medical societies get

together to develop guidance for the management of antibiotics in respiratory tract infections, for exacerbation of chronic bronchitis, and nosocomial pneumonia, and hospital-acquired pneumonia, the idea is not to look at these diseases as a single disease.

And we can clearly see, for instance, that in community-acquired pneumonia, we all agree that there are 3 or 4 groups of patients with pneumonia, and with nosocomial pneumonia, there are at least 2 or 3, or 4 according to the society.

And in acute exacerbation of chronic bronchitis, there seems to be that there are at least three groups of patients. And the classification of patients mostly is based on the severity of the disease.

And what we are trying to do is trying to help the clinician in selecting empiric therapy based on the likely resistant organisms causing the disease. And the problems that we are having is that we have antibiotics that are approved for all community-acquired pneumonia, and all acute situations in COPD.

When in reality we know that the patient with mild exacerbation, or I shouldn't say mild, but a patient with low risk, for an acute exacerbation of low risk, meaning that considering the three criteria

considered in the respiratory starts with FEV1, and considering the prior use of steroids, we know that these patients primarily are going to be infected with H. flu, and this is one patient.

And then the other end of the spectrum is that we have the patient with the high release for possibility for infection to due pseudomonas aeruginosa.

Then the use of an antibiotic for acute exacerbation of COPD, you probably need to contain the patients within a risk factor for resistant organisms, and trying to define again populations that we are discussing here with otitis media, and trying to define a patient that may have the resistant organism, or a particular organism.

I am trying to define antibiotic therapy more specific for a particular group of patients. I think we all agree that if you have only one of the criteria, you should not get antibiotics.

But with 2 and 3, and then the patient is hospitalized, there is no question that we get the feeling that antibiotics are necessary. I think that a stratification of the patient is critical in any one of these clinical trials.

CHAIRMAN RELLER: Thank you. Dr.

Patterson.

DR. PATTERSON: I would agree with Dr. Archer that the placebo controlled trials with escape for the Type II and III patients with AECV would seem appropriate.

I would be more concerned about the placebo controlled trials for the patient with the more severe disease, and perhaps maybe there are a large number of patients in this group, and that could be one place where you could use a smaller delta to evaluate that.

But I think also you could look at other outcomes or endpoints like the duration of time between exacerbations, and also not bacteriologic eradication, but the flora that is present at the recurrence of the exacerbation, and also to look at a comparison of therapy with symptoms, versus interval pulse therapy or prophylaxis, whatever you want to call that.

And looking at duration between exacerbations and also comparing susceptibilities of the flora at recurrence between those two groups, and would you get less resistance with one group versus the other.

Regarding other infections like otitis

media, I think that this has already been said today, but I think that the double tap is of interest and that bacterial eradication is an end point, although that is difficult to do in this country.

There are some centers that do that in other countries, and that is of interest as an end point. And regarding clinical outcome as an endpoint, I think it is another area where you could use a smaller delta because of the large population of patients.

CHAIRMAN RELLER: Dr. Fink, please.

DR. FINK: Well, speaking as a pediatric pulmonologist, I don't treat chronic bronchitis except in cystic fibrosis, and where we do see it rarely, but being familiar with the literature, I think there are some complicating features that using AECB as an example our important to point out.

This would be a situation in which international studies would in all likelihood be highly flawed, and the reason for that statement is that in the United States, we take cigarettes away when patients are hospitalized.

That is not done elsewhere in the world, and if you are going to deal with a controlled trial of chronic bronchitis, whether or not the patient has

access to cigarettes or not is probably going to have a significant effect on the response to treatment.

We also blame a lot on H. flu. There is a lot of newer data that says organisms such as RSV, chlamydia, mycoplasma, which often with the exception of RSV, and at least chlamydia and mycoplasma, often respond to the same classes of antibiotics that are used to treat H. flu.

And that these organisms may be playing a much greater role in exacerbations of chronic bronchitis than is currently recognized. So I think that part of what we need is better classification of chronic bronchitis. It isn't all the same.

And from a clinical standpoint, probably previous ICU admission is actually better than a scoring system for disease severity, in terms of risk of hospitalization.

So I think part of what we really need in chronic bronchitis is better classification, more comprehensive studies with a really good look at microbiology, including non-bacterial pathogens, and a better understanding of the disease before we can really design better trials.

CHAIRMAN RELLER: Dr. Ramirez.

DR. RAMIREZ: I will agree, because we

have been saying that serious infections, that you need to select the best therapy, and for this one, you need a small delta. But according to the recent identifications, patients with severe COPD has a higher mortality than a patient with nosocomial pneumonia.

And then we are going to be talking -- I mean, if we are one of these patients with prior hospitalization to an intensive care unit, that is another observation, and there is a very high probability that this patient is going to die during this hospitalization. And this is the type of patient that we need to be sure that we give the right antibiotics.

CHAIRMAN RELLER: Dr. Bennett.

DR. BENNETT: Several of us have commented about placebo controlled trials with early escape, and I am not certain that I really understand that. It sounds to me more like early discontinuation.

But if my understanding is correct, there are three things that we ought to take into account if we adopt a strategy of placebo control and early discontinuation.

One is that you would have to make a double blind. Otherwise, you would have people with

lack of confidence in the experimental drugs and stopping the drug for that reason.

The other is that I think you would have to have very rigid criteria as best you could for discontinuation. So it didn't become very center dependent on who wanted to stop the drug early, and particularly if the two drugs being compared were different in their toxicity, for example, and that one caused much more gastrointestinal distress.

And you are now mixing two end points, efficacy and discontinuation for toxicity. You would probably be well advised to have a blinded data review committee to look at all of the patients who had premature discontinuation, or who escaped if you will because you would want to see that there was some element of uniformity between centers, and that the study definitions were actually followed.

And the last was I am concerned that early discontinuation may not give one of the drugs a chance to show its effect. For example, if everyone got the drug for 1, 2, or 3 days, you may not be convinced that that was enough to actually give the drug a chance.

So perhaps those of you who understand early escape better than I do could explain how we

would get around these.

CHAIRMAN RELLER: Dr. Fleming.

DR. FLEMING: I wanted to comment on just that issue, and I don't know if you were commenting on something else. Well, I think you have raised a very important issue, and I am struggling with this as well.

I am not yet convinced that early escape would work here, and in my thinking I am going back to Dr. Thompson's slides, numbers 11 and 13. On 13, she is talking about success rates relative to what I understand the primary success definition is given to be in Slide 11, which is symptoms resolved within 21 days.

So if that is in fact is the primary end point, I worry if early escape means dropping off the placebo at some point before 21 days. If it is dropping off the placebo after 21 days, then I am not so concerned, and here is my worry.

The data on page 14 or 13, rather, is telling us that eventually we should expect on placebo convergence to a 55 percent success rate at 21 days. At 21 days, non-placebo, 55 percent will have resolution of symptoms.

But suppose though at day 10 it is only

half that large, and I have no clue how rapidly this occurrence of resolution of symptoms occurs, but let's say it is only half that large.

So let's say it is about 30 percent. There are 70 percent who have not yet resolved, and if a number of those people now escape placebo, and now you impute failure automatically, you are going to underestimate what the actual true success rate would have been on the placebo.

So if early escape means dropping off the control arm prior to the time period at which you would have achieved your full effect on the control arm, you are going to have a bias underestimate of the success rate on the control.

On the other hand, if early escape means, no, no, everybody will be on at least 21 days, and then they can escape thereafter, then my concern is not relevant.

CHAIRMAN RELLER: Dr. Temple.

DR. TEMPLE: There is a fairly narrow experience with so-called early escape, where its recurrence of symptoms like unstable angina is fairly easy, and there have been trials that have been successful using that.

The reasons for doing it though are

ethical, and so you have to choose an escape provision that satisfies your ethical needs. And I don't know whether going 21 days satisfies your ethical needs or not.

Intuitively, I would say somebody gets tremendously febrile and looks really sick, you get them out, and start treating them, even though you don't really know why that is happening, you just accept that.

But that is really a clinical judgment. clinicians have to sit down and say, okay, what scares me, and what makes me worried about the fate of this patient, and your obligation, and accompanying permission to use a placebo where there is arguably at least standard therapy, comes with some well-developed, mutually agreed on criteria for what constitutes actions that would protect the patient against going down the tubes.

But in the absence of a lot of examples, it is not easy to say what those are, and Dr. Bennett, who doesn't understand this at all, raised all the right questions, of course.

But nobody really understands it. There are examples that are easy. We have seen a withdrawal study with -- never mind. I am mixing two things. We

have seen early escape associated with randomized withdrawal studies, and that is probably the case where they have been used most.

And where people have looked at recurrence of initial symptoms, and there have been cases where blood pressure over a certain point in non-responsive patients who are being studied with a placebo got them out of the trial and on to therapy.

And you work it out on the spot, and I have no doubt that these early escapes probably decrease the apparent benefit of the drug. It depends on why you leave early. But you pay that price for the ability to get information in a setting where it is difficult to get it.

DR. FLEMING: Or they could lead to an exaggerated estimate effect if you are imputing failure in the placebo, when in fact further follow-up of that placebo patient would have led to a higher level of success.

My sense of interest in being able to do a placebo controlled trial, I share that with others here that it gives us in a real sense the truest way of determining whether or not the intervention is efficacious, is to do a head-to-head with the placebo.

And if in fact we can reliably assess that

in short term follow-up in such a setting, the early escape concept is appealing. If in fact though we are not able to follow the control patients adequately long through the period in which we can get an unbiased assessment of outcome, I think I would be more included to do a head-to-head comparison against a standard of care that is largely anticipated to be relatively ineffective based on what we are hearing from the data, at least in the less ill patients, where you wouldn't have to escape.

You could follow these people through 21 days and really establish superiority. So either doing a head-to-head comparison against standard of care, or in addition to standard of care, looking for superiority.

And then if in fact we truly believe there is interaction here indicating that there is adequate data establishing the antibiotics are effective in those patients that are more severely ill doing a separate non-inferiority comparison in that population, those approaches would be alternatives to early escape that should also allow us to determine whether or not we have truly added benefit relative to what is currently the standard of care.

CHAIRMAN RELLER: Dr. Shlaes, Dr. Wittes,

and then Dr. Powers, and then we will have hands up again, and we will get the next three.

DR. SHLAES: I just wanted to try to keep this in prospective a little bit, at least for me. So I think that most drugs that are developed for AECB are actually oral drugs that you would take as outpatients, and so I don't think this is directed at those patients who just came out of the ICU and are coming back to the hospital for another acute exacerbation, where they are going to get admitted again.

So I think it is really -- and to keep this in perspective -- the outpatient setting. The other thing is that I think the 21 day evaluation was not 21 days of therapy. It was just that that was the time, and I think they pulled that number out of the air.

I mean, I don't know why they picked 21 days in that study, particularly if anyone knows, and maybe Dr. Thompson knows why they picked 21 days in that study. I don't know.

But I think it was just a time when they could bring patients back and get another FEV1 that was realistic, but that is not 21 days of therapy. So you could have much shorter therapy, and withdrawal

during the shorter therapy, and still have a 21 day evaluation for FEV1.

And again I think the risk given outpatient therapy, or early antibiotics, and hurting somebody with a very severe disease would be small.

DR. FLEMING: By the way, I was assuming it as you had indicated as well, that the end point is follow everybody 21 days and find out what fraction resolved their symptoms, which would be something that I would want to know whether somebody maintained therapy for 4 days, 8 days, or 21 days.

And my concern is that if in fact natural history would show resolution of symptoms, and the rate increases as you follow people for a longer period of time, such that 55 percent have resolved by 21 days, and only 30 percent by 10 days, if we are pulling out in that 70 percent who haven't resolved by 10 days in the escape clause, and hence impute non-success, then we are going to have a final result of 30 percent success on an arm that really should have had a 55 percent success rate. That is the nature of the bias that I am concerned about.

CHAIRMAN RELLER: Dr. Wittes.

DR. WITTES: My comment has to do sort of in general with this, with the valuable percentages

which have disturbed me today.

And related -- and this is not unrelated to the early escape, but it seems to me that in these AECB trials, as in the others, I find this 35 to 50 percent invaluable rate just too high.

And somehow it seems to me that in order to evaluate whether a therapy is working or not, there has got to be a way of including end points for a higher proportion of people.

And in terms of early escape, and I fully agree with Tom, that the risk of this design in this sort of situation, where you are evaluating 21 days as the end point, if you have early escape designs, it may change the end point.

The end point may be time to more aggressive therapy, or time to being able to be off it, or something like that. So that the design and the end point should -- that the end point should help influence the way you choose the end point. It should not be locked into an end point and then all designs say that.

CHAIRMAN RELLER: Dr. Powers.

DR. POWERS: I had a question for Dr. Fleming that relates to something that Dr. Bennett said. Often times when we see people get discontinued

from therapy, it is hard for us to tell as medical reviewers why they discontinued from therapy.

And we used to get investigator comments or a printout of handwritten or typed out as to what the thinking of the investigator was at that point. We don't get that at all anymore, and so it is hard to tell why they discontinued, and I often think that perhaps the discontinuation is more of a measure of investigator nervousness than it is of the patient actually doing poorly.

