UNITED STATES OF AMERICA

+ + + + +

DEPARTMENT OF HEALTH AND HUMAN SERVICES
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH

+ + + + +

PULMONARY ALLERGY DRUGS ADVISORY COMMITTEE MEETING

+ + + + +

Tuesday,

September 6, 2002

+ + + + +

The meeting was called to order at 7:32 a.m., in the Main Ballroom of the Gaithersburg Holiday Inn, Two Montgomery Village Avenue, Gaithersburg, Maryland, by Dr. Mark Dykewicz, Committee Chairman, presiding.

PRESENT:

DR. MARK S. DYKEWICZ, Chairman
DR. ANDREA J. APTER, Member
DR. T. PRESCOTT ATKINSON Member
DR. VERNON CHINCHILLI, Member
DR. JESSE JOAD, Member
DR. PETER E. MORRIS, Member
DR. POLLY E. PARSONS, Member
DR. ERIK R. SWENSON, Member
DR. JAMES K. STOLLER, Voting Consultant
DR. DONALD PATRICK, Voting Consultant
WILLIAM J. KENNEDY, Industry Representative
MS. KIMBERLY LITTLETON TOPPER, Executive Secretary

SPONSOR REPRESENTATIVES AND CONSULTANTS:

DR. BURKHARD BLANK
DR. BERND DISSE
DR. JAMES DONOHUE
DR. PAUL JONES
DR. STEVEN KESTEN
DR. DONALD MAHLER
DR. SHAHENDRA MENJOGE
DR. ERIC PRYSTOWSKY
DR. THEODORE WITEK

FDA REPRESENTATIVES:

DR. BADRUL CHOWDHURY, FDA Representative
DR. LISA A. KAMMERMAN, FDA Representative
DR. ROBERT J. MEYER, FDA Representative
DR. EUGENE SULLIVAN, FDA Representative

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

PAGE

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

(7:32 a.m.)

CHAIRMAN DYKEWICZ: Good morning. Let's convene our meeting of the Pulmonary-Allergy Drugs Advisory Committee. I am Mark Dykewicz, Chair, and I am a Professor of Internal Medicine, and Director of the Training Program of Allergy and Immunology at St. Louis University School of Medicine.

And let's begin the meeting with introductions by each of us, starting with Dr. Kennedy. For each of you on the committee, when you do want to speak, push down on the microphone button, and then when you are done speaking, push it off so that you are not going to broadcast your comments all over.

DR. KENNEDY: Good morning. I am bill Kennedy, and I am the Industry Representative, and consultant to the pharmaseutical industry, and I was formerly vice president of regulatory affairs for it.

DR. SCHATZ: I am Michael Schatz, and I am Chief of the Department of Allergy at Kaiser-Permanente Medical Center in San Diego, and a clinical

professor at UCSD, and I am a guest speaker today.

DR. PARSONS: I am Polly Parsons, and I am a Professor of Medicine at the University of Vermont, and Chief of Pulmonary Critical Care at Fletcher Allen Health Care, and Chief of Critical Care Services there.

MR. MORRIS: I am Pete Morris, and I am an Assistant Professor in the Division of Pulmonary and Critical Care Medicine at Wake Forest, North Carolina.

DR. JOAD: I am Jesse Joad, and I am a Professor of Pediatric Pulmonary and Allergy at the University of California at Davis.

DR. STOLLER: I am Jamie Stoller, and I am a Professor of Medicine with the Cleveland Clinic, and Vice Chairman of the Medicine and Associate Chief of Staff.

DR. SWENSON: I am Erik Swenson, and I am a Professor of Medicine at the University of Washington in Pulmonary and Critical Care Medicine.

DR. APTER: I am Andrea Apter, Associate Professor, Allergy and Immunology, Division of Pulmonary Allergy and Critical Care Medicine,

University of Pennsylvania.

DR. CHINCHILLI: I am Vern Chinchilli, and I am a Professor of Biostatistics at the Penn State Hersey Medical Center.

MS. SCHELL: I am Karen Schell, and I am a respiratory therapist in rural Kansas, and I manage a respiratory care department.

DR. KAMMERMAN: I am Lisa Kammerman, and I am a biometrics team leader in the Center for Drugs.

DR. CHOWDHURY: I am Badrul Chowdhury, Acting Director, Division of Pulmonary and Allergy Drug Products, FDA.

DR. SULLIVAN: My name is Gene Sullivan, and I am a Medical Officer in the Division of Pulmonary and Allergy Drug Products.

DR. MEYER: I am Bob Meyer, and I am the Director of the Drug Evaluation II in CDER.

CHAIRMAN DYKEWICZ: Thank you. We will now receive the conflict of interest statements by Ms. Kimberly Topper.

MS. TOPPER: The following announcement addresses a conflict of interest with regard to this

meeting, and is made a part of the record to preclude even the appearance of such at this meeting.

Based on the submitted agenda for the meeting, and all financial interests reported by committee participants, it has been determined that all interests in firms regulated by the Center for Drug Evaluation and Research present no potential for an appearance of a conflict of interest at this meeting with the following exception.

Dr. Andrea Apter has been granted waivers under 18 U.S.C. 208(b)(3), and 505(n)(4) of the FDA Modernization Act for her spouse's interest in Pfizer, a co-marketer of Spiriva, and a competitor to Spiriva.

The stock value is between $50,000 and a hundred-thousand dollars. These waivers permit Dr. Apter to participate in the committee's deliberations and votes concerning Spiriva. A copy of this waiver statement may be obtained by submitting a written request to the Freedom of Information Office, Room 12A30, of the Parklawn Building.

With respect to invited guests, Dr. Michael Schatz, we would like to report that he is a

researcher for Aventis, Giaxo, and Astra, on inhaled corticosteroids. He also receives speaker fees from Astra for his talks concerning asthma and pregnancy.

In addition, we would like to disclose that Dr. William J. Kennedy is the non-voting guest industry representative. He is not a government employee, and hence we do not screen him for conflict of interests and we can make no comments on his actual or perceived conflicts of interests.

In the event that the discussions involve any other products or firms not already on the agenda for which an FDA participate has a financial interest, these participants are aware of the need to exclude themselves from such involvement, and their exclusion will be noted for the record.

With respect to all other participants, we ask in the interest of fairness that they address any current or previous financial involvement with any firm whose products they wish to comment upon. Thank you.

CHAIRMAN DYKEWICZ: Thank you. Dr. Patrick, would you like to introduce yourself, please.

DR. PATRICK: I am Donald Patrick, and I am a Professor of Health Services and an Outcome Research Specialist from the University of Washington in Seattle.

CHAIRMAN DYKEWICZ: Thank you. We will now begin with introductory comments by the FDA, starting with Dr. Robert Meyer.

DR. MEYER: Thank you. I want to leave the more formal introductory comments to Dr. Chowdhury, but I did want to make special note of the choice of having the meeting today. At sundown tonight, an important holiday for many of us in the FDA side, and on the committee, and I am sure in the audience, as well as in the company, begins.

And it was not by first choice by any means that we had the meeting today, but because of not wanting to hold the meeting in conjunction with September 11th, where travel would be necessary over that anniversary, and because of wanting to constitute the most full and expert committee possible, this was the only feasible day.

So I certainly offer apologies for the

choice of the day, but again we felt that we did not have a choice in having it today, and due in difference to the holiday beginning this evening, we did start the meeting early, which explains why we are all here at 7:30, and we will try to wrap up in a timely fashion to get folks home.

And now I will turn it over to Dr. Chowdhury for more formal introductory comments.

DR. CHOWDHURY: Good morning, Honorable Chairman, and Members of the Pulmonary and Allergy Drug Advisory Committee, I welcome you to this meeting, and thank you for your participation this morning.

This meeting is to discuss the new drug application of tiotripium bromide inhalation powder inhalation powder from Boehringer Ingelheim Pharmaseuticals. The materials to be discussed in this meeting, and opinions that we are seeking from you, are solely related to clinical issues of tiotripium.

Please bear in mind that the regulatory decision-making process to determine approvability of

the drug product, the agency takes into consideration various factors, in addition to clinical issues, such as chemistry, manufacturing, and controls for drug product, and pre-clinical considerations.

These are not being discussed in this meeting. This meeting is solely to discuss the clinical issues of tiotropium. Boehringer Ingelheim is seeking an approval for tiotropium bromide inhalation powder for the treatment of bronchospasm, and dyspnea, associated with COPD.

While all clinical issues related to tiotropium are open for discussion, we are asking for a deterred deliberation on the dyspenea claim because the specific indication of dyspenea is unique amongst all drugs that are currently approved in the United States for COPD.

As you can see in the agenda, Boehringer Ingelheim will first present an overview of the clinical data, following by the Agency's presentation. As you hear through the presentation, I would request for you to keep in mind the questions that are in the FDA briefing book, and also attached to the agenda,

since you will discuss and deliberate on these questions later in the day.

We look forward to an interesting meeting and again thank you for your time, effort, and commitment in this important public service. I turn it back to you, Mr. Chairman.

CHAIRMAN DYKEWICZ: Thank you, Dr. Chowdhury. We will now proceed with the presentation from the product sponsor, Boehringer Ingelheim, beginning with Dr. Burkhard Blank.

DR. BLANK: Good morning, Mr. Chairman, and Committee Members, and Members of the FDA, Ladies and Gentlemen, my name is Burkhard Blank, and on behalf of Boehringer Ingelheim, I want to thank you for the opportunity to discuss with you today Spiriva NDA in COPD.

COPD is a growing health problem worldwide. In the United States, it is the fourth leading cause of death, and further increases in its prevalence and mortality of being predicted. The disease is characterized by an increasing limitation of air flow, partly the result of bronchospasm present

in many patients.

Typically after many years of smoking, patients first develop chronic cough and increased mucous production. It is not, however, until they develop shortness of breath or dyspnea that most patients seek medical care.

This dyspnea is chronic and it gets worse, and eventually it limits the abilities of the patients to perform every day activities, and in unfortunate patients it may be present at rest. So far the only intervention that has been shown to change the course of the disease is smoking cessation.

Therapeutically, bronchodilators, primarily inhaled anticolonegics, and beta aganists (phonetic) are widely used for the relief of bronchospasm. Spiriva is an inhaled, long-acting, once-daily anticholinergic, and we have developed it for the treatment of patients with COPD.

The NDA contains data of over 4,000 subjects. In Phrase III, we enrolled more than 2,600 patients, roughly half of them receiving Spiriva. We performed six long term trials, which were conducted

as three replicate pairs.

Three, one year trials, comparing Spiriva against placebo done in the United States, two ipratropium one-year controlled trials in Belgium and The Netherlands, and somewhat later in the Phrase III, two, six-month trials with both a placebo and a salmeterol control group.