Would something like Dr. Bennett suggested firm rules for discontinuing patients address some of the concerns that you raised about underestimating the effect of placebo in those trials if you could at least discern why the patients actually failed? Now, that obviously brushes over the devil in the details of determining what is a clinical failure in making those rules, but would that address part of the problem that you raised?

DR. FLEMING: Probably partially, but not fully. Just to follow the example that I was giving. At 10 days, you have had 30 percent that have resolved symptoms, and 70 percent haven't. In that 70 percent, of those 70 who haven't, eventually 25 will over the next 11 days if your criteria for escape are

sufficiently stringent that none of those would qualify, it would resolve my concern.

I kind of doubt though that you are going to be that effective in being able to fully distinguish who those 25 are from the other 45. And so I think it would partially, but not fully, address the concern.

CHAIRMAN RELLER: Dr. Hardalo.

DR. HARDALO: I think you have actually brought up some very important issues. First, as Dr. Wittes said, the evaluability rate is one of the challenges that industry has to deal with since we sponsor most of the clinical trials, and has a lot to do exactly with investigator confidence.

But it also has to do with the lack of clarity that we see, and where we would want guidance from various stakeholders, including IDSA, and the American Thoracic Society as to how do they define treatment failure.

Is it failure to improve within the natural history understood by them for that disease, or is it clear cut deterioration and progression based on objective criteria.

That very much impacts exactly how can we detail discontinuation rates. But also it has a lot

to do with evaluability rates. If there is no clear cut objective criteria, what you have is patients coming off the study for rather soft reasons, which makes them unevaluable.

You just simply don't have enough data with your sample size to make any clear conclusions about the efficacy of the drug, or the safety of the drug.

In addition, there are a variety of factors, not the least of which are the clinical practice. If you are practicing in the United States, it is simply impossible to have patients come back for daily visits on an ambulatory basis. It just is not going to happen for most of the centers.

So you need to have a compromise as to what is getting done in clinical practice, versus what is a requirement for a clinical trial, so that you have good quality data.

And I think not the least of which is that we also have to have assessments which are practical. That although I really myself would like to have some studies that require TAPS or quantitative cultures, in reality, in managed care settings in the United States, and in most of Western Europe and Canada, simply microbiology has gone by the wayside because of

the emphasis on managed care that it ultimately does not affect what is done to the patient in terms of the choice of antibiotics.

Therefore, the only microbiology data that we do get is in the setting of clinical trials, and even then it is going to be quite limited. So, yes, we would love to discuss what would be relevant entry criteria, and what would be relevant interim evaluability criteria for discontinuation rules, and what would be relevant end point data so that all of us can get the best quality data from whatever sample size we agree upon.

CHAIRMAN RELLER: Yes? Please, your name and please comment.

DR. TALBOT: George Talbot, Barth. Sorry. Hiding behind the water pitcher.

CHAIRMAN RELLER: If I put my glasses on and I wouldn't need the introductions. So help me out in the afternoon. Thanks, George.

DR. TALBOT: This is an awfully long way away from you, and so I understand. I have a general comment, a big picture comment, as well as a specific suggestion.

The big picture comment is that it is very interesting to me to hear this committee talk about a

placebo controlled study design. I think that that is in some sense quite remarkable, and I would like to compliment the FDA, and the FDA presenters for actually presenting the group with the opportunity to break the paradigm of clinical trials in this indication.

I think that the opportunity this presents for the community to learn about this disease and how best to treat it is really quite remarkable. So I think it is a very good thing. Now, the problem with breaking the mode is that as you try to implement that, there may be resistance to change.

I could imagine resistance to change at the level of IRBs, of Investigators, and of other concerned groups. So I think relative to some of the points that have been made about violability, and about early escape designs, and so forth, that really it is incumbent to take these discussions to a working group level so that IDSA, and other groups of clinicians can offer the specifics which allow these changes in design to be implemented safely, appropriately, and with the confidence of the end users; that is, IRB's patients and investigators.

CHAIRMAN RELLER: I would like to follow up on Dr. Talbot's comments. We heard earlier that

some of these patients who are marginal in terms of gas exchange, and may be intubated, hospitalized, because the acute exacerbation throws them over in terms of respiratory pulmonary function.

Would it be important if we are considering placebo controlled trials to assure that those patients don't have pneumonia with a negative chest radiograph, so that we are really talking about acute exacerbations of chronic bronchitis?

And I was impressed in Dr. Thompson's review. I am not at all convinced that if patients were -- we had a randomized double-blinded control trial, with appropriate supportive measures -- bronchodialators, steroid use -- that we are at all confident that antibiotics contribute much or anything in these patients.

And if that be the case, I was also impressed by this morning's discussion of all of the subtle, sometimes covert, obtuse pitfalls in these non-inferiority trials.

Wouldn't it possibly be much -- and the dilemmas with the large number of patients, and the large number of patients who were excluded because they can't be a valuable.

Would the practice of medicine be advanced

by just going ahead and demanding rigorous double-blind placebo controlled trials for this entity as a more efficient way to see whether or not a drug is effective or not?

And related to that is I am confused about what it adds to have a placebo, a comparative agent, an active control -- a new agent and an active control, and a placebo, all in the same study, because it seems like you are making things almost impossible to sort out when you get into the discussion of deltas.

Why not just do a placebo controlled trial and get on with it? Dr. Temple.

DR. TEMPLE: Let me partly answer that. In settings where you are convinced that certain drugs are effective -- and depression would be a good example -- a three-arm study is an extremely informative study.

If you run the trial and your control agent wins, and your new drug loses, you find another drug, because you have learned what you needed to learn. This is a study that had assay sensitivity, and your drug could not be shown effective in that study.

If on the other hand both the control

agent and your drug fail, then the study couldn't distinguish active from inactive drugs, and you don't have any reason to be depressed.

Now, here it is more complicated, because from what I am understanding, nobody is entirely convinced that any drugs actually work. The only reason for including -- there are two reasons for including the active control.

One is to -- and as Tom said and others did, to see how the new drug actually compares with the other drug in a setting where you establish assay sensitivity, and that is not very important if you don't think they work very well.

The other is that in case that you really in your heart believe this other drug works, this allows you to distinguish from a setting in which you can't tell anything from a setting in which you can tell things.

So it can be an extremely informative design, and that's why people in depression and hypertension, that is actually the standard test now. Almost everybody does it all the time.

CHAIRMAN RELLER: Dr. Glode.

DR. GLODE: I obviously cannot comment on AECB as a pediatric infectious disease doctor, but I

just wanted to reiterate Dr. Talbot's points that have bothered me, and that is the issue of the sort of standard of care.

If somebody is writing 12 million prescriptions every year for this then patients and doctors have some belief in antibiotics. And so I am very worried about the introduction of placebo controlled trials relative to both the patient and the local IRBs, and sort of the issue of if the FDA says it is fine, does the world believe it, and are willing to approve it.

I think that is a big hurdle and that becomes a big hurdle if people won't enter the trial, or if you can't get it through your IRB.

CHAIRMAN RELLER: Dr. O'Fallon.

DR. O'FALLON: I think it is interesting that -- well, I will just say my point. We haven't really made enough of a point that what is under the surface of all of this is the overuse of antibiotics and what we are concerned about is the coming disaster of overuse of them.

So, in an issue like this one, or in a setting like this, it may very well be that there are all these prescriptions that are being written every year for something that the drugs aren't helping, and

we don't have the data to prove that they either do or they do not. So there is an issue here to stave off this growing wave of drug resistance.

CHAIRMAN RELLER: Drs. Ramirez, Cross, and Chesney.

DR. RAMIREZ: I just have a question. I have no problem to do it with a patient with mild COPD, a placebo-controlled trial, because I have not seen any data in the literature that indicates that antibiotics are better than placebo.

And I am sure that I am not going to have any problem to convince my IRB to say that if you have a patient with COPD, which was described as just a clinical entity.

But if you have a patient with COPD with mild exacerbations, and with just a couple of years of COPD, and if that were more than 75 percent, then nothing is going to happen to this patient if they don't take antibiotics.

And I am sure that at this moment I can convince the patient that we are doing a trial to see if we can avoid giving you antibiotics and develop 5 years down the road resistant organisms, and the patient is going to be happy to be in the placebo arm.

And then I have no problem, but the

question is that if I am in industry, and I come up with this new antibiotic, who is going to pay for this study to test my drug against the placebo?

Everybody wants to test their drug against the other drug, and to be sure that my drug is going to be on the market. I mean, how are you going to convince the industry to do a study of a new antibiotic that is going to be tested against a placebo? Who is going to pay for this?

CHAIRMAN RELLER: Let's continue around the table, and we will get everybody, including Dr. Temple and Dr. Nelson. Alan.

DR. CROSS: I would like to just follow up on a comment that Dr. Temple made about the three-arm study, and about including an arm that has the, quote, standard, drug. In our last meeting on sepsis, a slide was shown which the presenter made the point that there were at least 4 or 5 drugs that in the first trial were shown to be effective, which upon retrial were ineffective.

And I am just wondering in the area of infectious diseases do we have any examples of antibiotics, which on repeated trials have had about the same approximate point estimate of efficacy.

And I guess a corollary to that is simply

the second point which you made earlier about assay sensitivity, and can we measure the difference in effectiveness between drugs.

But at least from what we heard this afternoon, there is even a more basic aspect of the issue of sensitivity. And that is diagnostic sensitivity, especially when we talk about things like sinusitis or bronchitis.

And it appears that in the reviews that we heard that there were various criteria for making a diagnosis, such that it is really hard to even compare most of these studies, even if you did have an answer for my first question about reproducibility of results in these specific areas.

CHAIRMAN RELLER: Dr. Chesney, and then Dr. Temple.

DR. CHESNEY: Thank you, and I hope that I can keep my thoughts organized here. But I would like to echo a point that Gordon made, which is that -- well, first of all, how did we get here. We got here because colossal overuse of antibiotics by comparing one to another.

And I think several points -- and number one being, I don't think we know the natural history of a lot of these diseases. I don't think we know the

natural history of otitis media or sinusitis, or AECB, because we began using antibiotics before people recognized my second point, which is I think there are subsets.

And very clear subsets within these groups, and I think those of us in pediatrics could clearly identify subsets of children who had acute otitis media, and one of the big problems has been that they are all just put together in these studies, and they don't distinguish a two month old with a temperature of 106, with an 8 year old with no temperature sometimes.

And so I think that we really don't know the natural history of what we are using the vast majority of antibiotics for, and as that beautiful wheel diagram from the CDC continues to demonstrate.

So for me mild diseases is the real issue, and I don't know how we are going to get some of these answers without using placebo controlled studies. And I think a point that Dr. Talbot made that is so critical, is to get the right players together.

The people that are doing the double tap studies on otitis media have some very well defined concerns and ideas about how to do these studies with a very small number of patients, for example, for

acute otitis media, and we all heard Dr. Dagan a few months ago.

So I think getting the right people together and looking at the issue of subsets, and readdressing the whole issue of natural history for me are really the big points.

And determining what kind of delta to use, or what kind of study to use, is obviously important. But I think that is going to take a lot more discussion within the smaller groups of right players if you will. Thank you.

CHAIRMAN RELLER: Thank you. Dr. Temple, and Nelson, and then Metlay.

DR. TEMPLE: Presumably one of the reasons studies come out differently is in fact the difference in diagnosis, or the difference in the population that got into a particular trial.

If you had reason to believe that there was an effective therapy, the effective therapy accompanying the test drug helps you know whether this was a study that got the right people into the trial or didn't.

Now, if really there isn't any right population, and we don't know whether any of this works, then that is a different question. I just

wanted to comment on something that Dr. Talbot said.

We have recently gotten through about a year and a half in which many people, including the people who wrote the Declaration of Helsinki, asserted that you can't use placebo controlled trials when there is effective therapy, even for mildly symptomatic diseases, a headache or something like that.

So the discovery that FDA and the advisory committee wants to have placebo controlled trials of antibiotics for goodness sakes will draw attention. There is on question about it.

The answer I think lies in the very things that you have been discussing. You have real doubt about whether people are being harmed or helped by this.

You may be setting them up for a resistant organism infections later that will take their lives. So the case will be made on the credibility of those assertions, and the lack of information about whether there really is anything very effective.

But it will draw tremendous interest. I don't think there is any question about that from IRBs and others who are very nervous these days about placebos.

CHAIRMAN RELLER: Dr. Nelson.

DR. NELSON: I think actually a good lead-in to my question as the Chair of an IRB, it is unclear to me from a study design perspective that there is any difference when you can't tell between the placebo and an active control, and between an active control superiority trial and a placebo controlled superiority trial.

So I guess I am asking to be educated that if indeed physicians like me in an ICU who probably reprobate in the use of broad spectrum antibiotics as my patient is deteriorating, or families who are not going to be willing to go into a placebo controlled trial or patients.

And from a study design perspective, is there any difference between the active control superiority and placebo controlled trial in this kind of setting to where you can have your cake and eat it, too, on both sides, and placing the issue of resistance and over-use aside.

I mean, I am finessing that issue at the moment. Is there?

CHAIRMAN RELLER: Dr. Temple.

DR. TEMPLE: Well, Tom referred to this before, too. If there were reasons to think that one

drug was actually superior to another, then go ahead and do a superiority trial. That always works and it is interpretable.

The question is whether there is any reason to believe that that is true, and if it is not, then a superiority trial can't work, won't work, and there is not much point in it.

And your only choice is to do a non-inferiority trial, which Dr. Thompson explained can't be done. And sometime else, namely a trial against placebo, with appropriate are that people don't get hurt.

DR. NELSON: But, Bob, if the placebo and the active control are not shown different in any studies that have been performed, then what is the difference in selecting the active control over the placebo in that context?

DR. TEMPLE: No, I agree with you. If there is no reason to believe any of these things work, then there is not much point in not just going ahead and doing a placebo control trial, and only if you think that some of them do work in the right setting is there a reason to have that.

CHAIRMAN RELLER: Did I understand correctly, Dr. Nelson, that you are suggesting that

why not always have an active control if it really is tantamount to a placebo. Is that what you are saying?

DR. NELSON: No, no, I wouldn't want to go that far.

CHAIRMAN RELLER: Because if that be the case, then if we could think of some examples, then we would have examples of the very thing that initiated this whole delta discussion.

DR. NELSON: Well, placing the mild issue aside, if you want to carry this into a more severe disease setting, it is unclear to me if the argument that you can't determine a delta is based on the lack of difference or reproducible difference between the placebo and an active control in existing studies, it is unclear to me that from a study design perspective there is any difference then whether or not the control group is an active agent, or the placebo agent, based on those prior studies.

And so if indeed you are arguing on a feasibility that patients, families, and physicians, would be more accepting of an active control from a study design perspective alone, it is not clear to me there is any advantage of the placebo group.

That is the question that I am asking, as much as wanting to be educated from that, so that your

feasibility would actually be improved by the active agent if you did a superiority trial.

When I read E-9 and E-10, which I read to be educated, I see a lot of discussion about a superiority design is superior to the equivalence design. So it is unclear to me why that is constantly being sort of placed aside.

CHAIRMAN RELLER: Dr. Temple.

DR. TEMPLE: I'm sorry to keep doing this, but the distinction is between -- you can have an effective drug for which you nonetheless can't describe a delta.

We have thought about anti-depressants for a long time, and about half of the satisfactorily designed trials of drugs we know to be effective can't distinguish drug from placebo because the diagnosis is different or people get better. Nobody knows why.

But it is a fact, and which means that in any given study that you can't know what the effect of the active drug is, even though we are perfectly convinced that those drugs work.

And the situation here could be none of them work at all, and none of them are known to work, and there is no evidence of anything; or it could be that it is study dependent.

That is, that if you get just the right people, maybe it works then, and things like that. Those are reasons why you can simultaneously not write or not design a delta, not identify a delta-one, but might find it useful to include a putative active drug as a control.

You would never need to do that, but it might be informative, too. But the differences between the assurance of assay sensitivity in any given trial, and the overall effectiveness of a drug. There are many effective drugs for which you cannot design or describe a delta, a delta-one.

CHAIRMAN RELLER: Dr. Fleming.

DR. FLEMING: I think what Dr. Nelson is raising is a very important point, and if I am following what he is suggesting here, is that it is in a setting where standard of care is widely accepted, but thought to have relatively little impact on the end point, either favorably or unfavorably.

And then is it ethically more comfortable and easier to enroll in a robust fashion by randomizing patients to that standard of care against the experimental, where you still have to show superiority.

So you don't run into where we run into

troubles and if you are trying to show non-inferiority there, it is not acceptable because there is no legitimate margin.

But I think what you are saying is if you are truly intending to show superiority, isn't that an alternative approach to doing a placebo controlled trial that might be more ethically acceptable, and might allow for more rapid enrollment.

And my own sense about this is in fact it is, and it is not unlike the concept of doing dose response, giving a low dose and a high dose, where you are hoping that there is a gradient there such that the high dose is much more effective than the low dose.

And the risk to this approach is only if in fact the active comparator really is more effective than you think, and it is absorbing a fair amount of the efficacy of the experimental; or if it is adverse, and you are not recognizing that.

Many examples of this exist. Just one example of a trial that we were involved in, which was looking at reducing maternal-to-child transmission of HIV in developing countries, where the standard of care, when we did this study a few years ago, was still placebo.

And we designed a placebo controlled trial against a short-course AZT regimen against a short-course novarepine regime. Ethics Boards eventually closed down the placebo arm, but allowed you to continue the short-course AZT, short-course novarepine comparison.

And the short-course novarepine have the transmission rate of HIV relative to short-course AZT, an example of what you are talking about. Now, maybe the actual effect of short-course novarepine is even more than a halving, but it is sufficiently more potent that we were able to show a difference in a trial where it was judged ethical, because everybody was getting an active intervention.

So if in fact you believe that there is considerable uncertainty about whether the standard of care is effective, but it is widely accepted, and there would be serious concerns about doing a placebo, you could do a head-to-head superiority comparison against that active comparator.

And as long as it is relatively inert, in terms of efficacy and risks, you would actually get an informative sensitive answer to whether the experimental therapy is effective.

CHAIRMAN RELLER: To continue the train of

discussion, we went a couple of circles. Dr. Metlay, your turn, and then we will come back to the floor table.

DR. METLAY: Thanks. Well, first a comment on this discussion, which is that I don't think the issue is so much that these agents are ineffective, but that they are effective on subsets of patients that we can't readily identify.

And I think that is really the problem practically speaking. That said, I think that the idea of a placebo controlled trial is very appealing. There are some practical problems, and two of them have already been sort of teased out a little bit.

One of them is this issue when you said if we could just exclude the patients who have pneumonia from the AECB trials, and yet we are learning increasingly so that that distinction, at least even based on radiographic evidence, is problematic.

And I think that one could argue that part of the problem is, of course, that the way that we have created these diagnosis based on some relatively arcane tests now is really not the right way to guide therapy.

But nevertheless we are sort of stuck with them for the time being, and we are going to have to

realize these limitations as we start to think about actually giving people placebos.

The other issue is that I would agree that we really have to use clinical outcomes as the measures in these respiratory infections, and there I think we have a disconnect that we are going to have to deal with, in terms of enrollment, and IRB issues, and this escape issue.

And that is this belief that in fact people get better as they complete their therapy, when in fact the observational data would suggest that people's course of recovery is actually quite prolonged.

And I am always sort of amazed by the clinical trial data that suggests the proportion of people who are better by seven days, when you go out and sort of measure this in the real world if you will, and recognize how long it takes for people to get better.

And the consequence of that is that if you are in a trial in which there is a placebo, most people are not going to be better in a shorter or even intermediate period of time.

And so I think there is going to be a lot of emphasis on escape or switch. It is going to be

hard to resist that, unless we sort of significantly change the understanding of what we do know about the natural history of the disease, and which is as I would say in general that it is a lot longer than most people think.

CHAIRMAN RELLER: Dr. Chesney.

DR. CHESNEY: As Dr. Archer pointed out, I also am statistically challenged, and I wanted to ask Dr. Fleming that in your novarepine-AZT example, you called that a superiority trial. How does that differ from -- you know, this is very fundamental I'm sure, but how was that different from a non-inferiority comparison?

DR. FLEMING: Well, the analysis essentially was looking at differences in transmission rates of HIV maternal-to-child, and one of the primary end points was at six weeks. And the novarepine reduced the transmission rate from -- I think it was from 21 percent on AZT, to 11 percent on novarepine.

By achieving statistical superiority, we were able to conclude that single dose novarepine was very effective, and at least provided that 50 percent reduction, possibly more, if short-course AZT was effective.

If in fact those two rates had both been

11 percent, then the difficulty that we would have had is we wouldn't have known whether they were equally effective or equally ineffective.

And that was the loss of not being able to have the placebo arm in that trial. So the only way that study was able to conclusively establish benefit was by having a superiority difference.

If they had been the same, we would not have known if they were equally effective or equally ineffective, because there was no predefined margin that would have allowed us what short-course AZT did.

If we had known that short course AZT halved the transmission rate, and we saw comparable rates between novarepine and AZT, then we could have done a non-inferiority comparison.

CHAIRMAN RELLER: Mark.

DR. GOLDBERGER: Just a couple of things. One is in terms of really thinking about the placebo issue for AECB, it is probably worth you hearing where we are in terms of what is actually being done in trials, and i.e., the big trend in AECB, as it is in some other infections now, is to shorten the duration of therapy.

The last submissions to come into our office I think, one is 5 days of therapy, and I think

there may be one, although with a loner half-life drug, as short as 3 days of therapy.

So in fact in terms of thinking about early escape, we may be somewhat almost pass that if the duration of active therapy is so short. Perhaps the question may have to be in some of these regimens can we say something at the conclusion of the period of when the active drug was given of, of active drug versus placebo, that would sufficiently informative to help us in determining whether that person on placebo ought to receive therapy.

I mention that as an observation. It is just another issue as I see that Dr. Fleming is eager to respond. Well, it is good to see that at this late hour of the afternoon I have to say.

There was some discussion about maybe end points ought to be just keeping a person stable, and I think that if we start thinking about that in at least more severe disease, where at least there may be more comfort antibiotics doing something, I think that is something that may be worth talking about, or thinking about a little bit.

I mean, coming from the old school that the goal of antibiotics and infectious diseases is really to cure or very significantly mitigate

infection, keeping people stable makes me wonder a little bit.

Plus, the illness that we are talking about is acute exacerbation, and if exacerbation means getting worse, you would like to think that something actually could improve things.

And with that, I yield the rest of my time to Dr. Fleming, if that is okay.

DR. FLEMING: Well, just two quick thoughts. First, I would like to distinguish between the time that somebody is on a therapy and the time period over which that administration could affect their outcome.

Somebody might have been on therapy for 3 days, but the influence of that on their outcome might not be fully known until some period of time beyond 3 days.

Secondly, my concern in many clinical settings with looking at end points that are very short term, is that they may be missing the more global and clinically relevant aspect here.

And if we come back to here, and if we are using, for example, 21 day periods for resolution of symptoms, if we look over two days, we may get a relevant comparison over two days, but that may only

be the tip of the iceberg of what really matters to patients.

And I would argue that the clinical endpoints as best possible should capture the essence of what matters to patients, and so that factor should influence as well how long we have to follow.

DR. GOLDBERGER: I would certainly say -- and if you don't mind my taking back the last little nibit of my time, that I would certainly agree with you on the second point.

There is value I think in having some of these longer term outcome measures. Acutely, one might argue that if in fact the duration of antimicrobial therapy is so short that having the early escape at the end of that, there is perhaps a little less worry about giving placebo for only several days. That is what I am sort of wondering about.

Does that pose as much of a problem when we know that the active drug will be terminated at day 3 or day 5, and should we worry therefore as much about the consequences of using placebo.

CHAIRMAN RELLER: I would be interested in hearing perhaps additional comments from IDSA, Dr. Talbot, and others, and from PhRMA, Dr. Shlaes, and

others about where appropriate, should there be -- and leaving aside what indications that those might be.

But should there be greater consideration of the role of placebo controlled trials, looking at the issues that Dr. Chesney pointed out, and I am impressed as the discussions have gone on today that what started out as an emphasis on one thing may be as bringing into consideration that there are a lot of other issues that may help us get to where we want to be, having to do with what is the best way to assess efficacy, and recognize safety in the approval and study of new antimicrobial agents.

Any comments, Dr. Shlaes, or Dr. Talbot, or others?

DR. SHLAES: Well, I mean, I think we would certainly be interested in placebo controlled trial designs, assuming that they were ethical, and that we could carry them out, and that people would accept them.

I think that we mentioned that in our presentation this morning. So I think we are open to that. Obviously, they would have to allow us to carry out the trials in a way that provides meaningful information to all concerned.

But we are certainly interested in looking

at placebo controlled trials, absolutely.

DR. ANDRIOLE: I would be interested in looking at placebo controlled trials in this area, too, because I think the other way to do it is to go with the treatment control, and if you don't measure or don't feel superiority, you are not going to get approval in that area, but the other drug already has it.

It is kind of setting up a straw dog. So I think if it is really a question that that drug has any effect, I think I would rather go to a placebo controlled trial if it could get through an ethics committee.

CHAIRMAN RELLER: Dr. Talbot.

DR. TALBOT: Yes, thank you. There are difficult economic questions perhaps for the development of new drugs, but I think that the societal risk benefit issue requires that the consideration of studies, including placebos, be discussed not only today, but again and again.

And that some solutions be reached so that health care providers in the U.S. can be certain that they are giving effective drugs and not creating a public health risk, in terms of antimicrobial resistance.

Now, I do have to add a disclaimer that I am speaking for myself, and I am not sure that I am speaking for IDSA.

CHAIRMAN RELLER: Dr. Sumaya.