The objectives of Phase III were first to confirm that Spiriva, when inhaled once daily, provides to the patients reliable 24-hours of bronchodilation. For that purpose, and in line with the outcome of the phase and end of Phase II meeting with the agency, we selected the trough FEV1 response; i.e., the extent of bronchodilation present at the end of a 24-hour dosing interval as primary end point in all six trials.

The four, one year trials, as I indicated earlier, were performed first, and they included the measurement of dyspnea as a secondary end point in all treatment arms.

We found the results for Spiriva so encouraging that we decided to confirm these findings

in two pivotal trials. After consulting with the agency about our intentions, we amended the study protocols of the two six month trials to include as a co-primary end point the assessment of improvement of dyspnea when comparing the Spiriva group with the placebo group.

This amendment was made at a time when both trials were clinically complete. However, when the study blind how remained intact. Finally, the six long term trials allowed us to evaluate the safety of Spiriva in a broad patient population in COPD receiving long term treatment.

We are here today because the agency seeks your advice on a number of questions which all fall in these areas. First, does Spiriva really show 24 hours of bronchodilation, and secondly, are the observed improvements in dyspnea supported by measurements of a validated instrument and is the observed improvement of dyspnea meaningful.

Specifically, the agency asks the question was the responder definition that we choose clinically meaningful and is the difference in response rates

between tiotropium and placebo important.

Finally, as in regards to the safety, was the safety of Spiriva adequately assessed, and is the safety profile appropriate for the intended use. The agency makes specific reference to subtle indications that the use of Spiriva may be associated with cardiac events, especially in the category of heart rate and rhythm disorders.

In our presentation, we will present to you all data that we feel are helpful for answering these questions. First, we will show you the trough FEV1 primary end point across all six studies, and go through the consistency of the findings, and then show you the secondary spirometric and the secondary nonspirometric findings.

Following that, we will explain the BDI/TDI instrument, and argue to you what it is an appropriate tool to measure dyspnea. We will then show you the data on dyspenea from both the pivotal trials and from the four, one year trials. Finally, we will share with you the safety profile as it was observed in the Phase III, and the total of 2, 600

patients, half of them on Spiriva.

This safety profile, not unexpectedly, reflects the pharmacology as an anticholinergic compound in a very similar way to what we have seen from atrovent, which has been widely used over many years. Most importantly, we see no association with Spiriva and life-threatening events.

In reviewing today with you the clinical results of Spiriva, we hope that you will find the data convincing and in support of the proposed indication statement outlined on this slide. Dyspnea is the most disabling symptom for patients with COPD, and we will present to you data from two pivotal trials confirming an improvement of dyspnea by Spiriva.

These data, together with consistent supportive data from the four, one year trials, provide the basis to include the improvement of dyspnea in the products label, and we propose the indications and usage sections as the most appropriate place for this. Following me, my colleague, Dr. Bernd Disse, will show you bronchodilation data. Then Dr.

Jones will explain the instrument.

My colleague, Dr. Theodore Witek, will show you the data on dyspnea, and the safety profile will be provided by Dr. Kesten. Dr. Jim Donohue will share with you his perceptions as a treating physician on where he sees the place for Spiriva, and what does Spiriva offer to the patients with COPD, and I will come back with concluding remarks.

Since our presentation is built on each other, we believe that it is most appropriate for the objective of the meeting if we can answer questions at the end of our presentation. We are honored today to have not only Dr. Donohue and Dr. Jones with us today in the audience, but also Dr. Mahler, who developed the BDI/TDI instrument; and Dr. Prystowsky, who gave us his independent assessment of the cardiac safety of Spiriva.

Unfortunately, for reasons which Dr. Mahler addressed in his introduction, Dr. Prystowsky has to leave after the lunch break, and we would ask you that if you have questions that you want to direct directly to Dr. Prystowsky, please do so before the

lunch break. I would now like to hand this over to Dr. Disse.

DR. DISSE: Thank you, Dr. Blank. Good morning, ladies and gentlemen. I am Bernd Disse from Boehringer Ingelheim, and it will be my pleasure to introduce basic and bronchodilator efficacy results to tiotropium, and here is the overview of my presentation, and I will mainly focus on the Phase III spirometry results.

Basic cholinergic tone, as well as a major proportion of bronchospasm in COPD is mediated by isocolon (phonetic) and mass kirenreceptus (phonetic), or cholinergic receptus as they are often called in clinical medicine. And the standard bronchodilator used in obstructive lung diseases is ipratropium bromide, used 3 to 4 times a day, and the obvious room for improvement is duration of action.

Now, the new anti-mascorinic (phonetic) tiotropium is firstly more potent at about an affinity constant of 10 picomolar, which is very potent, but the most important quality of tiotropium is its long duration of action, and this is most likely brought

about by slow, very slow, disassociation from M3 receptors, and M3 is the receptor subtype responsible for smooth muscle constriction.

Tiotropium was first investigated in single dose studies in COPD, covering a dose range from 10 to 160 micrograms, and these studies established the pharmacodynamic duration of action to exceed 24 hours. A multiple dose study of four weeks treatment duration covered a range from 4.5 to 36 micrograms end placebo, and allowed us to select the dose for Phase III.

And this selection was based on the fact that the 18 microgram dose was approaching the pharmacodynamic plateau for FEV1, trough, average effects; and on the other hand, that the net dose, the 36 microgram dose, already had a slight tendency for increase in dry mouth, which is the most sensitive systemic side effect of anticholinergic treatment.

Tiotropium is a typical N-quaternary anticholinergic, and it shares all the positive properties of that compound class. For instance, it does not pass the blood-brain barrier. Now, from the

nominal dose of 18 micrograms, and up to an 8 to 10 microgram dose is delivered through the mouth piece, and the fine particle fraction of about 20 percent can be deposited in the lungs and eventually absorbed.

The coarse participles, the major proportion, deposits in the oropharynx, and is swallowed, and the remaining portion is cleared. As for absorption from the oral part, there is very low absorption, and this contributes minimally to the overall systemic load.

Now, to balance this, 3.6 micrograms reach the lungs, and distributes in the lungs, and gives rise to high tissue concentrations. Then absorbed systemically, it is 3.6 micrograms in the system which is diluted throughout the body, and gives rise to low tissue concentrations.

Here the overall as to pharmacokinetics. I mentioned already that the bioavailability by inhalation is about 20 percent, which gives rise to very low plasma concentrations. The molecule is metabolized by about 25 percent by P450 enzymes in the liver, and to some extent nonenzymatically, but the

major route of excretion is unchanged compound, 75 percent, via the kidneys.

The renal clearance is high, and exceeds even the creatinine clearance, and as may be expected for most renal excreted drugs in patients with moderate to severe degrees of renal impairment, we have seen increases in the plasma levels, but they never exceed more than doubling of plasma concentrations, and the consequences of this in older aged patients with renal impairment, we have seen some increase in the side-effect of dry mouth.

The half-life of this drug is about 5 to 6 days, and this is a pharmacokinetic half-life, leading to steady state in about 2 to 4 weeks, but the pharmacodynamic half-life, which depends on lung concentrations, is reached much faster, in about one week.

I will now focus on the long-term Phase III studies. Our proposed indication is tiotropium indicated for long-term, once daily, maintenance treatment of bronchospasm, and dyspenea, in COPD, and my focus will be long-term, once-daily, bronchospasm

and to provide the substantial evidence needed for this.

We conducted six major studies organized in sets of three repetitive studies, and all of these were randomized, double-blind, double dummy, if applicable, of course, parallel group comparisons, and the treatment, the active treatment was 18 micrograms of tiotropium by dry powder inhaler.

In the first set of studies, we had one year treatment duration and comparator placebo. In the second year, in the second set of studies, again one year treatment duration, and comparison to ipratropium by MDI, four times a day. And in the third set of studies, we compared the placebo and salmeterol two times a day.

Here is our patient selection. The selection was based on a clinical diagnosis of COPD, and we excluded patients with asthma, allergic rhinitis, or atropy, and everyone was required to be 65 percent of predicted normal, or less than 70 percent of the force vital capacity of these patients.

The age was higher than 40 years, and they

had to have a smoking load of more than 10 pack years. The exclusion criteria were defined as follows. We excluded unstable patients not able to participate in a long term study as judged by the investigator. Patients with a recent respiratory tract infection were included.

Further, patients with a recent history of myocardial infarction, cardiac arrhythmia, requiring treatment, or hospitalization for heart failure, were excluded; and anticholinergic class contraindications, narrow-angle glaucoma, bladder neck obstruction, or prostatic hypertrophy, were excluded.

These inclusions and exclusion criteria allowed us to recruit a COPD population with a broad range of significant and stable co-morbidities typical for the age group, and this is further outlined in the next slide.

Concomitant diagnosis of cardiovascular diseases was in about 50 percent of these patients, and among these the most prominent were hypertension, with about 20 percent, but also cases of coronary artery disease, cardiac arrhythmias, myocardial

infarctions in patients' histories, ranging from 2 to 12 percent.

Neurologic and psychiatric diagnosis were quite common, too, and most prominent, the class of depression, with about 21 percent. Patients with prostatic hypertrophy and micturation disorders ranged from 2 to 11 percent in this patient population, and the ranges do not really reflect differences in the populations, and I think that is more differences in the diagnostic habits in the countries involved.

Now, here is the demographics of our study population, and I should first mention that it was balanced between the treatment groups within the studies, and comparable in the sets of repetitive studies.

The included patients were mainly male, mostly Caucasian, and some African-Americans included, and there is a separate program ran in Japan to include the Asian population, not part of this in any way.

The mean FEV1 and percent predicted normally characterizes the severity and in the one

year placebo control studies, it was moderate to severe, slightly more moderate in the ipratropium control study, and in the six month studies in between. So the range of patients covered is from very severe, at about .3 liters, which is really very severe, to mild patients, at about 2.5 liters of FEV1.

Here is the overview of our primary end points. We have chosen Trough FEV1 as the primary end point in the one year studies, measured at 13 weeks, and at 24 weeks in the six month studies. In addition, we measured dyspnea as a co-primary end point in the six month studies at the end of the study, and this will be covered by my colleague, Dr. Witek.

Now, trough FEV1, and that is the mean of the pulmonary function breathings at 1 hour and 5 minutes before the next drug administration, and this reflects the maintained drug activity at the end of the dosing interval, and so this is why we made this choice.

As secondary end points, we measured the time course of FEV1, in clinic measured forced vital

capacity, to support pulmonary function measurements, and home-measured peak flows. The shuttle walking test was included in the six month studies. This test has not been shown to separate drug treatment effects in the literature, and we also have not been able to show the separate effects of tiotropium, spirometry, or placebo, with this test.

However, symptoms and exacerbations of COPD in patient recorded outcomes may lend support of overall and consistent patient benefit. Now, here is the key spirometry results. In the next few diagrams, I will always use the same scheme. The FEV1 is on the y-axis, and please note that it has depicted changes higher than one liter, and the x-axis is at the time after administration, and it is not entirely to scale.

Now, this is the first dose effect of the placebo adjusted for a common base line, and you do see an appreciable bronchodilator response. After eight days of treatment, we reached a steady state, and now patients wake up at an elevated daily base line, versus a study base line, and they present in clinic already with a better lung function value.

So this represents sustained activity for 24 hours, and 90 days of treatment brings us to our primary end point, and the trough FEV1 value versus placebo was significantly elevated, and we high significancies throughout the day, and so peak, average and trough, and that all time points were significant at a p-value of 0.0001.

Now, here is the value at the end of the study, and again lung function measured for three hours, and you do see that the lung function profile is unchanged over time, and that means that we have maintained efficacy over the one year treatment period, and there is no indication of tolerance whatsoever.

To be mentioned, we conducted two studies of this kind, and Number 115 is really absolutely comparable, and I don't need to present these data as they have been outlined in the briefing documentation. Now, as to the comparisons in ipratropium, the active comparator, the day one and day eight response, and again you see on the first day that ipratropium, the green, and tiotropium, the yellow, is in the beginning

comparable, but then tiotropium is more long acting.

On day eight, patients wake up with their improved lung function (inaudible) base line, and so the trough value is elevated, and our next dose again shows an increase in FEV1, which is substantial, and the end of study shows that you only followed for three hours, and the end of study shows the lung function profile is unchanged for both drugs. So this has maintained efficacy over the one year period.

Here are the results of the comparison study to placebo and salmeterol and the interesting feature here is that we measured lung function over the date, and that means the 12 hours. You do see the profile on Thursday, and a substantial increase over the 12 hour period, and unrepeated dosing reaching a steady state, and we have the elevated draft effect.

And again an increase over the day, and the efficacy is sustained over the 24 hour period, and maintained over the one year treatment period, or over the half-year treatment period. I'm sorry. And here is a comparison to salmeterol, and the profile on day

one, on day 15, and on day 169.

Here is the replicant sister study, and only measured lung function for three hours, and so this was done somewhat simplier, but the lung function profile was essentially the same, and so you do see the first dose effect, and the trough effect elevated over baseline, and maintained over the half-year period, and again in comparison to salmeterol, day 1, day 15, and day 169.

And I would like to summarize the magnitude of spirometric improvements. Tiotropium elicited an appreciable magnitude of response, and this table is compiling the mean response reached in comparison to placebo at the end of the treatment intervals. The mean response reached 190 to 250 milliliters at peak, and this is about 17 to 24 percent improvement from baseline compared to placebo.

And even in trough, an improvement of about 13 percent of baseline versus placebo is reached, and this is a lot considering that the trough reflects the minimum effectiveness reached and sustained over the entire 24 hour period, and the

interesting feature is that tiotropium reaches a trough to peak rate ratio of 53 to 72 percent, and this sets the standard for 24 hour effectiveness.

The values for salmeterol are explained on the right-hand side of the stable, and they are numerically lower at peak than trough in both studies. We conducted subgroup analysis, and it was analyzed in the combined replicate studies for an influence of age, gender, smoking status, severity of disease, previous atrovent use, and most important, concomitant medications, and it can be stated that tiotropium was similarly effective in all subgroups analyzed here.

Now I would like to report the supportive information obtained from the secondary end points. The vital capacity was assessed in all the studies, and as an example, I show the peak and trough force vital capacity response of a one year treatment period for the combined studies, comparing the placebo for one year.

And, of course, the statistical evaluation was based on the individual studies and it can be stated that the results were significantly with P-

values less than 0.0001, and at all time points after reaching steady state.

As you can see, tiotropium treatment provides maintained improvement of the fourth vital capacity over the year, and that the trough reaches values of 290 milliliters, and at peak it reaches values of 440 milliliters, always in comparison to placebo. This effect can be interpreted as improvement of air flow limitation and reduction of hyperinflation, leading to reduced breathing, and should be associated with an appreciable symptomatic improvement in these patients.

Also, home measured peak flow rates were assessed in weekly intervals and weekly means, and as can be seen here, the morning peak flow was increased by 10 to 30 liters, and the evening peak flow was increased by 15 to 40 liters, and these results were significant at most time point, again evaluated in the two individual studies.

As a secondary end point, we assessed exacerbations of COPD, and they were defined either as an exacerbation diagnosed by the physician, or as a

complex COPD related symptoms, cough, wheeze, dyspnea, sputum production, two of these, lasting at least three days, and reported as an adverse event.

Now, when analyzing our four, one year studies for this secondary end point, we saw encouraging trends and occasionally nominal P-values of less than 0.05 in the individual studies. For this reason, we conducted retrospective exploratory analyses in the combined replicate twin studies, and pre-specified it as combined analysis in the six month studies, which were conducted somewhat later.

And we would like to share these interesting scientific results with you. A common way to analyze exacerbations is by Kaplan-Meier analysis, and the probability of no exacerbation is depicted here, versus the days on treatment in the placebo-controlled one year study, and you do see an appreciable improvement of tiotropium of the placebo, and the appropriate way of statistical analysis is the time to first exacerbation, and a nominal p-value of less than 0.05 could be assigned.

And in a similar graph here, the

comparison to the active comparator, ipratropium, and again an appreciable advantage of tiotropium in the probability of non exacerbation, versus ipratropium and the time to first exacerbation has a nominal p-value of less than 0.05, and the same for the six month comparison of tiotropium versus placebo again, and the time to first exacerbation has a nominal p-value of less than 0.05.

With this, I would like to outline the results we obtained with the COPD specific health status assessment with the St. George's Respiratory Questionnaire and the instrument assesses patient's health status in symptoms, activity, and impacts, and gives the total score, and a decreasing score indicates improvement, and the score change of more than four points is suggested to be clinically meaningful.

And here as an example for all, the results from our one year study, and in Study 114, you do see an improvement of the score, which increases over time, and becomes significant after a year, and it is approaching at the end of the year a threshold

that is suggested to be clinically meaningful.

And in the second study, it is a similar picture, and so improvement over time, and at the end of the study the clinical meaningfulness is reached. The findings with the St. George's Questionnaire support the impression of overall benefit achieved with tiotropium.

In summary, tiotropium, once daily, provides clinically meaningful improvement of spirometric measures sustained for 24 hours, and the improvements were maintained over one year with no evidence of tachyphylaxis.

The analysis of secondary end points, as well as exploratory analyses show improvements of related lung function measures, such as the force, vital capacity, and peak flows. Exacerbations of COPD appear to be reduced, and improvements in health status as measured by the St. George's Questionnaire meet or approach the threshold of clinically meaningful change with prolonged treatment.

Thank you for the attention, and I would like to hand over the podium to Professor Jones, who

will introduce the assessment of dyspnea.

DR. JONES: Thank you, Dr. Disse. I am Paul Jones, and I am a pulmonologist, but I have developed health status instruments in the past, and I have also worked in the field of dyspnea measurement, although I was not involved in either the development or the validation of the BDI and the TDI that we are discussing here.

This presentation will be in three parts; the measurement of breathlessness, the validation of the instruments that we are discussing, and the identification of a clinically significant threshold.

Dyspnea is a principal symptom of COPD, and there are multiple causes for it. Expiratory airflow limitation, increasing static lung volume, and the dynamic hyperinflation that occurs at exercise onset. Recent studies have shown that these two components are more important predictors of dyspnea than expiratory airflow limitation, but they are complex measurements.

And in fact there is no simple physiological measure, whether complex or simple, that

can be measured as a surrogate for dyspnea. So dyspnea should be measured directly. Dyspnea is a sensation, and for that reason it should be related to a known level of stimulus.

In the laboratory, that is easy. We can measure breathlessness, and relate it to a known level of work rate, minute ventilation, or oxygen consumption. But the requirements of laboratory exercise tests are far too complex to be included in large multicenter Phase III clinical trials.

For that reason, breathlessness is related to reference points in daily life. For example, being breathless when getting washed or dressed, or walking up hills, and in fact these reference points were used as the basis of the MRC and the American Thoracic Society grading systems for dyspnea.

You will appreciate that what we have here is a ranking of activities based on metabolic demand. It is important to understand this, because you then realize that the breathlessness measurements are grounded in physiology. Thus, in contrast to functional disability, or health status measurements,

there is more grounded in patient's perceptions.

There is a multifactorial relationship between dyspenea and activity. There are activities that cause dyspenea, and activities that become more difficult because of dyspnea, and activities prevented by dyspnea. And it was an understanding of this that led the developers of the BDI and TDI to develop this particular construction.

It has three components; functional impairment, magnitude of task, and magnitude of effort. There is also a focal or total score. The two questionnaires are related, but have different properties. The BDI is cross-sectional, used for distinguishing levels of dyspnea between patients so that it is discriminative.

The TDI is grounded on the BDI, but it is longitudinal, used within patients to evaluate changes. We now look at the psycometric properties of the BDI, and we find that it has good internal consistency, good inter-rater reliability, and test-retest reliability. The panel should understand that a questionnaire with poor psycometric properties will

tend to underestimate the true effect of a change that is reparent.

Perhaps more importantly the question is do these instruments measure dyspnea, Unfortunately, it is not possible to address this question in one step. We have to set up a number of hypotheses, and then test with the questionnaires related to physiological impairment, other measures of dyspnea, and health status.

The next few slides summarize the evidence for this. First, we find that there is in fact to my view a relatively surprisingly good correlation between FEV1 and the BDI. That is, the expected level of correlation with exercise performance and with other measures, with established measures of dyspnea; the ATS questionnaire, the Oxygen Cost Diagram, and the more recent Shortness of Breath Questionnaire developed at UCSD.

The BDI correlates with disease specific health status, measured using the CRQ and the SGRQ, and generic health status measured with the SF-36 and the QWB. If we turn now to the TDI, we find that it

has good inter-rate reliability, and in terms of its responsiveness to change, we find that following pulmonary rehabilitation the TDI score correlates with change in the CRQ dyspnea score, and following recovery from a COPD exacerbation, again there is a very good correlation with change in the CRQ dyspnea score, and really quite a surprisingly good correlation with change in FEV1.

If we now turn to the issue of clinical significance. There are a number of different ways in which this can be assessed, but historically the first and perhaps the most widely used is the humanistic approach, and perhaps best described in the seminal paper from Dr. Guyatt's group in 1989, in which he defined the minimum clinically important difference was that difference in score which patients see as beneficial, and would mandate in the absence of troublesome side effects and excessive costs a change in the patient's management.

I should also point out that this approach was used for the development of the threshold for the Juniper Asthma Quality of Life Questionnaire, which I

believe is now accepted by the agency. If we turn to the TDI, and look at one of the components, the magnitude of task, we see that there are three grades of deterioration, and three for improvement.