DR. SUMAYA: One issue which I have heard, but maybe not as strong, is where do we focus our energies? Do we focus the energies toward looking at trials in the mild, moderate group of patients, or should we focus major energies on the severely ill?

And can we do that altogether in one trial, or do we have to separate that, or do it in stages, or phases? My prior experiences are that you go to the severe, and then you go to the mild.

In this case, I am not so sure about that, because the mild brings in more things with overuse, potential resistance, but the severe deals with potentially greater mortality issues, and disease burden, and complications.

What I see is that the all-need criteria needs area need to be much better defined, and criteria for entry into any trial, for monitoring during the trial, and for the end points.

So, obviously uniformity , clarity, and those definitions across all those high areas would be very important. If the energies go towards the mild

form, mild to moderate, then I think a placebo control makes very good sense.

If we go more toward the severe forms, then I think some type of comparison, perhaps the standard care as Dr. Fleming had mentioned, versus a test drug, would be the most appropriate.

But again where do we focus the industry focus? Is it a mild to moderate issue, and/or the severe.

CHAIRMAN RELLER: Dr. Ramirez, and then Dr. Leggett.

DR. RAMIREZ: Yes. I think if we have an infectious disease, and the infectious disease is caused by bacteria, antibiotics will always be beneficial.

Then the question is that we know that a patient with mild COPD has bacteria in the airwave, but we don't know if this is an infectious disease. We don't know if bacteria are part of this cycle or inflammatory process.

Then we are asking the industry to define a clinical question. Is a patient with a mild acute exacerbation of COPD having an infectious disease, and are antibiotics necessary.

And I think we are here as doctors, and

the only mention of this is the industry, and there is the agency, and there are clinical investigators. This is a great question for clinical investigators.

Do we need to use antibiotics in patients with mild to acute exacerbations of COPD? But I still don't understand why I need to ask a drug company to generate a new antibiotic and trying to test a basic question to see if a person with a disease requires antibiotics.

I want to ask the industry do answer the question if this person has an infectious disease. I mean, this is not supposed to be the industry. This is supposed to be the clinical investigators answering the question.

Once we find that this is an infectious disease, and the patient has a bacterial infectious disease, then we decide to use a antibiotic. We understand that acute bronchitis is an infectious disease, and is caused by viruses.

And we are not asking the industry to give us antibiotics for acute bronchitis. We just closed the case. The problem is that we don't know if mild exacerbation of chronic bronchitis is still an infectious disease.

CHAIRMAN RELLER: Dr. Leggett.

DR. LEGGETT: One point that maybe I didn't understand, but the most severe definition of the criteria was cough and purulent sputum. I mean, to me that is not very severe.

So in other words, I think that just talking about the COPD patient in the ICU is three standard deviations away from the first that I think of as even having an AECV. Maybe I didn't understand.

CHAIRMAN RELLER: Before moving to question two, does anybody have anything additional they wish to say about acute otitis media or acute sinusitis? Dr. Chesney.

DR. CHESNEY: Just one quick thing. I think I terms of thinking of the natural history, we don't know the natural history of resistant organisms in acute otitis media and sinusitis. And we have good reason to think that it wouldn't be different.

But I just wanted to make that point, that we are dealing with new infections to some degree here by very resistant organisms.

CHAIRMAN RELLER: Dr. Chesney, do you think it is possible to assess efficacy of new or existing agents against resistant pathogens, especially streptococcus pneumoniae, without tympanocentesis puncture studies with sinusitis?

DR. CHESNEY: No.

CHAIRMAN RELLER: Dr. Talbot.

DR. TALBOT: I have one comment about the Chairman's comments about AECB if I could to follow up on Dr. Ramirez's point?

CHAIRMAN RELLER: Please.

DR. TALBOT: I think you raised a very good point. As Dr. Thompson mentioned the current conundrum with AECB is that there is no generally accepted study at this point that definitively proves that active antibiotic therapy is better than no treatment, or placebo.

So let's say theoretically that such a study was done that conformed to all appropriate statistical, and clinical, and regulatory standards. And Antibiotic A was shown to in fact be superior. Would that not potentially obviate the need for successive placebo controlled trials?

Or would the committee think that AECB is inherently more like depression with a lot of variability in presentation and clinical course, such that even after that first demonstration there would be a continued need for placebo controlled trials?

So is it just one that is needed, or does there have to be a uniform and continuing inclusion of

placebo controlled trials? And I don't know if that answer is known at the moment.

CHAIRMAN RELLER: Thank you. Dr. Ebert.

DR. EBERT: I just wanted to comment briefly on the third part of the question on the dose response trials, and I am a little bit unclear as to exactly how that fits in here.

But I think in general I would be somewhat leery about using dose response trials as a measure of efficacy without good pharmokinetic/pharmodynamic data to form the basis for those clinical studies.

And given what we have talked about so far, and the possibilities of drugs being either equal to placebo or not showing a clear definition, I would be a little bit concerned that we may find in a dose ranging that a, quote, subtherapeutic dose does in fact show some clinical efficacy.

And subsequently would just contribute to a use of the drug at that dose, which might lead down the line to resistance because of an in essence a subtherapeutic dose.

CHAIRMAN RELLER: Dr. O'Fallon.

DR. O'FALLON: I am concerned about the fact that we really aren't talking much about the possibility that a treatment can actually be damaging.

And I am thinking not so much about a short term basis, but rather say that a treatment clears things out, but then it leaves a patient at risk to have a prompt recurrence, or a more difficult recurrence, or something of that sort.

I mean, I don't know the diseases well enough, but that in planning these studies, we should be open to the idea that actually a treatment might be damaging.

Now, the second thing is about the dose response or something. You know, they can do 3 and 4 arm studies, with one of them being a placebo, and the others being 2 or 3 supposedly active, and ones that are believed to be active therapies and that can be done in one.

That would probably require the industry to cooperate, but they could get it done. They could get a lot more done with one study perhaps, one large one.

CHAIRMAN RELLER: Thank you. The last comment on question one.

DR. FINK: This relates probably to many studies of lung disease. I think you have to be careful when we start talking about the value of PK/PD data. It is almost always blood levels, and

penetration into airways and the lung parenchyma itself may bear no relationship to PK/PD data from the blood.

So I think if we are really going to try and use PK/PD data to extrapolate, you would have to talk about doing it in experimental animals where you have bleed them out, and then sacrificed them, and actually measured tissue penetration and clearance from the lung tissue itself, which has rarely, if ever, been done.

CHAIRMAN RELLER: Thank you.

DR. GOLDBERGER: Dr. Reller, would you want to summarize or attempt to summarize what you have heard as to question one? It is always a big help for us.

CHAIRMAN RELLER: Actually, Mark, I was gong to propose that as the transition sentence and do next number two. We were asked to comment in relation to the proposed approach for selection of delta and non-inferiority (equivalence, clinical trials).

And what I have got out of hearing all of this discussion is perhaps as important, or more important, is the delineation in acute exacerbations of chronic bronchitis.

Those patient groups or definitions of disease -- and it may well be all such patients with acute exacerbations of chronic bronchitis -- is consideration of placebo controlled trials, or superiority trial design, and not the equivalence design in the first place, and so rather than getting constrained by what should the delta be, is to consider the nature of the trial design in the first place in which patients are included.

And secondly that for other respiratory tract infections, especially acute otitis media, and sinusitis, that smaller numbers of patients with knowing exactly what you start out with, and what you end up with, with the importance of emerging resistance, would be far more useful in delineating efficacy of new compounds, including ones for resistant organisms, than the discussion of -- and not that it is not important.

But again spending the emphasis on the delta and power in non-inferiority equivalence trials. Or to put it another way, that the precise entity being studied, and what is the best trial design, and what would be reasonable assurance of efficacy in the first place, may be more productive than simple discussion of delta.

Question Number 2. Please discuss implication of the choice of deltas in clinical trials for serious infections, and include in our discussions efficacy of new drug compared with currently available measurements for hospital-acquired pneumonia and meningitis.

And I think that these entities are so different as has been amply pointed out that we should consider them separately. So let's take perhaps the more -- well, let's just take meningitis first.

Trial design for meningitis, where one of the messages that we heard clearly from IDSA, and from industry, from Dr. McCracken's presentation, is there is a very serious clinical entity with grave consequences, the number of patients involved is small, and some of the design considerations with the non-inferiority trials for a level of confidence looking at clinical outcomes would require numbers of patients that are either clinically or economically, or both, not reasonable.

So how do we assess with meningitis? What is the most efficient approach to establish efficacy and safety with new drug development? Comments from the committee. Dr. Bell.

DR. BELL: I think the best insight I

heard was your comment earlier about perhaps using different deltas for different outcome variables if I heard that right; microbiologic, clinical.

You know, I am very concerned that when you have a serious infection that you don't want the comparator drug to be much less effective than the standard.

CHAIRMAN RELLER: Now, one way of perhaps getting at this and zeroing in on it is that one could talk about whether it is at 24 hours, or 48 hours, or both, and is the committee, the ISDA, PhRMA, others, are we in agreement that unless one can sterilize the CSF, one doesn't have a drug for meningitis?

DR. SHLAES: Yes. We actually talked about this over lunch, and we were saying that if in fact you had a drug that didn't do that, then probably you would stop development pretty quickly.

CHAIRMAN RELLER: What beyond that -- and the precise numbers, and timing, and how to assess that could be -- those details could be worked out.

I mean, it would require obviously not only an initial diagnostic effort, and you would have to have a repeat lumbar puncture, and assure adequate microbiology that people would accept as being rigorous, decent, and something akin to the

tympanocentesis -- TCs -- and sinus punctures.

But having that, what in addition, in terms of trial design, follow-up, numbers of patients, deltas, would be wise to have? Dr. Ramirez.

DR. RAMIREZ: In prior meetings, we were discussing sometimes the lack of correlation of microbiological resistance and clinical deterioration. And we always blame the consideration that we did new composition in the lungs, and any antibiotic gets good penetration in the lung.

And in the presentation this morning, Dr. McCracken mentioned that quinolones for meningitis is going to be a reality, and the reality is because of streptococcal pneumonia resistant to penicillin.

We tend to agree that this is the area where we are going to see the single failures, and our pediatricians are telling us that they are failures with cephalosporin, and there has been some delay with vancomycin.

At least in our Children's Hospital now the empiric therapies is cephalosporin, vancomycin, and rifampin, until you prove that the pneumococci infection has been resolved.

Now, we have the quinolones, and the quinolones are supposed to have good penetration and

are supposed to have good activity against the streptococcal pneumonia.

And to me this is the idea situation to prove superiority. I mean, you cannot be as bad at the third generation cepthalosporins, again in resistant pneumococci, because otherwise, why try the quinolones.

I mean, to me this type of trials is trying to achieve superiority and resolve the problem of the delta, and resolve the problem of the number of patients, I think we should look for superiority in trials of meningitis in pediatrics and looking for the pneumococci resistance, because this is why we want to use the quinolones in pediatrics.

CHAIRMAN RELLER: Dr. Archer.

DR. ARCHER: I would like to kind of raise another issue. As was brought up, as the rate incidence of meningitis decreases in this country with vaccinations and so forth, and in virtually all of the cases are recruited from abroad, it may have increasingly less relevance for what we do in this country, in terms of practicing medicine.

It may in fact be that bacterial endocarditis might be a better example of a rare infection that meets the same criteria that in fact

meningitis does.

That is, that you have got bacteriologic end points, and you have got clinical end points that are very clear. It is a disease that is not decreasing in this country and you probably could enroll enough patients just in this country alone with our standard of care to affect a disease that would be relevant.

That is, we will continue to see it; as opposed to meningitis, which we hope will become less and less relevant. So as a paradigm, endocarditis might actually be a better paradigm for this kind of delta consideration than meningitis.

And I wondered if anybody from industry had any comments about that?

CHAIRMAN RELLER: Dr. Hardalo.

DR. HARDALO: Two points. First, about meningitis. I think as Dr. McCracken adequately pointed out, there are certain factors that are beyond the control of the treating physician, not the least of which is the duration of symptoms before the onset of effective therapy.

In order to prove superiority for any other outcome other than bacterial eradication, we need to have some clarity as to how do we standardize

the populations that we are studying so that when we look at differences from the time of symptom onset to the time of initiation and treatment that we can compare the end points.

It would be useless to compare drugs if one study population had a delay in treatment of four days, and another study population had a delay of one day or one hour.

And you will never be able to do a reasonable comparison of the superiority trial in that type of a setting. The second would be that although I would like to believe that pneumococcal disease is going away in the United States, he did show evidence that it clearly is not, even with the advent of vaccines.

The only thing that vaccines really have done is reduced H. influenzae, but not necessarily taken care of some of the other pneumococcal diseases. So it will have to be something that we do study in the United States, as well as rely on data from our colleagues abroad.

And I think there it really becomes again an issue of the training of the investigators, understanding what is reasonable natural history of the disease, and criteria for discontinuation, as well

as criteria defining failure.

As was said, there are certain aspects of the natural history that some investigators feel are failure, but clearly are not. So I think it is a standardization with input from the key stakeholders like IDSA, and specialists in pediatrics, to understand what should be the entry criteria, and getting into the study, and what are the definitions for treatment, failure, or progression, so that when you go to a superiority design, we are all talking the same language, in terms of being able to determine efficacy and superiority.

For endocarditis, I agree. The time has come that we need to look at the same types of cidal therapy for determining drugs that are better than what we currently have in the armamentarium.

But again it is distinguishing the inflammatory sequelae of disease from bacterial eradication, and asking not only to demonstrate that you have sterilized the blood stream, but somehow that sterilization has some impact on the long term natural history for that patient.

And picking the most relevant clinical criteria, and the most relevant time points for that determination. And I don't think we have yet

determined what that may be, and that may be a subject for a workshop.

CHAIRMAN RELLER: A lot more is known about this criterion. I mean, clearly one would in the endocarditis trial nowadays need patients who were entered to have transesophageal echocardiograms so that you could even out those who had valve ring asepses, and those persons who came to surgery.

And in addition, you know, time to sterilization of blood, and follow-up afterwards. Dr. Archer, along those lines, if you were to design such a trial with four weeks and six weeks of therapy, and with all of those other things, and adequate training of investigators, consistency in entry, to have some reasonable assessment after therapy of cure, would you not allow -- and sometimes what is done is the oral suppressive therapy after a rigorous cidal regimen?

DR. ARCHER: Well, that certainly could be part of any kind of a study. I think that everything is wide open. I think Dr. McCracken's point though about bacteriological eradication, versus sequelae, that are irrelevant to the antibiotic.

I mean, flipping emboli from a vegetation after the vegetation is sterile or a valve leaflet rupturing after the vegetation is sterile, are not

really necessarily indications of the efficacy of an antibiotic, and now well it worked in sterilizing the disease.

And so I think just like meningitis, I think the bacteriological end points, which are easily measurable in terms of sterilizing the vegetation, are very good surrogate end points in endocarditis, just like meningitis, and might be equally accessible to therapy and therapeutic measurement, and delta calculations.

And I think the issue of oral therapy, and the issue of the length of therapy, there is a whole bunch of things that need to be tackled, and with new antibiotics coming out which are potentially more bactericidal, and might even shorten the course of therapy, I think is the opportunity to do that now.

CHAIRMAN RELLER: Wouldn't one -- there clearly are issues that affect clinical outcome that are not -- they may be related, but they would not be a reason to discount an effective drug for sterilization in the spinal fluid, or a vegetation in the bacteremia associated with endocarditis.

But shouldn't those differences be evened out if there were really good design and randomization to treatment arms? That is, those patients with

delayed therapy, and enrolled in an meningitis study, and those who had embolic complications, or who came to surgery because of either failure, or vegetation size, or emboli, or whatever happened?

But wouldn't that even out if you had proper randomization? Yes? Dr. Glode.

DR. GLODE: But you have the example of that in the study presented by Dr. McCracken, where again I think you would have to prioritize whichever agency was advising approval.

You would have to prioritize those outcomes. So if you look just at his example, then he had bacteriologic success, and could have had a delta of 5 percent, and that would have flown, passed, right?

But on clinical success, which should have again by what you just said, by randomizing people to ceftriaxone, or trovafloxacin, you should have randomized appropriately in the mean duration of symptoms prior to therapy was the same, and the two groups, et cetera.

So clinical success should have been the same if you are assuming right that neurologic sequelae were independent of antibiotic other than duration prior to therapy.

But it didn't, and so it fails the 15 percent delta on clinical success. So if you are the committee, and it passes the 5 percent on microbiologic, but it fails the 15 percent on clinical, then what are you left with?

You have to say, well, I guess microbiologic is more important, and I don't know why it came out differently. It should have come out the same.

But do you see that by putting those extra end points that you have to prioritize which ones are more important to you than the other ones?

CHAIRMAN RELLER: I also got the impression in his presentation that there were some questions about the quality of the data collected at different sites, and that's why I put the emphasis on proper randomization and control, et cetera.

Now, we have lots of hands. This really opened up the discussion, which is great. Let me try to go in a reasonable order. Dr. Nelson, Dr. Shlaes, Dr. Talbot, and Dr. Wittes, and there will be others, but that is a start.

DR. NELSON: Well, two quick comments. What I took away from the fact that 11 were on the investigational agent and two were on the control

agent, is that there was an adequate block randomization by study sites, and that they would have had to somehow control.

But if there were 6 and 5, or 6 and 7, then that might have fallen out as not being an issue. One question again, and not that I would like to be educated on, but this notion that a surrogate criteria of bacteriological clearance, is there any evidence at all that the relationship of a particular drug if it has a different mechanism of action, of its cidal action, could induce a different inflammatory response that could be qualitatively different from patient to patient, to where one would then assume no relationship between the surrogate marker of the bacteriological clearance, and the eventual clinical outcome just based on the host response?

Is that possible, or is there any evidence to suggest -- you know, same bug, different drug, different inflammatory response, depending on the drug?

CHAIRMAN RELLER: Well, you know, I think there may be, and in one of the issues clearly there may be differences in safety, going back to some very old studies and some quite provocative titles, like "With Endocarditis, Dead or Dead," and titles to early

clinical trials. Dr. Shlaes.

DR. SHLAES: I just wanted to bring us back to the reason, as you were trying to point out, I think, as to how we got to microbiological end points for meningitis, and now for endocarditis, at the beginning.

And that was to make the trials doable with diseases that are very severe, but have very low incidence. So it is clear to me from the discussions this morning that the way we got there was to use surrogate markers, such as microbiological efficacy, to allow you to enroll a smaller number of patients, and you would sacrifice therefore a number of the clinical end points that you would normally use to be able to use the surrogate end point, which you have confidence.

And certainly in the case of meningitis, and I think as Gordon Archer pointed out, probably in the case of endocarditis, where you have confidence that the microbiological eradication would be correlated with clinical outcome in some reasonable sense.

So I think that that is still a very reasonable approach to these diseases which are severe, or where the incidence is small, and we must

keep the trial size small in order to actually be able to practically carry out the trial.

CHAIRMAN RELLER: Or put another way, that some of these entities to have smaller number of patients that are well studied may provide more useful information than a larger number of patients, where the quality of recruitment, and the quality of follow-up, and the rigor of the randomization, et cetera, is not there.

Now, Dr. Talbot was next, and then Dr. Wittes, and then we will get a fresh list. I cannot handle more than four at once. Dr. Talbot.

DR. TALBOT: Thank you. I have two comments, one on behalf of Dr. Edwards, who sends his regrets that he had to leave. His comment was that IDSA wishes to emphasize that its clinicians, even right now, are limited in their therapeutic options for some very serious illnesses, such as meningitis, endocarditis, fungal diseases.

So from a clinical perspective, and as front line people in the battle against infections, I think the IDSA membership feels that this is an acute problem and that's why we are here.

But certainly the IDSA would like to see some meaningful progress today. So, I have tried to

distill a little bit what I have heard. The last time I did this, Bill Craig told me that it was his job as Chairman, and not my job as participant, but I am going to risk it anyway, Barth, if you don't mind.

You know, the issue with these serious diseases, it is exactly as Dr. Shlaes mentioned. And I think with serious illnesses, there are two questions that are critical.

First of all, do regulators and clinicians, and pharmaceutical companies, want data on how drugs work in these diseases. The answer is yes.

The second question is do these same stakeholders want some certainty, statistical certainty, about the results, and I think the answer is clearly yes, and that has been adequately mentioned already today.

So I think that there are potentially two choices with a fallback position. To allow the studies to be done, one has to change the delta to widen it if necessary, but that is for reasons that we have heard, and not particularly appealing, given that these are illnesses with severe morbidity and mortality.

A second option is to change the end point, but use a strict delta. That is what Dr.

McCracken had mentioned before is what you were saying Dr. Shlaes, and I think given the state of advancement of anti-infective drug development in a moment, that should be feasible for things like meningitis, endocarditis, Dr. Archer, and possibly others.

That would allow you to have statistical certain, but it would require that you have confidence in that end point, and that is where workshop discussions could generate a consensus about whether such end points existed.

Finally, if you had a situation where there was no acceptable surrogate, you might be able to fall back to the GC paradigm perhaps, where you said that if you have a drug that gets 95 percent clinical efficacy in a small subset, 80 to a hundred patients, in a serious infection like meningitis, that is going to be good enough.

So I wondered if -- I hope that overview helps focus the discussion with one hour to go.

CHAIRMAN RELLER: Thank you, George. Dr. Bell.

DR. BELL: I would like to come back to the concept of different deltas for different types of end points -- surrogate versus clinical outcome -- for a couple of reasons.

One of them is that I think that no matter what we or the FDA agree on the clinical community of practicing physicians out there is going to be much more comforted seeing clinical outcome data, than simply surrogate data.

And to promote a drug based solely on surrogate data might become problematic when there is some inevitable reports of failures, or uncertain successes. They will want to see some evidence that clinical outcome actually was better.

I think the place where this has not been the case has been in HIV, where as we were discussing at the break, the viral load now is widely accepted as the surrogate outcome for many good reasons.

But the difference there is that this is a uniformly fatal disease, and where there never was a cure. And so people were happy to use the surrogate outcomes to get the new drugs quicker.

But as we start talking about diseases where there are clinical cures, and it is just a matter of losing the antibiotics, people are going to be very uncomfortable no longer getting information on clinical cures.

And I just wonder if the FDA could take -- I think it was you, and maybe it was Dr. McCracken,

that different deltas -- well, maybe the delta for the surrogate marker could be much narrower.

And the delta for the clinical one could be greater to deal with the patient accrual problem. But that also eventually there would be something, and if there was some paradoxical and unexpected effect for reasons that we don't understand, this clinical outcome really was worse, and at least there was some framework in place to monitor that.

CHAIRMAN RELLER: David, I had brought that up, and just to follow up your analogy of HIV infection, maybe it is not so dissimilar. I mean, if one has bacterial meningitis, or bacterial endocarditis, with staphylococcus aureus, and there is no sterilization of the blood to vegetation or the CSF, I think there aren't any cures either for practical purposes.

But that does not mean to say that there wouldn't be differences in therapy of drugs that can sterilize the CSF, in terms of rapidity of doing that sequelae with hearing, et cetera, like there are differences in the art therapies with tolerance, and side effects, and other outcome measurements apart from controlling viral replication, and viral load.

So I think that has been brought up, and I

think it is one of the things that has come out of the discussions today that what the emphasis would be in end-points, and Dr. Talbot has pointed that out as well, could be indeed, or probably should be different with the different clinical entities under study.

And exactly what those criteria and their prioritization is, Dr. Glode pointed out every one has recognized at the outset of this meeting all these loose ends. They are not going to be tied up this afternoon.

But the heteriogentity of the appropriate responses I think is a message that is coming across very clearly in today's discussions. Dr. Patterson.

DR. PATTERSON: Well, I would agree that especially in meningitis that you want to know about clinical outcome, as well as bacterial eradication, because for instance you could have an antibiotic that is more rapidly cidal, and with increased cytokine release, more cerebral edema, and it could be better at bacteriologic eradication.

But you might have a worse clinical outcome, and so I think especially for meningitis that you are also interested in clinical outcome, and I think that Dr. McCracken suggestion that at the end of his talk to continue the 300 patients, 20 percent

delta, for clinical outcome, is a good one.

And perhaps then for bacteriologic eradication would be of interest and a smaller delta could be used for that.

CHAIRMAN RELLER: Dr. Fink and Dr. Leggett.

DR. FINK: I was just concerned with Dr. McCracken's comment that I am not sure what the applicability of clinical outcome data in meningitis is when you go overseas to populations where the patients are malnourished, and where 30 percent were HIV infected.

What is the meaning of clinical outcome in that population, when it is so different from what is treated in the United States, that an adverse clinical outcome does not necessarily mean that the drug is bad.

I am worried, because I think clinical outcome is important, but I think if you are going to do measures of clinical outcome that you would at least have to do it in a population that has similar socio-economic status, similar societal status, to that of the United States if you are going to use the results here.

CHAIRMAN RELLER: Dr. Leggett.

DR. LEGGETT: I would like to echo the comments of Dr. Patterson and Dr. Talbot before, who took the words out of my mouth because you would not look this way.

But I would like to point out that setting a rigid delta for things that the drugs can control for, and for things that the drugs cannot control for, seems to me to be fundamentally different.

If we are talking about a bacterial eradication, whether it is endocarditis or meningitis, up near 98 or 99 percent, you could sort of keep this sliding scale, and whether you modify it in this sort of modified Lewis criteria thing or not.

But it seems that there is more noise to your clinical outcomes, whether it is form embolic disease or from cytokine release, that you have to leave room for a larger delta, and for the practicality of doing the studies.

So to affix 10 percent and say that it is 10 percent, no matter what the cause of the difference is, I don't think is going to help us down the road.

CHAIRMAN RELLER: Thank you. Dr. Tally, and then Dr. Hardalo.

DR. TALLY: We have gone through the rationale of studying endocarditis, and indeed we have

a proposal at the FDA right now that we have been talking to them about.