Let us concentrate on the smallest degree of improvement and look at an example. Here we have a patient who was dyspneic when walking on the level, or perhaps even when washing, and now has become dyspneic only when walking up a gradual hill or carrying a light load on the level.

To clarify this and set this into a broader setting, let us return to the ATS Dyspnea Grade that I have simplified for the purposes of presentation. I should just point out that COPD is a chronic and incurable disease, and it is not possible to convert a patient who is severely disabled, such as they are breathless when they are getting washed and dressed, to someone who can undertake strenuous exercise.

But worthwhile improvements can be achieved, and to illustrate that, what a change of one unit in the TDI can mean, we may have a patient who is

still breathless on walking up hills, but is now no longer breathless when walking at a normal pace on the level.

Another example would be a patient who is breathless when they are walking slowly on the level, but they are now no longer breathless when washing or dressing. These changes I would contend are not trivial, and they more than exceed the criteria for minimally important improvement as defined by Dr. Guyatt.

So in summary the BDI and TDI have reliable measurement properties. Their scores are valid measures of dyspnea, and we can attach clinical significance to them. I would like to thank you for your attention, and pass over to Dr. Witek, who will present the results from the tiotropium studies.

DR. WITEK: Thank you, Dr. Jones, and good morning, ladies and gentlemen. My name is Ted Witek from Boehringer Ingelheim, and as my colleagues have mentioned, dyspnea is a unique claim in our proposed indication, and I would like to spend the next 15 minutes describing the data and the application of the

instrument in the program to help us in our deliberations today.

I will briefly review the assessment of dyspnea in clinical trials, and how we applied the BDI and TDI in our program, and then I will review the response of the TDI and the related measures to tiotropium. Now, in the assessment of dyspnea in clinical trials there are several things that we needed to consider.

Particularly with tiotropium, where we have a long term maintenance treatment, it is important that we evaluate and find an instrument that can assess the effects of dyspnea over time, and in fact, knowing that the TDI have been previously used in a two year perspective study, where we saw the drop in TDI of about .7 units over two years, indicating the natural decrement in dyspnea, this was one of the elements in our selecting the BDI and TDI.

Also, the instruments need to be practical for a multi-center, and in our case, in multi-national programs, where the instructions for the use of the instrument are in the uniform training and

investigator meetings. Secondly, dyspnea assessments need to be in the context of a clinic visit where there are many measurements.

However, it is important that we have supported measurements in our assessments to both determine and help determine the validity of the instrument in practice, as well as evaluating the consistency of related measures.

Now, briefly, just some key protocol elements to keep in mind. The TDI evaluations were performed at clinic visits. For example, in the six month studies, in days 57, 113, and day 169. As noted by Dr. Jones, the TDI assessments referenced the BDI scores, which were collected at baseline, and the TDI is completed after the SGRQ, and prior to the post-dose pulmonary functions.

Now, this further evaluates or illustrates the domains that Professor Jones had mentioned; the functional impairment, the magnitude of task, and magnitude of effort. If I just focus on the BDI for one moment, here we have scores that range from zero to plus four, and that gives us a focal score range of

zero to plus 12 units at baseline; zero being very severe dyspnea, and 12 being little or no dyspnea.

And if we put some real numbers to the BDI, this is the distribution of the BDI score baseline in our population in the one year study, and here you see the BDI focal score, and on average the patients that were enrolled in our clinical trial have a BDI focal score of six, indicating moderate dyspnea.

And a BDI focal score of six, for example, could be a patient who recorded a grade of two in each of the three domains; and if that was the case, a BDI focal score of six may be describing a patient who abandoned at least one activity due to shortness of breath, became short of breath with an average task, such as walking up a gradual hill, and become short of breath with moderate effort task performing, with occasional pauses, and requiring longer to complete than the average person.

Now I would like to turn to show you our data. I will describe for you the two studies where TDI was listed as the co-primary end point, and there our specification was responder analysis, and the four

studies where TDI was a secondary end point, and in those reports we had originally looked at the mean TDI response.

Now, in our discussions with the agency, we have discussed the advantage and disadvantage of the two ways to evaluate or to express the data in responder analysis or means, but what was agreed was that whatever we do select, it should be stated in a formal protocol amendment, which we did do and which was outlined by Dr. Blank, and that both illustrations should be provided; i.e., responder analysis, and the means.

So I will do this for you. I will show you these TDI improvements, list the supporting endpoints from our secondary measures, and would like to point out the consistency across the time of the trial, as well as cross-studies, which we do feel is a strength of our data.

Now, just a point on the responder analysis and mean response. We chose the responder analysis, and what this is, is the proportion of patients achieving a meaningful response, which we did

define a priori as a plus-one unit change in TDI focal score.

So this one unit change as described by Dr. Jones is inherent in the instrument, and of course this responder analysis will then reflect the individual patient changes from baseline. The analysis of means is also important, because that does reflect the overall population change, and you are able to illustrate the differences you see from the drug relative to placebo.

And here a positively and significant delta, versus a placebo, indicates an overall benefit. So I will show you all of the data from particularly the placebo control studies, where a drug effect could be evaluated.

These are the data from Study 130 described by my colleagues, which was the first of the two, six month studies. On the y-axis is the percent of patients responding; i.e., the percent of patients that reach the plus one unit change or greater, and the x-axis is the three study days.

Here you see the proportion of patients

responding to tiotropium relative to the placebo, and all three of the study days, particularly day 169, which we pre-specified as the primary end point analysis. As noted in these trials, salmeterol was included, and here you see the proportion of patients salmeterol relative to placebo.

And in this study, the proportion of patients relative to placebo in salmeterol was not significantly different. In Study 137, the sister study, here we see again the proportion of patients responding to the TDI relative to placebo, and again the consistent response across the three time points, and importantly also in the day 169 that was pre-specified.

The study again also included the salmeterol comparator, and in this study, salmeterol was significantly greater than placebo in the proportion of patients responding relative to the placebo as I mentioned, with no difference between the tiotropium and the salmeterol.

Now, showing you the mean TDI focal score for the population, that is what is illustrated in

this slide. On the y-axis is the mean TDI focal score, focal score units, and the x-axis is time, the same three time points that I mentioned to you.

And here you see the effects from the improvement with tiotropium relative to placebo across the three time points. Here at the end point, day 169, the effect size is 1.02 units in TDI focal scores, and so that is the mean difference between tiotropium and placebo. In the second study, again you see the improvement with tiotropium, the significance indicating the effect relative to the placebo mean, and the mean effect size in the second study was 1.2 TDI focal score units.

In this study, salmeterols were included as I mentioned, and here in the first study, you see the effects of salmeterol in the mean TDI relative to placebo; and in the second study, as was the case with the responder analysis, there was a significantly higher mean effect relative to placebo for salmeterol, with again no difference between salmeterol and placebo.

Now, I will review for you the one year

studies, the one group of studies relative to placebo, and the second group of studies relative to tiotropium that Dr. Disse showed you for lung function.

So in the top panel is study 114, and the bottom study is 115, and the y-axis is the proportion of patients responding, and here you see the higher proportion of patients responding in tiotropium relative to placebo, and in the second study, you see the same pattern, with the asterisks indicating a nominal p-value of p-less-than-0.05.

When we go to the tiotropium controlled studies, here we see in yellow the proportion of patients responding to tiotropium relative to ipratropium bromide, and you see that similar pattern in both study 122a and 122b. These asterisks are indicating a nominal p-value of p-less than .05, even versus the active control in ipratropium bromide.

Now, to complete this, I will show you the mean TDI focal score for the one year studies, and again the y-axis is the focal score, and x-axis is time. So in the first placebo controlled one year, we see the improvements with tiotropium relative to

placebo, and here is the p-value, p-less-than.05 on all test days.

And similar in study 115, the improvement with tiotropium relative to placebo, and once again the nominal p-value significant at the level of p less than 0.01. These are the data in the ipratropium control trials. The same axis, and here you see a rather atypical response, but the improvement with the tiotropium that wanes over time, and a parallel response to the ipratropium with the difference between the two drugs, and the mean and TTDI focal score are still evident.

And here are the nominal p-value are all but one test day of less than 0.05. And in study 122b, you see the improvement with the tiotropium that is maintained over the one year, and in the ipratropium group, the increase, and again these two are paralleling each other with the differences being p-less than 0.05 on all test dates.

So these are the temporal pattern of response under two, one year studies. Now I would like to turn to the secondary supportive end points

that one would expect to see improvement with a change in the TDI, and this is the shortness of breath score, and to highlight the scale for you here, this is a scale on a unit of zero through three.

And this is a simple assessment, diary assessment, where you have here the placebo response, and the improvement in tiotropium relative to placebo, and all nominal p-values relative to the placebo are significant, and in the sister study, study 137, you see this improvement with tiotropium, and on the last day the nominal p-value is lost in this study, the last time point.

Looking in the one year studies, and this is the same pattern following, and here is this zero through three score, and four point scale, and the improvement with tiotropium. This delta, with this effect size, has a nominal p-value of p-less-than-0.05 on all days, and that pattern is also illustrated in the second one year study.

If we turn to the physician's global evaluations that were described, here again this is a scale representing a range of zero through eight

scale, and here you see the improvement with tiotropium that is maintained and that is relative to placebo.

And a similar pattern of response in the 137 study, and again we see that drop at the end with these differences, tiotropium less than placebo having a nominal p-value of p-less-than-0.05 on three of those test dates.

And again the one year studies, and the improvement with tiotropium relative to placebo in both trials, and in both of these studies, nominal p-values were achieved. Now, the last secondary end point that I will review for you is the supplemental albuterol use in patients who were allowed supplemental albuterol use, and here this is expressed and we are looking at weekly averages of the daily use.

And in study 130, we see the drop early on in the study that is maintained over the course of the six months, and in the second study, we see that initial reduction in albuterol use, and that initial reduction was not maintained later on in that second

study.

And in the one year studies, however, we saw the reduction in the use of supplemental albuterol that was maintained throughout the study, and that was also the case in the second one year study, where albuterol reduction is illustrated.

And again all nominal p-values were p-less than 0.5 or greater. So Dr. Disse had reviewed for you the FEV1 trough, and I briefly reviewed for you the TDI data, both in terms of response and the TDI mean from placebo. In all of those cases where we have listed our primary end point, we have achieved the statistically significant level versus placebo.

And importantly in those secondary end points that I just reviewed for you, it was with rare exception that we did not achieve a nominal p-value, indicating the drug effects of tiotropium in these secondary measures. So in summary, we believe that we have selected and utilized a validated instrument, and not only in the literature that was reviewed for you briefly with Dr. Jones, but also in our own internal program, where we looked at these similar

correlations.

We gave pre-specified the primary end points, and the key statistical significance in the two independent studies. The proportion with a meaningful change that we selected as our primary analysis was supported by the responses and the dyspnea response was reflected in the related measures that I have just shown for you.