And Gordon is right. There are a couple of drugs coming down the pike that have the characteristics as defined in previous studies on endocarditis that mag be suitable to treat endocarditis.

And particularly the new endocarditis that is now representing approximately 30 to 35 percent, and that is staph aureus. So when you have appropriate models and blood levels, and the initial data to support that you can go into that, then we had been in discussion to look at this.

Now, these are difficult infections, and you need to be in special hospitals, and where you can do the transesophageal to apply the new criteria, and to who does have endocarditis.

But again we have heard around the table that the treatment of this disease is multi-factorial, because you need to have cardiac surgery there, because that is part of the treatment of staph aureus and endocarditis.

And that is not drug driven, and it may be needed initially when the patient presents. It should randomize out, but again what David brought out, and I

think it has been brought out today, these difficult diseases, and that are difficult to study, which have these very hard bacteriological end points, can be studied in a prospective manner, and not to get hung up on the real delta in the beginning.

But to look prospectively and looking very carefully, and I think the two responses that I think you need in endocarditis is to initially bring the endocarditis under control, and sterilize the blood.

That is very hard, and I think if you have not done that in a certain period of time, it is clear cut. It is a failure and the new drug is either going to be equal standard of care therapy rate now or it is not.

And I think we can come to that when we develop that data. The second evaluation does take in these other factors, and the one with the long follow-up is the relapse rate that comes afterwards, and was the drug effective.

And I think you need a good number of patients to say that, but I don't think you need the 500 patient studies. I think you can do it with a smaller number of proven cases of endocarditis. And that is the discussion that we are in now, and I think we could be moving forward to try and answer some of

these questions.

But I think this is the one where you really have to be in dialogue with the regulatory agency and be in dialogue with your investigators, prospectively monitoring very closely to make sure that you don't get in trouble because of the high or deleterious effect of a failure rate is usually in this one severe morbidity and death.

CHAIRMAN RELLER: Dr. Tally, you raised some very important points, and I wanted to ask do you think it is important to emphasize that this quality of investigator, the centers where patients would be recruited and enrolled, would have the capacity to take care of these patients properly.

And with endocarditis, as Dr. Archer mentioned earlier, these are studies -- I mean, they would not be exclusive to the United States, but the United States, and Western Europe -- I mean, these require -- I mean, a standard of care that we would accept requires a sophisticated center where to study fewer patients well may provide better answers than missing data that people aren't going to be able to evaluate at the end of the day.

DR. TALLY: Well, I think if you stick to institutions that are approved for cardiac surgery,

and can do valve replacement, the you already are at a level of care that is a higher standard, I think, then routine hospital care in the United States.

CHAIRMAN RELLER: Dr. Hardalo.

DR. HARDALO: I think all of these things point out the need for developing consensus on exactly -- within clinical outcome, we have heard multiple end points. The hierarchy of those end points from those which are most directly related to anti-bacterial or antimicrobial efficacy, and down to those which are more related to anti-inflammatory treatments or other sequelae of the disease.

In endocarditis, we have heard embolism, immune complex disease, other sequelae which have little or nothing to do with the anti-bacterial clearance of the infection, and that has a lot to do with the duration of disease and prior underlying history for that particular patient.

Indeed, the need for cardiac surgery may not necessarily have anything to do with antibacterial therapy. It may have to do with other host factors. For meningitis, clearly there is a difference in terms of when you do your clinical outcome, but in what kinds of patients.

I am sure as the pediatricians in the

group can say, that getting an auditory test on a 2 year old, and trying to get a reasonable indicator of whether you have auditory sequelae, is quite different than trying to get one on an eight year old.

And trying to interpret that as you follow the patient over six weeks, and six months, can lead a certain amount of noise in interpreting the results. And so you have to have some consensus on how much noise you are going to allow based on the populations you have tried to study.

Certainly the efforts by the industry as the information becomes much more critical, and as these patient populations become much smaller, is to really go through extensive efforts to qualify your investigators.

It is no longer the standard just to take all-comers who want to do critical investigations. We have been held to an increasingly high standard in good clinical practices for exactly this reason.

We want to believe the data at the end of the day that we have put so much into developing the protocol, and there is so much resting on this in terms of delivering good quality data to our clinicians.

CHAIRMAN RELLER: Thank you. Dr. Ramirez,

and then Dr. Ebert.

DR. RAMIREZ: I just would like to emphasize what you just mentioned. This is critical. Plenty of the discussions of the patients at the end of the trial will have data to evaluate is because of the investigators.

And really I can summarize the meningitis presentation by Dr. McCracken, and say that he has a problem with an investigator. There was a bias against the quinolones, and every patient that was on quinolones was a failure.

I mean, there was not a problem of the assignment of the trial, and we don't need to increase the delta. We just need to change the investigator. But the study is supposed to be blind, and how come investigators are going to know that my patient with meningitis was getting quinolones, or is getting the standard therapy?

But essentially we just need to have good investigators. I think that in this regard really we don't need to blame the FDA. We just need to blame the industry and with an intention to get patients in empirical trials.

I mean, I liked what you just mentioned, highest standards for investigator, but what we see in

different universities around the country is that the clinical trials are no longer there.

The clinical trials go to the very busy private physician, who has a nurse running around and drawing everybody. We can't even say that these are bad investigators when there are no investigations to begin with.

And another thing is that I don't think we need to travel all over the country to find bad investigators. I mean, we can do it at the center trials here. And why is it that we are having such a poor quality in our research? It is probably because we are not selecting good investigators.

DR. HARDALO: I would really want to argue with that. Part of the reason is that when you have to do a trial of 2,000 patients in the United States, especially if these patients can have no prior antibiotic therapy, and especially you want to get resistant pathogens, you are not going to find them in the United States or in many areas of Western Europe.

And that has been shown in time after time when you look at the trials that are enrolled. Again, we would love to work with United States centers, but some of the realities of making a trial feasible requires us to go outside of the country.

And you are absolutely right. The investigator selection issue, it is a monitoring issue, and we can do what we can in real life. But the investigators are clinicians and who also have their obligations to do trials according to good clinical practices.

CHAIRMAN RELLER: Dr. Ebert.

CHAIRMAN RELLER: This is a comment I think more of surrogate outcomes in general, but certainly the examples that were used regarding microbiology I think were very compelling.

But I think something that as we start to develop surrogate outcomes for other diseases and try to use those in lieu of clinical outcome, we need to keep in mind that as we try to reduce the delta for the use of these clinical outcomes, or excuse me, these surrogate outcomes, we need to be sure that those surrogate outcomes are achieved at a fairly high level.

In other words, a very high percentage. For example, the sterilization of nearly a hundred percent. If the frequency at which these surrogate outcomes is achieved is at a lower level or similar to clinical outcomes with regard to the frequency, I don't think we have really accomplished anything, and

using the small delta is just going to drive up the sample size again.

CHAIRMAN RELLER: Dr. Metlay and then we will -- we included infective endocarditis, which was not in the question, but I think some very important points have been raised related thereto for future drug development.

And then we need to have any comments, if there be any, for hospital-acquired pneumonia before going to the final, but shorter, third question before concluding at 5:30. Dr. Metlay.

DR. METLAY: I guess what I am struggling with to some extent is to what degree do these surrogate end points, bacteriological eradication, really are a solution, or just an occasional exception to the rule.

One of the insights, for example, in the last couple of years, and perhaps relevant in the treatment of community-acquired pneumonia, is that therapy within 8 hours saves lives.

It seems plausible to me that if we were measuring bacterial eradication at 24 hours, for example, or even 48 hours, that we would fail to detect benefits of some therapies, or some strategies within that kind of a window, because our measure is

not sensitive enough.

It is not inherently the case that bacterial eradication is a more sensitive measure for the efficacy of the drug given that in the end what we are interested in are patient outcomes.

So I think that there are lots of applications of meningitis, and in some ways like an ideal one, but I think how well that would generalize and get you to a lot of other solutions is not clear to me at all.

CHAIRMAN RELLER: Dr. Cross.

DR. CROSS: Well, just as a follow-up to that, in certain disease processes, especially infections with bacteria in sterile sites, a prerequisite is that you have to clear the site of infection.

In the case of pneumonia, it is a lot more complicated pathophysiology of which the clearance of bacteria perhaps is only a small point. But I would agree with Steve's comments that if we do have a surrogate end point -- and so far the only surrogate end points that I have heard have been bacterial clearance from sterile sites has to be very high.

But to reemphasize a point that Jan made, perhaps there ought to be serious consideration given

to the clinical outcomes in the situation where the antibiotic itself does have an effect on the inflammatory response.

And Dr. McCracken mentioned about the inflammatory response in meningitis, but also there has been perhaps more made out of it than it ought to be.

But people have tried to compare differences in, for example, ceptazam (phonetic) versus enepenam (phonetic), both of which can clear the blood of GRAM-negatives very rapidly, but one of which may liberate in the process of that killing a pro inflammatory agent more than the other.

So in that situation, I think on the one hand we can have a small delta for the clearance of the bacteria, which is a prerequisite, but on the other hand, I think we still ought to allow for some potential differences from a difference which may arise not as a result of the pathophysiology of the disease which we might not know anything about. but because of the mechanism by which that antibiotic may work.

CHAIRMAN RELLER: Dr. Chesney.

DR. CHESNEY: Just two quick comments. I

think George made the comment this morning that sterilization of the middle ear correlates very well with clinical outcome, and I think that is something that we have just learned in the last few years.

The other thing is that I just wanted put a little bit of a plug in here for quality of investigators. In terms of the NIH having put so much money into the PPRUs, which are the Pediatric Pharmacokinetic Research Unit, that some of you may not know about.

But these are wonderful research units -- I think there are 13 in the country -- that have been set up exclusively to study drugs in children and to maintain that, and set the standard for that kind of quality. So I think as pediatricians that we would like to thank them at every opportunity that we get.

CHAIRMAN RELLER: Contributions to the discussion for hospital-acquired pneumonia. Dr. Archer.

DR. ARCHER: I would like to start this off again. As a comment about the dichotomous nature of these infections, I think somebody mentioned it earlier, but that hospital-acquired pneumonia is an excellent example.

For instance, there is a population of

hospital acquired pneumonia, and people in the VA know about this very well, and in the extended care facilities, patients who develop pneumonia while in the hospital and who don't make it to the ICU, and don't get ventilated.

And post-operative patients developed hospital-acquired pneumonia, and those are very different than the hospital acquired pneumonia patients who are ventilator dependent.

And I think the bacteriology is different, and so I think you could also argue that you could have different populations of hospital-acquired pneumonia patients, some of whom may do better than others as well.

And I don't know that those have been well separated out in studies, at least the studies that I have seen. And a second comment about hospital acquired pneumonia, particularly those in intensive care units, is that it is way too easy to get bacteriology as they are suctioning patients out in Q-5 minutes I think, and a lot of these people are -- and there is bacteria everywhere, and they are cultured frequently.

And I think this is a slippery slope. If you include these in studies, then you have to have

some measure of eradication of the bacteria that are within the spectrum of the drug that you are using.

And I think that is very difficult with hospital-acquired pneumonia, because as has been said, the presence of bacteria don't often correlate, and nor do I think the eradication correlates very well.

And I have not seen a lot of study design where attention is paid to the effect of the drug on the bacteriology of the pneumonia, or the organisms that are recovered from the sucrate.

CHAIRMAN RELLER: The guidelines that were published in the collaborative effort with FDA and IDSA in 1992 were a giant step forward from the former days of lower respiratory tract infections when it was delineated as community-acquired pneumonia, hospital-acquired pneumonia.

What I do not recall, and maybe a further distinction is necessary and an important message to send from this committee to the next iteration is the separation in hospital acquired pneumonia into those patients who are intubated and those who are not.

I don't think that currently exists in the hospital-acquired pneumonia guidelines. Correct me if I am wrong.

DR. ROTSTEIN: Well, there is a

differentiation that the ATS had based on organisms, the types of organisms that people would have, and whether they were admitted to the ICU with hypotension, et cetera. So the ATS does differentiate somewhat.

CHAIRMAN RELLER: Right, the ATS, but in the guidelines, the points to consider documents, Dr. Albrecht, currently the agency does not make that distinction in clinical trial design?

DR. ALBRECHT: It is correct that we don't have a separate guidance for ventilator-assisted, or associated pneumonia. I think there is mention of it, but not a separation at this point.

CHAIRMAN RELLER: Because maybe we would -- you know, in addition to, and apart from the delta, if the committee thinks that is an important distinction to make, in terms of evaluation of clinical outcome, or bacteriologic outcome -- I mean, outcomes, whatever the end points are, we should get that point across clearly. Dr. Ramirez.

DR. RAMIREZ: My opinion is that there is a significant difference. I mean, pneumonia is a continuation of disease from community ambulatory care, to the patient who is going to be in intensive care unit and on a ventilator.

And there is definitely a continuation of the disease in study, after study, after study indicated, that early nosocomial pneumonia -- and how you define early in different investigations is defined differently in days.

But 5 days, or 7 days, whatever is the definition of early, early nosocomial pneumonia, you look at the pathogens, and they are exactly the same pathogens that communicate community pneumonia.