So given the importance of dyspnea as a COPD symptom, and given our demonstration of dyspnea relief, we believe that dyspnea should be included as an indication for tiotropium's use. So, I thank you for your attention, and I would like to turn the podium over to my colleague, Dr. Stephen Kesten, who will review for you our safety analysis.

DR. KESTEN: Good morning. My name is Stephen Kesten, and I am the medical director of the International Spiriva Program for Boehringer Ingelheim. My task today is to summarize an extensive safety program in a focused and concise presentation.

And in a manner that provides you with the critical information necessary to allow you to judge

the safety of tiotropium, and respond to the questions posed by the agency. The data will demonstrate a safety profile consistent with ipratropium bromide, an inhaled anti-cholinergic, used in the treatment of COPD in the United States for approximately 15 years, and approximately 25 years globally.

For background information, the anti-cholinergic effects appearing in the most recent version of the U.S. label for ipratropium bromide are listed in this slide. These include the more common events of dry mouth, and less common or infrequent events seen such as tachycardia and Superventricular tachycardia.

These events are those that you might expect to see with a drug such as tiotropium. Our early experience in healthy volunteers indicated that we could elicit anti-cholinergic effects with tiotropium when administered in high doses and over multiple days.

Single dose studies of up to 282 micrograms, however, failed to show any effect on ECGS, vital signs, pupillometry, or salivary

secretions. Multiple dose studies of 70 and 140 micrograms could show anti-cholinergic effects, such as decrease in salivary secretions and reports of dry mouth.

However, even at these doses, we cannot see any effects on vital signs, ECGS, and pupillometry. The COPD experience with tiotropium in a dry powder formulation is illustrated in this slide. There were 1,723 patients randominzed to receive tiotropium, and 414 received tiotropium in studies of up to six weeks in duration, and 1,308 received tiotropium in long term studies ranging from 6 to 12 months in duration.

The safety profile of tiotropium has been characterized through a variety of measures which are depicted in this slide, and are illustrated in your briefing document. Abnormalities that would be expected in patients with COPD were observed.

The majority of our safety information comes from the clinical adverse event reporting. However, with these other measures, I would like to highlight a few aspects. A vital sign evaluation

showed no effect of tiotropium on heart rate.

Lung function testing indicated that acute inhalation of tiotropium was well tolerated. The laboratory evaluations showed no influence of tiotropium, a finding consistent with what we would expect from inhaled and quatemary anti-cholinergics, and we performed several characterization studies evaluating different attributes of tiotropium which supported the overall safety of the compound.

The next few slides will summarize the electrocardiographic monitoring in the tiotropium program. Twelve lead ECGs and two-minute rhythm strips were performed as part of a four week, multi-dose, dose ranging study, with doses up to 36 micrograms daily.

ECGs were performed before, and at 1, 3, and 5 hours after dosing, these serial ECGs being conducted at baseline, and then at 1, 2, and 4 weeks. There were 134 patients who received tiotropium in this evaluation, yielding over 2,000 ECGs.

Twenty-four hour holds for monitoring was conducted before and after six weeks of treatment in

72 patients who received tiotropium, and in the long term studies, there were 12 lead ECGs performed at baseline, and then at 3, 6, and 9, and 12 months in the one year studies, and in the baseline and end of treatment in the six month studies.

ECG abnormalities in these long term studies were recorded as adverse events if the investigator deemed them to be clinically significant, or requiring treatment, or leading to the discontinuation of therapy.

Now, this slide summarizes the ECG findings on heart rate, rhythm, or conduction in the four week multi-dose, dose-ranging study. The ECG abnormalities in those categories are listed here. The numbers refer to the number of patients who had the associated ECG abnormality at any time while on treatment.

Now, I recognize that this is a busy slide, but what it illustrates is that there is no pattern here suggesting an association of these findings to any of the treatment groups. That is, this study demonstrated that there was no findings

that could be on heart rate, rhythm, or conduction associated with tiotropium, and as expected with an inhaled anti-cholinergic, there are no suggestions of prolongation of QT interval.

In addition to the prospective evaluations of ECGs in the one year studies, we retrospectively obtained the ECGs and sent them to a central laboratory for high resolution measurement of cardiac intervals.

There were no difference between groups in the proportion of patients who had an abnormal rhythm on any ECG. There was also no difference between on treatment groups with a mean change in heart rate from baseline, nor in the mean maximal change seen on any on treatment ECG.

The only finding that we observed was a 0.6 percent increase in the number of patients, or the proportion of patients who had at any time an ECG read as having tachycardia. Now, this constituted 12 patients in the tiotropium group, 10 of which only had this on a single occasion, and none of them had it on all occasions.

At the bottom of this slide, I have illustrated the heart rate ranges for the maximum heart rate that was seen on any of these tachycardic ECGs for these 12 and these 6 patients, which is down here at the bottom showing that most of this was in the range of 100 to 110 beats per minute.

All of the aforementioned ECG findings, including the Holter studies, have been reviewed by independent external cardiology consultants and the only suggestion of a finding has been the small imbalance in tachycardic ECGs. The remainder of this presentation will focus on the clinical adverse event experience.

There were eight short term studies of patients with COPD receiving the dry powder formulation in doses of 4.5 to 72 micrograms, with most patients receiving the intended dose of 18 micrograms daily.

Overall, there was no difference in the proportion of patients having an adverse event, and the only event seen that was associated with tiotropium was a dry mouth, and there was some

evidence of a dose response.

For completion, I have included a summary of the serious adverse events and deaths, and there was no difference in the serious adverse events, and no association with increasing dose.

Two of the deaths occurred many weeks after completion of the study, and the last death was in a placebo treated patient. The long term trial population consists of patients who had participated in two, or four, one year trials, and two six month trials.

Two of these one year trials were tiotropium controlled, and two were placebo controlled. The number of patients within a treatment arm, and the number of patients receiving tiotropium is illustrated in this slide.

And as described by Dr. Disse, these were mainly men, age around the mid-60s, and who had a mean FEV1 of about 40 percent predicted, and these patients had numerous co-morbidities.

The adverse event profile in the six month studies was similar to the four, one year studies. I

will therefore highlight the one year studies in the initial adverse event presentation of these long term trials.

Given the demographics of the population as described, and the duration of exposure, it is not surprising to see that approximately 90 percent of patients are observed to have at least one adverse event during the participation in the trial.

However, tiotropium was associated with a lower proportion of patients who had adverse events that were characterized as serious. Tiotropium also had a lower portion of patients who had adverse events leading to treatment discontinuation.

Fatal events were relatively few in these trials, with similar proportions among treatment groups. The next two slides characterize the most common adverse events observed with tiotropium in the one year trials.

The first two columns represent the description of the adverse event according to WHO adverse reaction terminology, with the first column being system/organ class, and the second column being

the preferred term listed alphabetically.

The four numeric columns represent the proportion of patients within a treatment group showing an adverse event. The most common adverse event associated and attributed to tiotropium, with the largest difference between treatment groups was dry mouth.

Dry mouth often resolved during continuation of therapy, and only lead to discontinuation of treatment of 3 of 906 patients. The remainder of the events are shown in this slide. The most frequent adverse events overall were COPD exacerbation, and upper respiratory tract infection.

The next three slides illustrate all serious adverse events occurring more than once in any treatment group in the one year trials. As previously noted, there was a lower proportion of patients with tiotropium who had serious adverse events.

As you can see, for any individual serious adverse events, the frequency was relatively low and the differences among treatment groups were relatively small. As with the first slide, again the frequencies

of these serious adverse events are low, with small treatment -- with small differences between treatment groups.

And we did see a difference with myocardial infarction, .5, versus .3 percent; and .8 versus zero percent. However, this pattern was not seen with coronary artery disease and angina. The more frequent adverse events are those that you might expect in a COPD population.

The most frequent serious adverse events overall, as would be expected, occurs with lower respiratory system disorders. There was a higher proportion of patients in the control groups in both sets of one year trials who had serious adverse events secondary to COPD exacerbations.

This slide illustrates all fatal adverse events in the one year trials, and have been aggregated according to system organ class due to the relative infrequency of any individual cause of death.

The proportion of patients again are illustrated according to the one year trials. In order to facilitate your review in line with the

questions posed by the agency, these four system organ classes, which encompass potential cardiovascular causes of death, are going to be broken down into their individual preferred terms for the one year and for the six month trials.

Here you see those identical system organ classes, and on top is the number of patients, and please note that there is unequal randomization, and what I am illustrating in the columns now is the absolute number of patients unadjusted for this unequal randomization.

In this case, you are now seeing the six month trials, and there was one death with tiotropium, five deaths with placebo, and six with salmeterol. There were two fatal outcomes in heart rate and rhythm disorders with tiotropium in the one year placebo-controlled trial not seen in placebo.

However, this was not observed in the ipratropium controlled trials, and in the six month trials, there were two with placebo and none with tiotropium. There were also or there was also one myocardial infarction and three myocardial infarctions

here in the one year trials, not seen in the control groups.

However, this was not observed in the six month trials. As you can see the numbers here are overall few. Given the infrequency of several of the relevant adverse advents, including the causes of death, we have conducted an additional analysis by pooling all placebo-controlled data and standardizing for patient exposure in order to permit a more precise evaluation of adverse events and to reduce random error.

We can do this because we have similar protocols and similar populations, as well as a similar pattern of response. For the placebo-controlled pooled analysis, we have computed incidents rates calculated as the number of patients with an event, divided by the patient years of exposure.

It is going to be expressed in the following slides per 100 patients years. The rate difference is hence the incidence rate in the tiotropium group, minus the incidence rate in the placebo group.

A positive rate difference indicates a higher rate with tiotropium, and a negative rate difference, a higher incidence rate with placebo. P-values have also been calculated to take into consideration the statistical reliability of these rate differences.

The events included in this analysis were selected on the basis of clinical relevance to the compound; that is, anti-cholinergic effects, or to the patient population, particularly cardiovascular and spirotory events.

This slide illustrates the population taken for this additional analysis, and we combined the one year placebo controlled trials, and identical arms from the two six month trials, and standardized them for patient exposure, and that just adding these patients gives you 952 tiotropium treated patients, and 771 in the placebo group.

The incidence rates and the rate differences for the pertinent cardiac events are illustrated in this slide. The top row shows the patient exposure, and note with this combined

analysis, we can achieve 679 patient years of exposure to tiotropium.

The end refers to the number of patients, and the rate is the incidence rate, and the RD is the rate difference obtained simply by subtracting these two rate columns, and the last column is the p-value associated with the rate difference.

Again, the p-value is there to assess the statistical reliability of these rate differences. The rate differences and rates again are expressed per 100 patient years.

The rate differences for all of these events are low, and you see both positive and negative rate differences, and the p-values are all high. As an additional step, we sought to further understand these cardiac events by combining terms that might indicate physiologically similar events.