The patient is in the hospital for X-amount of days, and develops nosocomial pneumonia, and at least in our hospital guidelines, we don't use anti-nosocomial regimen, because these patients are going to have H. flu, streptococcal pneumonia.

These people don't have the time in the hospital to be colonized with the nosocomial resistant pathogens. In early nosocomial pneumonia, in any studies from Europe -- and in our intensive care unit, we have a trauma unit.

And if you go to the unit, you are on a ventilator. You develop pneumonia, and you have early nosocomial pneumonia, bronchial, or haemophilus influenzae, number one.

If you are smoker, you have early nosocomial pneumonia, and you don't need -- there is

no question that nosocomial pneumonia is a single disease. It is different.

Now, here you have multiple medical co-morbidities, and you are in the unit, and you have been in the hospital for 2 weeks. There is no question this patient is going to be colonized with whatever organisms is living in your hospital.

And to me the distinction of early development of nosocomial pneumonia versus other organisms, these are two different pathologies. This is one person with a community organism versus another person.

And another thing I would like to say since I have the microphone is that in the delta-1 question in nosocomial pneumonia, and the two studies that were presented, one was 90 percent mortality with placebo, and the other with 10 percent mortality, if we don't have data for one disease, I think we have to look at similar diseases and translate the data.

We know that in community-contacted pneumonia and the pre-antibiotic era that you have bacteremia pneumococcal pneumonia, and there was 80 percent mortality.

And then intuitively, I would agree with the 90 percent mortality. With nosocomial pneumonia,

you don't use antibiotics. If we know that you have nosocomial pneumonia, and you don't use any antibiotics, then I would not say that only 10 percent benefit for antibiotics.

It would be more towards 80 or 90 percent benefit, or probably even 100 percent benefit with antibiotics compared to placebo. Then I would resolve the delta-1 question with this.

Now, the delta-2 question is the question that we have been discussing, and the problem with nosocomial pneumonia for delta-2 is that the problem is not a problem with the drug. It is the problem with the clinical diagnosis.

In any clinical trial, approximately 50 percent of the patients don't have nosocomial pneumonia. And then this is the problem, because 50 percent of the patients, it doesn't matter whatever you use, they are just going to have the natural cause or the ARDS, or whatever other disease they have that we call pneumonia, because we don't have any better way to make the diagnosis, and the delta-2, I don't know how to resolve the problem.

CHAIRMAN RELLER: Dr. Patterson and then Dr. Leggett.

DR. PATTERSON: Okay. I would like to come

back to a point that Dr. Powers made in his presentation that looking at overall mortality I think is not the right outcome in hospital-acquired pneumonia, because there are a lot of other things obviously that these people die from.

And so l think looking at attributable mortality, although that is sometimes difficult to tease out, would be a much more important outcome to look at. But overall mortality, I think wouldn't be the right outcome.

And then also based on a Fagan study that showed improvement in outcome in people who were diagnosed with the associated pneumonia with a protected specimen brush, versus those who were empirically treated based on what was in their sputum and sort of the traditional way of diagnosing it, what are the critical care people think about using the protected specimen brush with quantitative culture more in the setting of diagnosing and studying ventilator-associated pneumonia?

CHAIRMAN RELLER: Yes, please?

DR. ROTSTEIN: Just another comment about pneumonia, and hospital acquired pneumonia. This is one area that we really could look at resistance, because this is where resistance occurs.

These people are often on multiple antibiotics over prolonged periods of time, and this is where we see our resistant organism. So any trial that does look at this really should look at resistance issues as well.

CHAIRMAN RELLER: Thank you. Dr. Shlaes, and then Dr. Leggett.

DR. SHLAES: Actually, I just wanted to comment that I think that this particular area of hospital-acquired pneumonia is the most difficult of the areas that the committee is considering, and that the FDA is considering.

And because of the heterogeneity of the population included in this umbrella, and in addition, actually the CDC is thinking about changing their definitions, in terms of what is community-acquired and what is hospital-acquired.

I am hoping that the CDC is talking to the FDA about their considerations, and that may help in fact in helping us dissect out these two populations. Actually, there are probably 3 or 4 populations, in nosocomial pneumonia.

And it may be that some of those things that we have been calling nosocomial pneumonia are actually community acquired pneumonia, and would fit

better in the new CDC definitions when they come out.

And that may be an easier way for us to start teasing this apart a little bit. So I really think this is a challenging area, and this is going to require stakeholders that are not just industry, and IDSA, and FDA, but is actually going to require some help from CDC, and perhaps others, to just figure out some of these definitions.

DR. RAMIREZ: The CDC is going to use it after seven days to nosocomial?

DR. SHLAES: I don't know what they are going to do. David is here, and maybe he can tell us what they are going to do. But they are reconsidering their definitions of community-acquired, versus hospital-acquired infection in general.

CHAIRMAN RELLER: I think there are multiple manuscripts from different places under review, and that the data aren't in yet. But basically health care associated infections may look more like nosocomial infections than community-acquired in the strict sense.

And the proportions shifted, and not everybody who comes in from the community has not had recent association or be it extended care. But I think the issues are very important.

And that the definition of community acquired pneumonia and hospital-acquired pneumonia will need some redefinition, and including the ventilator, and those complicated by the need for ventilatory assistance fall into a different category, in terms of expected response, and distribution of pathogens. Yes?

DR. CROSS: I think that in hospital-acquired pneumonia the bacteriology in this will be a real bear and has to be really clearly defined. I think as has been said that the bacteriology of ventilator-associated pneumonia is quite different.

But the other thing to consider, especially as we talk about hospital-acquired, and community-acquired, is the rather extensive, and very formidable data from 20 years ago looking at the role of underlying illness, in terms of colonization with GRAM-negative criteria.

For example, on day one of entry into the ICU, J. Sanford and Reiner showed about a quarter of the patients are already colonized with GRAM-negative bacteria.

Similarly, the classic studies of Valenti showed that the likelihood of colonization with GRAM-negative bacteria, even people walking in off the

street, is a function of their underlying health status.

Therefore, it isn't a simple breakdown to say that people who are in the less than 48 hours, or 96 hours, will have a certain amount of or certain types of bacteria, in the absence of actually defining those critical factors which have already been well-defined in terms of health status, and bacteriology.

CHAIRMAN RELLER: So you are getting at the importance of this attributable mortality issue as well. Dr. Ramirez, and then we will have a comment from the back. Go ahead.

DR. RAMIREZ: I just want to clarify that when I mentioned the early versus late -- and I totally agree with the GRAM-negatives -- is that people can come from home with klebsiella, E. coli, and they have multi-medical co-mobilities.

But the multi-resistant pseudomonas, you are going to get in the hospital. Another thing is that sometimes when we see studies done for the drug companies, they want to test this particular drug against the others.

We have seen in ciprofloxacin versus emipenam, and in all the latest studies of nosocomial pneumonia. But in reality, what I see happening in

critical care units is that the patient may have ventilator-associated pneumonia, and we suspect pseudomonas, and the tendency is to use combination therapy.

And the problem that I sometimes discuss with industry is that we don't want to use antibiotics. I just want my antibiotic. But we are using more and more combination therapy in an attempt to prevent the development of resistance and improved outcome.

Wouldn't it be more realistic to do studies of combination therapy based on ventilator-associated pneumonia, and with the more severe form of nosocomial pneumonia?

CHAIRMAN RELLER: I think there are big differences in terms of Western Europe and the United States and those who believe in the importance of quantitative cultures from bronchoscopy specimens. I know that in our own center there are brushers and non-brushers, believers and non-believers.

And I think that one of the messages that comes across is before a discussion of deltas, that one has to spend considerably more time in delineating what it is that we are talking about with hospital-acquired pneumonias as a prelude to a meaningful

discussion of what kind of equivalence or non-inferiority trials, and what the numbers should be.

I need some help from those who wish to make comments who are not seated around the table with their nameplates. So, please introduce yourself, and then comment.

DR. SCHENTAG: Hi, I'm Jerry Schentag from the University of Buffalo. I am presenting the triad of people who harass you folks with PK/PD type comments, but I am the only one here today.

So I felt obligated to speak, and I think on this nosocomial pneumonia thing, if you do a multiple logistic regression analysis, and include all the clinical factors that you can dig up on nosocomial pneumonia patients, and you add to it the activity of the antibiotic.

And then you plot that against how long it takes to kill the bacteria -- and not whether or not you kill it, but how long it takes to kill that bacteria on serial culturing.

And if you do the serial culturing, you can get about 80 percent of the variance in the relationship killing that organism over time just from the antibiotic activity, leaving about 20 percent of the remaining variance in that logistic regression to

be explained by the other factors.

Now, I agree with you that this is not an easy scenario to assign which one the pathogen is when there is lots of organisms, and there is lots of drugs, but it is relatively easy to assign an outcome to that organism, which I do believe from studying this now for quite a few years of multiple different antibiotics, and we looked at maybe 15 or 20 antibiotics this way over the last 10 or 15 years.

And I do believe that you could show differences between concentrations to activity ratios of each of those drugs, which makes sense. In other words, it is the activity of the drug that determines the microbial outcome.

What I don't know is whether it always determines whether you perceive the surrogate end point of cure to follow that or not. And ventilator-associated pneumoniaes, it probably does reasonably well.

In the non-ventilator associated pneumoniaes, it is probably like a lot of other pneumoniaes; cures don't always follow eradication of the organism. There are other factors that aren't quite so closely linked.

But cure is nonetheless the surrogate,

because the effect of the antibiotic is on the bacteria. Now, have we been able to find any evidence of endotoxin storm or any of those other things contributing to outcome?

Well, my submission on that point is that we have tried awful hard with the sepsis drugs, and we haven't been able to show much of an additive benefit, and we just managed to find a small one not so long ago, but it is by and large not a dramatic effect if it is there.

Most people would look at all of those trials and agree with that. So I guess my comment is that you shouldn't reject microbial end points so easily as surrogates, given that they can almost always show superiority with very small numbers of patients in each group between two antibiotics, or in fact between combinations of one handful of drugs, versus the other handful of drugs when you want to start looking at that as cumulative activity, just assuming additivity.

Thank you for letting me make that comment. I had to get that off my chest. Jerry Schentag from Buffalo, okay? Just in case.

CHAIRMAN RELLER: For the record, thanks, Jerry. Dr. Bennett, and then we need to move. We

could use Dr. Schenag's comments as a transition to question number three. Dr. Bennett.

DR. BENNETT: I wanted to give Art Goldberger my two cents about deltas since we spent the morning talking about deltas, and very early of the afternoon.

But what I think I have learned is that 10 percent is not for everybody, and not for every trial, and not for every indication. So that a 10 percent delta as a receipt in general is too inflexible.

But I have also heard that the STEP function is also inflexible in a different way, and not very useful. So what I am taking home from this meeting is that you are going to have to come up with guidelines that are specific for indications, and maybe even have some protocol definitions built in.

And then you will be able to get deltas. So your goal of having us bless a given delta, I just don't hear that. And that is why I think we are not talking about it.

DR. GOLDBERGER: That in fact wasn't really the goal of the meeting. I don't think we recognized upon reflection that a fixed delta for everything was necessarily the best way to proceed, which is why I during my introductory remarks made

some points about this as being the beginning of a process, rather than the goal of coming up with a judgment at the end of the day.

So we would agree with your comment, that I think it would be very difficult to squeeze in everything under a single delta.

DR. BENNETT: The only reason that I made the remark the way I did was that I had the impression from conversations in the hall that that is what the FDA had been doing; that is, using a 10 percent delta for many different indications across the board.

And that was raising some appropriate hackles, but that apparently was not correct.

DR. GOLDBERGER: It is fair to say that there was at one point what I would describe as a communication breakdown, which hopefully we have satisfactorily rectified with regards to that.

I wouldn't want to say that those people who were upset were upset entirely based only on their imagination, because I don't think that is a fair statement.

But I think we recognized that this was in fact not the preferred way to proceed, which was the reason for trying to get as broad an input as possible, for instance, at today's meeting.

DR. BENNETT: Thank you for clarifying.

CHAIRMAN RELLER: Question Number 3. The FDA announced that they were not going to slavishly follow the STEP wise. What they were going to do was perhaps prematurely anticipated.

But what Dr. Bennett just summarized, is where I think the parties at this meeting are fairly concluding, is the reality that there must be a diversity in what goes into a fair assessment, and realistic assessment, of efficacy balanced off with safety of anti-infective compounds, and that will be different by different indications, and other very important issues need to be addressed explicitly.

And in some cases, objective end points; and in others, a redefinition of what constitutes the appropriate study populations with, for example, hospital-acquired pneumonia.

Question Number 3. Discuss any other factors or characteristics of a drug product other than the confidence intervals, the deltas, that could be included in risk-benefit analysis supporting FDA regulatory decisions.

Now, actually, these things have already come up in the discussion. So, it would be in addition to what has already been said.

Any comments about safety considerations, PK/PD considerations, and the availability of alternative therapies in this balance of safety and efficacy which is the fundamental basis for regulatory approval? So, Dr. Fink, Dr. Glode, Dr. Shlaes.