This slide shows the combination of terms in a similar display, and you see the rates, rate differences, and p-values. And when we combine a tachycardia superventricular tarchycardia and atrial fibrillation, it showed a positive rate difference of

1.4 per 100 patient years with a lower p-value.

And it suggests that there may be an anti-cholinergic effect of tiotropium on increasing heart rate. The lack of, or the relative lack of findings in the evaluation of vital signs and thousands of ECGs is actually consistent with this analysis, in that it indicates these events are infrequent or rare, and are predominantly transient in nature.

We have combined angina and angina aggravated coronary heart disease and thrombosis coronary, and separated it from myocardial infarction as myocardial infarction could reasonably be considered a more serious manifestation of ischemic heart disease.

A combination of these terms shows a positive rate difference lower than the preceding one, and a weak association to treatment. However, turning to myocardial infarction, there is no difference between treatment groups.

Now, as you have seen, most of these deaths that occurred in the long term trials were from cardiovascular disease, or appear to be from

cardiovascular disease. We therefore evaluated total cardiovascular mortality and further distinguish them into ischemic deaths and arrhythmic deaths.

For arrhythmic deaths, we have taken the most conservative position and any event reported as cardiac arrest, sudden death, arrhythmia or death, we will assume it is related to arrhythmias.

When we have done this, you see that there is no difference in ischemic deaths, arrhythmic deaths, or total cardiovascular mortality. And finally we have looked at all that cause mortality, and this shows a negative rate difference that is a lower rate with tiotropium.

The pooled analysis also confirms the expected pharmacological effects of tiotropium. Anti-cholinergic effects, such as dry mouth and constipation, and positive rate differences, and low p-values.

We also saw positive rate differences for upper airway events. However, the most profound respiratory effects were with COPD exacerbation, with a much higher rate in the placebo group. There was

also a higher rate difference, a higher rate with adverse events reported as dyspnea in the placebo-treated patients.

Micturation disorders, urinary retention, and urinary tract infection showed positive rate differences, with low p-values, suggesting an anti-cholinergic effect on bladder contractility.

To summarize then, the core of the clinical adverse event analysis has been based on long term studies involving over 1,300 COPD patients participating in long term trials. The analysis of these long term trials, in combination with the evaluation of vital signs, lung function testing, lab testing, and thousands of ECGs, has allowed us to characterize the safety profile of tiotropium.

Events have been observed that are consistent with anti-cholinergic pharmacology and include superventricular tachycardic arrhythmias, dry mouth, constipation, and urinary tract disorders.

While there appear to be some numerical imbalances between key treatment groups, the results of our analysis show that there is no association of

tiotropium with life threatening events.

In conclusion, the safety profile of tiotropium is consistent with establishing anti-cholinergic therapy that has been used in the treatment of COPD. I thank you for your attention today, and I would now like to turn over the podium to Dr. James Donohue.

DR. DONOHUE: Thank you, Steve, and good morning, Mr. Chairman, and members of the committee, and members of the FDA, and guests. It is a privilege to have the opportunity to speak to you again on behalf of this medication.

My role is as a practicing pulmonologist for the last 25 years or more, and I have been involved in clinical trials with bronchodilators since the early 1980s. The first point that I would like to make is that as far as our patients with COPD go, there is a huge unmet burden in the United States and around the world.

We have a very large number of people with this condition, many of whom are not diagnosed, and many of whom are under-treated. A couple of weeks

ago, David Mannino published in the Morbidity and Mortality Weekly Report these statistics on COPD from 1971 to 2000.

And a couple of points are very meaningful. First of all, of course, the death rates have gone up, and the number of women affected with COPD also has gone up. But very importantly there is a large number of patients who have not yet been diagnosed, and of those who have COPD, over 38 percent said that their activities were limited.

Another piece of information came from the survey confronting COPD in America a couple of years ago, 58 percent of patients with COPD complained of dyspnea daily, and 24 percent had dyspnea even at rest, and 70 percent of patients who walked up a flight of stairs had shortness of breath.

So there is really a very, very large burden of unmet needs out there in our country today. On the other hand, we have patients who have COPD, and who tend to be older folks, and they are often in their sixties, and many have co-morbidities.

So we have to be very careful, of course,

with medications that we used. And at the present time our therapeutic options are somewhat limited in taking care of patients with COPD. First, we had the methylzanthines, and the theophyllines, and they are limited by drug interactions in older people, and a narrow therapeutic window.

The short-acting beta agonists have been around a long time, and they are limited because they have to be used every six hours, and not that beta-specific, and there is some tolerance with them, and some cardiac toxicity.

We have had oral beta agonists, which also suffer from an adverse -- in some cases adverse toxicity type of profile. The longer acting beta agonists of course are an improvement, but have to be dosed every 12 hours.

We have oral systemic corticosteroids, but they suffer from really very severe side effect profiles, and as we discussed back in January with you, the inhale corticosteroids have still not been approved for COPDs.

So we are limited in what we have to offer

our patients today with this condition, and I would like to just focus for a moment on the well-known Fletcher and Peto curve, describing the natural history or the life history of a patient who suffers with COPD.

And on this axis function from a hundred percent to 25 percent, and on this axis is as we age from 25 to 75, and there are patients who are on the blue curve here, and on the top line would be an individual who does not have COPD, and after our lungs are grown, we lose about 24 to 30 MLs per year.

Our patients with COPD are on this curve here, and often come to medical attention in their 50s and 60s when they are becoming short of breath. And they lose, 50-60 MLs per year, and patients with alpha1-antitrypsin as Dr. Stoller here has shown, will lose a lot more rapidly, maybe a hundred MLs, or something like 85 to a hundred MLs per year.

But the loss of 50 MLs is very, very significant when we think about how long the lung function is of our patient, and when we look at bronchodilator effects, even small changes, like we

see with the bronchodilators and COPD are often highly meaningful in a patient who is losing function at this rate.

The bronchodilator effects that we have heard today I think are to me as a clinician, and also as an investigator are very impressive in their magnitude. First of all, what about the bronchodilator effect. I think a once daily dose really is important for our patients.

They don't have to get sick, and require dosing themselves every six hours, and they don't have to -- they can just get by with once a day dosing, and this will I believe enhance compliance, and it certainly has with other medications, but I think it is going to make life a little bit easier for our patients.

Now we want to talk a little bit about the trough and what do they numbers mean, and that is the value when you first get up in the morning, and that is the worst time of day for patients who have COPD, and if anybody in the audience has it and comes forward, they will tell you that.

And that in the early morning hours, and at 5:00 and 6:00 a.m., and at 7:00, they are more symptomatic. In fact, when I was a young lung doctor, we used to have our surgeons do thoracic surgery in the afternoon because our patients with COPD do so poorly in the morning.

So to me, looking at a trough level of 140 mls is very, very significant, and I think it really will help our patients, and particularly in the early morning hours. We heard the -- Bernd showed the average effects. Remember that this is not asthma. This is COPD, and so these changes of peak effects in the high 200s, and the average effects in the 200s is really excellent for a drug that we use to treat people with COPD.

And I guess I was very impressed by the forced vital capacity changes. As Bernd mentions, hyperinflation is really one of the major causes of dyspnea that our patients suffer with, and these improvements in the FVC are very similar to what we got years ago when we were studying aerosol solutions for patients with COPD.

So I think the magnitude of those changes to me as a doctor are pretty impressive. They also, even though the trough effect is still 60 percent of the maximum effect, the patient still derives extra benefit every day by taking the next dose, and they still get an additional peak effect.

We have consistent sustained bronchodilation through the day, and I think that will translate into patients having less symptoms over the course of the day. And more importantly, these are long studies, one year and six months, and there has been no evidence of any loss of efficacy.

And you have got to look at that with the idea again that the patient with COPD is an older individual who is losing function and going downhill, and so I think that the fact that the drugs are still working are very impressive. I think those big bronchodilator effects explain what we are seeing with dyspnea.

Now, dyspnea, of course, is why our patients come to medical attention, and this is why their activities are limited. But like everything

else in COPD, it is really a hard thing to quantitate and get a real handle on. People know that they are short of breath, and if the drug works, they can tell you that they are no longer short of breath, and that is certainly very, very impressive.

But it is a highly complex subjective symptom, and this is why we have trouble with it, and patients will alter their activity to avoid this unpleasant sensation. They will sit down and become couch potatoes.

So a real big part of our comprehensive COPD program is to get the patient up and moving again. So you have to take into account the patient might not complain of this because they are not doing anything.

And so that is very important and very key when we are analyzing again some of the instruments and what have you. Individual patients vary considerably in their evaluation, and as you know,for many good clinical trials, there is a substantial trial effect or placebo effect. Patients get very good medical care and they tend to get better just

because of being in the therapeutic trial.

Also, we have the co-morbidities. Lots of our patients with COPD are anxious and they hyperventilate, and many are either overweight or underweight, and many are deconditioned, and may have co-morbidities, like heart disease. So this influences greatly the evaluation of dyspnea.

Nonetheless, what impressed me about these studies, and to be a robust effect, we have here a multi-center, multi-national study, and we still have consistently pretty good effects when it comes to the dyspnea scales.

Now, again, I am no expert in the scales, but we certainly used them widely. The Mahler Transitional Dyspnea Index is widely used in clinical trials. Gosh, all the studies we are doing now, we have that in our program.

We have had consistent results across these six studies at least at the end point of a one unit improvement. When we look at some of the other studies that we have done, we have not seen that as consistently with other medications that we use and

the bronchodilators that we use in COPD.

When we go to other outcomes, like rescue albuterol use, and some of these symptom scores, there is correlation. Also, one of the things that has been helpful to me was that I was just looking the other day and reading the paper in the New England Journal of Medicine on primary pulmonary hypertension, a disease that causes terrible shortness of breath, and they used for a new medication the TDI and the patients improved 1.4, just to give you a different disease perspective.

Now, what about the responder rate? Consistently we were seeing over 40 percent, and in the paper that I published, 43 percent responder rate. And when we look at the enormous number of people who suffer with COPD, this is a very big number to me, and I think is highly meaningful in taking care of patients.

I think that as all of you here who are experts in clinical trials know, the placebo responder rate, regardless of the type of study that you are looking at, is always quite high in clinical trials.

And, of course, I would like to focus on the very impressive number of patients who have responded to this on medication with improvement in their dyspnea. Now, what about the safety?

Well, the anticholinergic class has been around a long time, and we have been using them since 1987, and they really are the -- and atrovent in particular, and comparable drugs that we have used in a variety of clinical situations; with outpatients, with inpatients, and in critical care.

And they really do have a very, very strong safety record, even in the most severely ill patients. The thing that I liked about this program, as opposed to other clinical studies years ago, where we went six weeks, or three months, here we went one year with four studies, and six months with two others.

So we have a very long duration of exposure to our patients. In my view the patients in these studies are reflective of the patients that I see every day in my practice. They are the same age group, in their 60s, and 10 years of diagnosed COPD,

and same tobacco history, and same numbers of co-morbidity.

So I think the patients are highly reflective of what we have seen in other clinical trials, and what we see on a daily basis when we are taking care of patients with COPD. We had a very low incidence of adverse effects, and as most of you know, these are the anticholinergic effect, the dry mouth.

And usually these patients will work through that and they will continue on taking the medication. So I was greatly reassured in this older population that the safety data were fairly good. Boehringer Ingelheim asked me to make a comment; where do I see this drug being used.

And I think based on the very strong bronchodilator data, as well as the efficacy as far as dyspnea goes, I would see that as a first line chronic maintenance therapy for patients who are symptomatic with COPD, and really of all variance severities, from mild to severe.

Just a comment. Nationally, the government has articulated a project called, "Healthy

People 2010," and a number of health goals. And there are two goals that are relevant to COPD. One is a reduction in the mortality rate, from 119 per hundred-thousand, to 60 by 2010.

And at this point, no medication has been shown to affect mortality. The second goal though I think is more relevant, and that is to reduce the number of people whose activity is limited by breathlessness, from 2.2, to 1.7, and I am hoping that tiotropium will be an effective tool to help us accomplish this.

On a personal level, I think there is still a lot of our patients whose needs have not been met, and I think that the increased awareness of dyspnea might lead to more diagnosis of COPD and a more willingness on the part of doctors to try medication in this population.

I think at the present time that our (inaudible) is quite limited for what we have to offer patients, and I am very optimistic that tiotropium will provide a worthwhile addition to our (inaudible).

I want to thank you very much for the

chance to express my comments, and Dr. Blank will make the concluding remarks. Thank you.

DR. BLANK: Thank you, Dr. Donohue. In my conclusion, I want to come back to the questions that the agency brought to this committee, and share with you Boehringer Ingelheim's position on these topics.

The safety of Spiriva was studied in one of the largest programs conducted in COPD so far. The safety profile shows anticholinergic pharmacology for Spiriva, including an association with rare superventricular tachyarrhythmias, and it is very similar to what has actually been widely used for many years.

The safety profile is described in the label that we have proposed in our submission to the agency, and most importantly there is no association with life threatening events.

Twenty-four hour bronchodilation, after once daily inhalation of Spiriva, has been consistently demonstrated in all six week studies, and its effect remains fully sustained throughout chronic therapy.

The improvement of dyspenea was shown in two pivotal studies with a validated instrument. The improvement by one unit in the TDI, which is the definition we used for treatment response is relevant to an individual patient and for the COPD population as a whole.

We have met the regulatory requirements for the indication of the relief of dyspnea associated with COPD. In medical practice, most patients with COPD seek medical care because of their dyspnea, and physicians monitor their patients according to their symptoms.

Spiriva improves dyspnea, the key symptom of COPD, which has the greatest impact on the patient's lives, and this improvement should be described in the product's label. We believe that the most appropriate place for this is the indications and usage section as outlined in my last slide.

Thank you very much for your attention and that brings us to the end of Boehringer Ingelheim's part, and we will be glad, my colleagues, and I, to answer any questions that you may have.

CHAIRM DYKEWICZ: Thank you. Before I entertain questions, Dr. Atkinson had joined us just before the product presentation, and if you could introduce yourself to the group.

DR. ATKINSON: Yes. I am Prescott Atkinson, and I am allergist/immunologist from the University of Alabama at Birmingham.

CHAIRMAN DYKEWICZ: Thank you. First of all, I would like to compliment the BI people for staying on time. My first personal question is related to slide 16 of Dr. Witek's presentation, which shows the date of -- I believe from the two pivotal studies relative to mean TDI focal scores, and I wanted to make sure that I understood the data. If we could perhaps have that projected, please.

DR. BLANK: Dr. Witek, please.

DR. WITEK: Slide 16, please. If you could supply that. Well, just to reexplain the slide here. This is looking at the mean TDI focal score in Study 114 and Study 115. So these are the two separate, one year studies.

On the y-axis is the focal score and the

x-axis is time. Day 50 was the first assessment point for the TDI, and the mean TDI score in the tiotropium group was a little bit over one unit as you see here, and the placebo group, .2 units. So a mean improvement needs group to that magnitude, and this is describing the TDI changes over the course of the year, and then in this second graph we see the same pattern.

CHAIRMAN DYKEWICZ: All right. Now, it has been suggested that the clinically meaningful difference in the TDI scores is about one, and we will of course be discussing that as a committee later.

But looking at that, it seems to me that at least in terms of Study 114, for the five time points, did not achieve that difference, at least between placebo and the tiotropium. Is that correct?

DR. WITEK: That is correct.

CHAIRMAN DYKEWICZ: Okay. My other question was looking at the document that was given to us by the FDA relative to Mahler's screening instrument, the chest article from 1984, it was suggested that inter-rate variability using that

instrument would be no more than one.

Now, I presume that during these large-scale studies different raters were rating people at different time points?

DR. WITEK: Well, in each of the clinical centers, we had a study coordinator, and whenever possible, that study coordinator would be the same individual. However, that wasn't always the case.

CHAIRMAN DYKEWICZ: Thank you. Let's open up the floor to questions from the committee. Dr. Patrick.

DR. PATRICK: While we are on that, could you just explain how the TDI was administered, and at what point it was administered after the SGRO (sic), and I believe before some physiological measures? Were the results of the SGRO (sic) available to the raters of the TDI, and were the people who did the rating trained to some level of kappa agreement prior, and similar to other clinician rating scales?

DR. WITEK: No, there was -- you know, no inter-rater analysis in the multi-center or large-scale studies among those coordinators. To go back to

your first question, Dr. Patrick, regarding the SGRQ. There was actually specific instructions in the trial for the coordinator to review the last page of the SGRQ, and this is atypical for the TDI instrument.

However, that was done to help the coordinator and patient remind them of their activities of daily life, which is what is listed on that last page.

CHAIRMAN DYKEWICZ: Ms. Schell.

MS. SCHELL: I noted that you start the interviews on day 50. Were there any pre-interviews done regarding their level of activity?

DR. WITEK: Yes. In the clinical trials described here, the long term clinical trials, those were the first assessment points. We do have in the one year tiotropium controlled trials, an assessment as early as day eight, and that was relative to ipratropium bromide.

And there we did see responder rates, and mean effect size, and a rate higher in the ipratropium relative to placebo. We have other small studies where we have earlier measurements submitted in the

NDA.

CHAIRMAN DYKEWICZ: Dr. Joad.

DR. JOAD: Yes, I have a question for Dr. Kesten, which is about the Holter monitors that were done in the one study. I know that you didn't bring that up, but in our briefing packet it mentioned decreased heart rate variability, which I understand is associated with morbidity and mortality to heart arrythmias, and I wondered if you would care to comment on that.

DR. KESTEN: Thank you for the opportunity to clarifying that point. First, the issue of heart rate variability and applicability. As you noted, there has been association in context with a clinical event.

There has been no associations of pharmacological induced changes in heart rate variability in such events. That being said, the differences in heart rate variability changes were extremely small and just suggested that we could see a pharmacological event.

And I would actually like to turn this

over to Dr. Prystowsky who has reviewed the information on that as an electrophysiologist for his opinion.

DR. PRYSTOWSKY: Thank you very much, and I appreciate the opportunity to address the panel. I am Dr. Rick Prystowsky, and I am a clinical electrophysiologist, or known more to my patients as the electrician of the heart, and basically I have had an opportunity to review all of the Holter data and actually all of the cardiovascular data from the study.

I have had a major area of interest in my own research career in autonomics in the heart and heart rhythms, and this whole -- and let's talk first of all about the issue of the heart rate variability. This is sort of the test du jour of some of the researchers in our field.

We have seen these patterns, as I am sure in pulmonary, go in and out at what people like to look at research wise. There has been an association and it actually dates back several decades of looking at heart rate variability in a more simple matter in

patient's post-MI, patients with significant left ventricular disfunction.

And there appears to be a correlation of lower heart rate variability in these patients, and in some studies there appears to be an increased incidence of sudden death in overall cardiovascular mortality.

Sometimes the -- and the multi -- you know, the regression analysis, it just barely will make it, even though it is an independent predictor. But in a population like this, there really are no data or no data at least that I know of that any kind of heart rate variability means much.

It has been known for decades in diabetics that as they get parasympathetic disfunction, there is a lower heart rate variability. So I think it is there, and it would make sense knowing the effects of the drug.

Any anticholinergic -- you know, man is basically a vagal animal regarding the effects of the sinus node with autonomics. The parasympathetic is clearly prepotent over the sympathetic, and you can

take a patient who is getting isoproferenol infusions in your lab, and if they get a vagal effect, they can go totally asystolic for 10 or 15 seconds, even in the presence of high sympathetic tone.

So the vagas runs the heart as far as sinus rate goes, and of course heart rate variability is clearly related to that, and if you sort of put that all together and you say now I will give somebody an anticholinergic, even if it is a minor anticholinergic, one would anticipate from the pathophysiology or the physiology of the autonomics, that you would have a slight decrease in heart rate variability.

In a trial of patients that are reviewed that are sick, and clearly we have got probably more cardiovascular disease than came out in the questionnaires with all of the smoking history and their ages.

We have not seen, at least from my review of the data, any increase in cardiovascular mortality. So I think that it is an interesting point that you raise, but I think that there is really no data to

support it that has any meaning in these patients, and it follows the physiology of the drug.

DR. JOAD: While you are still standing there, I would like you to comment on the fact that people who had arrhythmias, and that were on medicines, or who have had recent Mis, were excluded from these studies. Yet, in real practice, probably it won't be as rigidly prescribed.

So what do you think the effect of that exclusion criteria had on the cardiovascular risks of this drug?

DR. PRYSTOWSKY: I think that point is very well taken. As you probably know, I was not involved in any of this until a few moments ago when I was asked to review the data. I am not part of the trials, and I am obviously not a pulmonologist.

And I think that this is something all of us suffer as clinicians, too, with any trial, is that one never sees the exclusions on any basically, and yet in real life we have to deal with that, and I think your point is very cogent.

I will tell you my feeling based on all

the data that I have reviewed. First of all, there is no reason to believe an anticholinergic agent will exacerbate arrhythmia, certainly not from the ventricle.

There is no data to suggest it and in fact in my line of work I have much more commonly stopped the sympathometic agents in patients who have come to me, and I don't think I have ever stopped any of the known -- you know, the known anticholinergic agents.

Ironically, and it doesn't make sense physiologically, I saw an increased incidence in atrial fibrillation, and that doesn't make any sense because the classic model to produce afib in a lab, either in man, even the data on this, as well as animals, is vagomimetic effects.

So if you want to produce afib, a high amount of vagal tone will do it. Why there was a little bit of discordance in the afib is hard to explain the known effects of the agent.