DR. FINK: I am not going to address safety considerations, but I think the one thing that is glaringly missing from that list is patient acceptability.

Ease of administration, perceived burden of the administration of the drug; is it once a day, four times a day; does it give you an upset stomach. I can't get a parent to give amoxicillin when it gives their child diarrhea.

So I think you have to really look at what is going on outside of the controlled clinical trial that affects real world adherence to use of the drug in an appropriate manner. And that that needs to be very high on the list of alternate considerations.

CHAIRMAN RELLER: Dr. Glode.

DR. GLODE: I will need possibly Dr. Fleming's comments on this as well, having both served on the Vaccine Advisory Committee, and dealing with from the perspective of safety, and how many children do you need in a trial to assure safety.

So I guess I don't know the answer to the question of before a new antibiotic comes to a Phase III efficacy and safety trial, in Phase I and Phase II, how many hundreds or thousands of individuals have been studied for safety?

Because if, for example, you take the meningitis example, where you might use as an end point bacteriologic sterilization of the spinal fluid so you can use very small numbers of people.

Then, you know, you compromise your ability to look at safety issues it seems to me. Now, if they have already been looked at, but it is so detrimental to everyone concerned, starting with the patients, when a drug is withdrawn from the market after approval due to an adverse event that was not recognized during the preclinical trials.

And I was wondering if anybody has gone back and looked at the last 10 drugs removed and sort of asked the question were they adequately studied in the first place?

Well, by the time that a new antibiotic gets to Phase III, is there some approximate number of patients who have received it to assure safety, or are we relying on a Phase III study?

CHAIRMAN RELLER: Comments from the FDA?

In some of these past meetings, I think there has been considerable discussion on some events that the numbers simply can't preclude knowing until a drug is approved.

I know that these issues came up in the electrophysiologic effects QT intervals, and arrhythmias with fluoroquinolones, that there may be some effects that are simply not knowable until actually put into clinical practice. Comments, Dr. Goldberger, or others?

DR. BENNETT: We could to use your example, electrophysiologic effects. You can do a dose escalation study with a drug, 10 or 12 patients per arm, with careful monitoring of QT and establish whether the drug has some effect on QT.

But absent an enormous prolongation, the chances of seeing anything in a clinical trial database of 5,000 people are essentially zero. You are up there probably needing tens of thousands of people in post-marketing databases to see anything, if in fact there is any type of signal, just to use that as an example. Bob may want to add some other things.

DR. TEMPLE: If I understood the question, the question was how much do you know at the end of Phase II. There is an unfortunate idea that you know

a great deal about safety at the end of Phase II, and that is just completely wrong.

If you are lucky, you will have a few hundred patients. Well, you only know a very little bit about safety from that. Phase III, which in antibiotic terms, given multiple indications, will typically have several thousand people, gives you much more assurance about events up to the order of one in a thousand, or something like that.

But what Mark was describing is how we use surrogates for toxicity in fact, a drug that prolongs the QT interval a lot probably won't be approved unless it does something really spectacular.

A drug that causes certain kinds of liver test abnormalities probably won't be approved because we believe certain findings that are not lethal themselves, predict ultimate lethality.

So that all of those things go on, and nonetheless, some slip through, and have to be taken away later. But you only know a very little bit at the end of Phase II because you just can't find out that much in a couple of hundred people about real events.

CHAIRMAN RELLER: Dr. Shlaes.

DR. SHLAES: I just want to make a comment

that PK/PD. I mean, in the anti-bacterial realm, we have had very good animal models, which have been very predictive of general success in the clinic for a very long time.

We have had proposed guidance, I believe, that came from your predecessor, Dr. Reller, who was Dr. Craig, suggesting that PK/PD be used much more in consideration of approval for certain indications.

I think we are going to talk more about this tomorrow. But we have known about this in the anti-bacterial realm a lot longer than the HIV people have known about it.

And yet they have -- and as a matter of fact, I am not sure how much PK/PD they have compared to what we have, in terms of our confidence and ability to predict success.

Yet, they are using it much more routinely compared to us. So I think it is about time that we had a little confidence in the predictability of these animal models, and our ability to do PK/PD to get antibiotics approved, especially for those indications which are difficult because of low patient populations. And I think we are going to talk more about that tomorrow. Thank you.

CHAIRMAN RELLER: Dr. Metlay.

DR. METLAY: Well, I guess I would just add as an extension to that that the whole issue of the impact of the agents on microflora, oral and icteric microflora, I think really very much that we had a lot more data on the impact of cross different drugs, and we have been in cross-classes.

Because I think in the end that a lot of our indications and recommendations are going to ultimately come down to those kinds of considerations. So that we could be better minimizing the impact on resistance emergents.

And I know that is the theme for tomorrow, but it seems to be quite integral in this discussion as well, and I am trying to understand whether there are new compounds out there that really is value added.

CHAIRMAN RELLER: Dr. Cross and Dr. Shlaes.

DR. CROSS: I would just like to follow up on Dr. Shlaes' comment, and just ask as a matter of information, how good are the animal models for lots of the things that we look at?

For example, in the sepsis field, it is accepted that there is no one good model which is predictive of any therapy consensus. I know from

personal work in animal models, for example, that there are very few animals of staff orates.

And that certain organisms, like klebsiella, are not pathogenic in mouth-to-mouth except for one type. So it would be very hard to test the drug for ESBL, for example.

So as a point of information, how good are the animal models, in terms of the PK/PD? Well, I think as you know, you can carry out and do Bill Craig's model, which is the thigh infection model, and get I think very good information on the critical pharmokinetic parameter based on blood levels.

So whether it is AUC, and whether it is peak, and whether it time above MIC, and then you can use that to make predictions, knowing PK and people about what the efficacy will be under various circumstances.

And in fact, Jerry Schentag, and Bill Craig, and others, have carried out studies on people, and you do see very good correlation between the PK/PD predictions that you get from an animal model like the thigh model, and what you see in people.

Sometimes you have to do additional studies on people, and as somebody brought up earlier the issue of drug concentrations in the lung, and in

the ELF. Those studies can now be done in people, and you can get very good PK/PD information in people.

And frequently this does correlate in what you see in analysts. So I think that is on example where those correlations work quite in predicting the kind of doses that you might have to use, and the kind of concentrations that you might have to achieve in people.

CHAIRMAN RELLER: Dr. Ramirez, and then Dr. Soreth.

DR. RAMIREZ: Yes. Regarding my wish list for risk-benefit analysis in clinical trials, I would like to add a better determination of cost of treatment, because at this moment when we have a new antibiotic on the market, the only thing we know is that it is going to be less effective as the old antibiotic for the management of the particular infection that this is.

And then when we are on the P&T committee trying to define what is the most cost effective therapy, if one antibiotic costs $30 and the other costs $25, the one that is most cost effective is the one that costs $25.

And this is because clinical drugs do not allow us to define what is the most cost effective

regime. And I think that matches perfectly with the discussion of looking at other outcomes besides clinical outcome.

I think we need to be looking at other outcomes for costs, and for acute exacerbation of chronic bronchitis was already mentioned, and the time that the patient takes to return to work, and these types of issues need to be in the protocol.

For community-acquired pneumonia, there are large studies which indicated more or less (inaudible), and probably we know the time to (inaudible), and we can define in the hospital/patient time to switch therapy, because we know that switched therapies are associated with early hospital discharge.

And then I don't care too much if the two antibiotics cure the patient the same at 30 days. If the antibiotics decrease the length of a stay for two days, this is going to be the most cost effective, regardless of the cost for the antibiotic.

And for nosocomial-acquired pneumonia, issues such as time of exacerbation of days in the intensive care unit, because a decrease of one day in the intensive care unit is going to be definitely the most cost effective antibiotic for nosocomial-acquired

pneumonia.

And I would like to see incorporated more outcomes that are going to help us physicians when we are admitting the P&T and try to define ways that are the most cost effective antibiotics incorporated in the clinical trials.

CHAIRMAN RELLER: Dr. Soreth.

DR. SORETH: I just wanted to make a comment on safety considerations that Dr. Glode had raised. I think in addition to clinical trials fundamentally not being powered to tell us much about or elucidate much about uncommon adverse events, we also have to recognize that in the clinical trial setting we are studying patients under ideal conditions.

And that the amount of information that we might have in the development program about the use of concomitant medications, about underlying co-morbid conditions, disease states that affect drug metabolism, and excretion, and so forth, can be quite limited.

And once a drug is on the market, and thousands, and hundreds of thousands of patients are exposed under less than ideal conditions, under real conditions -- concomitant meds, states of hydration

varying widely -- only then do we really understand the full safety or toxicity profile of a drug, but unfortunately not at the time of an action.

CHAIRMAN RELLER: I would like to thank those attending -- yes?

DR. YUH: Can I make two comments?

CHAIRMAN RELLER: Please.

DR. YUH: I think we are getting close. My name is Lianng Yuh, and I am representing PhRMA. Actually, I am speaking for myself. I think the sense of urgency is that we would like to know of any interim solutions before we come up with any real good guidance on antibiotic development, because a lot of the companies have experience with different guidance, and I think it has been there about -- longer than a year now.

So we need some interim solutions before we have a better solution. I agree with Dr. Goldberger that we need to welcome different indications, different special cases, to come up with better solutions. But interim solutions are important to us.

Secondly, I would say that any designs we are discussing, hopefully we can also address the concerns from other regions, and not just the United

States or North America, because we tried to harmonize our experiments.

There is a word they say, that patients are waiting. There is a sense of urgency and that we have to move forward. Thank you.

CHAIRMAN RELLER: Dr. Fleming.

DR. FLEMING: Are you still soliciting responses to Issue 3? A resounding yes, I think.

CHAIRMAN RELLER: For you, yes.

DR. FLEMING: Well, I will be brief. I am actually kind of folding my answers to Issue 2 and Issue 3 as well. When I think of the factors that should be considered, I think this is a little bit just stating the obvious.

But I think it is still worth stating, and that is that I am assuming that this question is written with the understanding that in many, if not most, cases the primary confidence interval we are talking about here is on the primary end point, which I would hope would usually be a direct measure of clinical benefit.

And in that context, then certainly other factors that should be considered are secondary measures of clinical benefit, such as hospitalization. And mortality results, safety, tolerability, drug-

drug interactions, will weigh in, as will as we have already heard convenience, acceptability of administration measures.

And I had mentioned this morning in my presentation that when defining margins, if one anticipates substantial differences in issues relating to safety, tolerability, drug-drug interactions, or convenience, those issues in fact could influence the actual final choice of the margin.

External results from interventions that are members of the same class are certainly factors that would be considered. And I mention last, not because it is the least, but because I want to address it separately, are measures of biological activity.

And I have no concern about the fact that clearly they are, such as bacterial eradication, measures that influence your overall sense of strength of evidence of effects having been established.

My concern arises in those settings that we advocate their use in lieu of understanding results about efficacy directly, or results about clinical end points directly; i.e., as a surrogate marker that is a replacement end point.

Just as a reminder of these classical complex issues, one has to understand the disease

process well enough to be confident that this specific measure that you have is really in essence fully capturing the mechanism by which the disease process influences the end point.

And furthermore, one has to be confident that there aren't significant unintended mechanisms of action, anti-inflammatory activities, or other factors, that could influence the critical end points that are not being captured by this marker.

So we run into some fairly complex issues. We have mentioned specifically in question number two that for the specific setting of meningitis the use of the marker because of the fact that there is a quite clear understanding of the biological mechanisms here, and could be an appropriate replacement for a cure end point.

Let me just mention that it is not completely obvious thought that that gets you a very low sample size. In HIV, when we are using viral load, we are looking for differences that are easy to quantitate that are very large in magnitude, and that allows us to get a much smaller sample size.

I think that Dr. McCracken was mentioning this morning that with standard therapies that we might be able to achieve 99 percent bacterial

eradication, and we should be able to with this marker be able to clearly see differences.

Well, if we wanted to discern the difference between 99 and 98, that would take about 6,000 patients. So that is no easy answer here. If on the other hand, we were trying to discern the difference between 99 percent bacterial eradication versus 93 percent, then we are down to around 250 people.

So my question here isn't so much whether bacterial eradication is an important thing, but how much can we fall away from 99 before we care, and that is a critical question to find out whether use of that marker truly will give you a much smaller sample size.

CHAIRMAN RELLER: Thank you, Tom. Dr. Goldberger, we have tried to have forthright comments on all of the questions that you posed, and a rigorous discussion, which I think has taken place.

And I would like to in closing thank Dr. Shlaes and the Pharmaceutical Research Manufacturers Association, his colleagues, industry, and Dr. Tally, and Dr. Talbot, and other members representing the IDSA, as well as of course all of the members of the committee, including those who were added to the committee for discussions from the pediatric

subcommittee and other advisory committees with expertise relevant to the discussions today.

So thanks to all, and we will reconvene for Phase II tomorrow morning at eight o'clock with discussion of the development of drugs for emerging resistance.

(Whereupon, at 5:47 p.m., the meeting was adjourned, to resume at 8:00 a.m., on February 20th, 2002.)