So what I would anticipate as far as arrhythmias go, I would anticipate no particular problems. I would not be worried about it at all,

even though some of these people were excluded, because clearly in the COPD population you are going to have people with afib.

There is no reason to believe an anticholinergic should exacerbate it. I guess the only thing which isn't tested here that one could argue that there could be, could you possibly have a slightly increased ventricular response in someone with known afib, who gets an agent that is anticholinergic?

I guess it is conceivable. It would be unlikely because whereas the sympathetic - whereas, the parasympathetic and sympathetic effect on the sinus node is markedly power sympathetic, the aging node is balanced.

Sympathetic and parasympathetic in humans is pretty much a balance situation. So the slight amount of anticholinergic effect, it may be a few beats a minute. I don't think much more than that.

So I am not at all worried about any of the arrhythmia issues. The one area of anything when I reviewed the data that I would have some concerns,

and I have expressed this to the company, would be a group that was not looked at, and probably appropriately so.

Someone with unstable anginal could be so critical with a lesion that even a slight increase in rate could push him into having an anginal episode. So I think in an unstable anginal in a patient, without any data, I would have some personal concerns about using any agent that might increase heart rate, even a little bit.

But I am not worried about arrhythmia components. It would just be more from an anginal standpoint. Otherwise, no, I don't have any real concerns based on the known long term effects of anticholinergics in these patients that we have seen for years.

CHAIRMAN DYKEWICZ: Thank you. Dr. Apter.

DR. APTER: I have two unrelated questions. The first question is the TDI, and underscores Dr. Patrick's question. I would like to hear more details of how it was administered and how the observers scored it in these trials, and then I

will come back to the second question.

DR. WITEK: The TDI was administered in the morning when the patients reported to the clinic, and so after the questionnaires on adverse events, and then the SGRQ was administered, and then the TDI was administered.

The TDI administration, as was pointed out by Dr. Patrick, patients and the caregiver, or the coordinator, referred to the last page of the SGRQ, which listed activities, which was again a catalyst for the TDI instrument, where it is an open-ended interview to look at the change relative to the BDI.

So, additionally, the BDI is looked at, and in that open-ended interview, the coordinator makes the rating as was described here.

CHAIRMAN DYKEWICZ: This is the last question.

DR. APTER: So the subject reports on what would be moderate activity, and what would be strenuous activity?

DR. WITEK: Yes, that would be in an open-ended interview.

DR. APTER: So it would be different for different people, and how would it be compared against a person, just in terms of moderate, severe?

DR. WITEK: It is within patient assessment. So it is hard to say how it would compare at those different levels. I don't know if this would address your question, Dr. Apter, but in various levels of severity, whether it be FEV1 severity, or BDI severity as an example.

We did show the same effect of the drug in those that had mild BDI and severe BDI.

DR. APTER: And maybe even Dr. Mahler could answer this, how the TDI has been and the BDI has been used in other populations, and the extent of experience.

DR. MAHLER: Thank you. My name is Don Mahler, and I am a pulmonary physician at Dartmouth Hitchcock Medical Center. I want to comment that these instruments were developed under the direction and mentorship of Dr. Alvan Feinstein, and he has been instrumental in the effort to develop instruments to provide clinical outcomes and clinical measures.

The TDI right now, again, the development -- do you want me to describe more of the development, or its use at this point in time?

DR. APTER: Both quickly if you could.

DR. MAHLER: Sure. These instruments were developed in a four-step process. We first looked at the available instruments, predominantly the MRC, and saw its limitations as Professor Jones described.

We then met with pulmonary physicians at Yale University, where we developed as Dr. Richard Matthay, and Dr. Jacob Loke, Dr. Herbert Reynolds, and kind of had informal discussions about how best to expand the MRC into these other components.

We then had a pilot testing at the VA Medical Center in West Haven, Connecticut, and that pilot involved 15 patients with COPD, and I interviewed the patients using our BDI/TDI instruments, and we had a pulmonary function technician doing the same thing.

We then both met with the same patient and said, well, you told me this, and you told me this, and based on feedback from the patient, we then put

together the final BDI/TDI, and then I guess what I would describe as the fourth step is we then applied it in these 38 patients, both at a baseline state and then at a follow-up state, and then published that information in 1984 in "Chest" is kind of the first experience validation responsiveness of these instruments.

As far as its use at the present time, it is amazing as Dr. Donohue said that these instruments are used worldwide, and there are at least 30 publications in peer review journals using the BDI/TDI predominantly in COPD population.

And that involves bronchodilator therapy, pulmonary rehabilitation programs, inspiratory muscle training, and lung volume reduction surgery. It is being used by at least 10 current companies, pharmaseutical companies in the United States, not only in bronchodilator therapy, but looking at, say, monoclonal antibody treatment, and the psydokines in COPD.

It is currently being used in an investigation of interstitial pulmonary fibrosis. So

at least from the information that I have available, I would say that it is the standard instrument that is currently being used in the pulmonary community to measure dyspnea at baseline, and particularly the responses to an intervention.

CHAIRMAN DYKEWICZ: Thank you. We are going -- oh, Dr. Stoller, one last question, and then we will have an opportunity later this morning by the way to ask some additional questions.

DR. STOLLER: I guess my question is to Dr. Witek, and it regards the administration of the dyspnea index, and having worked as Dr. Mahler knows with Dr. Kleinstein, I had an appreciation of this outcome, and my question regards whether in knowing that this is a several page instrument, and having worked with it, and the temptation to actually give the written form to patients to complete on their own, but recognizing that the instrument was developed to be administered in a questionnaire.

And my question regards whether there were any protocol violations with regard to the forms being given to patients to complete themselves, as opposed

to interviewer-lead, and what the prevalence of that. I imagine that it happened.

And what the prevalence of that was, and whether there was any concordance between subsequent interviews and the self-administered, and which instrument was recorded in the data set.

DR. WITEK: Sure. As was pointed out in your briefing document, there was one incidence in the FDA audit where patient handwriting was noticed on the diary, and that was subsequently responded to by the investigator.

Essentially, it does follow the SGRQ, which is a patient-administered instrument in the sequence of case report forms. In that case the patient had begun to fill it out. Now that was noticed, and the coordinator corrected that with a re-interview and initialed that, and that was formally responded to with the agency.

Given that, we did go back carefully again to check particularly all the U.S. centers, and we checked the U.K. centers, and we did find one other case reported by the CRA visits, which were conducted

every 4 to 6 weeks in these studies, and that was the case as I think you could appreciate, was a new coordinator that came into the study, and that visit, the first visit with that coordinator, the patient had completed the diary.

This was noted in our routine monitoring and that case was corrected. So those are the two cases, and our analysis was based on the U.K. centers, and the U.S. centers from 130, subsequent to the agency's comments and the briefing document on that.

CHAIRMAN DYKEWICZ: Thank you. We will resume in 15 minutes.

(Whereupon, at 9:33 a.m., the meeting was recessed and resumed at 9:53 a.m.)

CHAIRMAN DYKEWICZ: Welcome back everyone. We will now resume the meeting with the FDA presentation, starting with Dr. Lisa Kammerman on the transition dyspnea index.

DR. KAMMERMAN: Good morning. I am sitting up here on this stool because I am recovering from a broken leg. So when my colleagues told me good luck, and don't break a leg, that was the wrong thing

to say.

So I will be discussing the issues surrounding the use of the Transition Dyspnea Index, and the tiotropium development program. I want you to know that the primary statistical reviewer for this application is Dr. Jim Gebert, and he has done the nuts and bolts for looking at all of the data and intricacies there.

And my role really is to focus on the use of the index in these studies. My presentation will focus on the use of the TDI in the tiotropium clinical studies, and I am going to first give you an overview of the baseline dyspnea index as well, and as you consider the requested indication for the treatment of dyspnea.

It is important to keep in mind the history of TDI and how it was actually elevated from a secondary end point to a co-primary end point in the six month studies. This is a very important point because it has many implications for the clinical trial design issues that I will be discussing.

Other issues to consider include the

development and the validation of the TDI, its implementation in the 6 month studies, and the definition of a clinically meaningful difference.

The Transition Dyspnea Index, or the TDI, is the end point that is being used to support the indication for the treatment of dyspnea. Moreover, the rest of my talk will be focus on the TDI. I just want to go over again an overview of both the BDI and the TDI.

As you know, they were both developed by Dr. Mahler and his colleagues, and they described their indicates in the 1984 paper that appeared in "Chest." And copies are in your FDA background package.

Each has three components, and the focal score, which is actually the total score, is simply the sum of the component scores. It is also important to recognize that each component is actually a single item, and because you heard a lot about BWI earlier this morning, I just want to comment that two of the components in the BDI are actually highly related.

So when you look at the distribution of

BDI baseline, you see that lots of people are at six, and it might just be reflecting some double-counting when the source of the items are added together.

As you have heard, the indices are administered by interviewers who ask open-ended questions. The interviewer interprets the responses and selects the score. In order to implement the TDI, the BDI needs to be established first at baseline, and so when the TDI is actually scored, the interviewer and perhaps the patient, which we will get into a little bit, needs to refer back to the BDI.

The scores for each TDI component range from minus-3 up to a positive-3, from a major deterioration, to a major improvement. In the next set of slides I will show you the definition of minor improvement for each of the components.

And the reason that I will be focusing on minor improvement is that a TDI score of at least one was used to define a responder, and as you see on this slide, a plus-one is the same as a person with a minor improvement.

So just to remind you, the interviewer and

patient need to refer back to baseline in order to assign a score for the TDI, and here we see the definition for a minor improvement in change and function of impairment.

And this reads, "Able to return to work at reduced pace, or has resumed some customary activities with more vigor than previously due to improvement in shortness of breath."

This definition illustrates an issue with recall to baseline, reduced pace, and more vigor, and implies that either this information was recorded at the time that the BDI was administered, or that the information needs to be gleaned from the BDI itself, or that patients need to be relying on their memory.

Here is the definition of minor improvement for the change in magnitude of task that causes dyspnea, and if it has improved less than one grade from baseline, a patient with a distinct improvement within grade, but has not changed grades, there are two more points that I want to make from this slide.

First, the criteria for minor improvement

are very subtle, and the second point again is the need for a recall to baseline.

Again, for a magnitude of efforts, relatively subtle improvement could be graded as plus one; able to do things with distinctly greater effort without shortness of breath.

For example, may be able to carry out tasks somewhat more rapidly than previously. Again, this requires the interviewers to assess a subtle change, and to remember what was occurring at baseline.

The total score, which is the focal score, is obtained by adding together the scores for each item. The focal score can range from minus-9 to positive-0, where the positive number indicates an improvement from baseline.

So why was the TDI developed? The goal was not to address differences in drugs using clinical trials, but its goal was more clinical in nature. Until 1984, as you have heard, dyspnea was assessed in the clinic by looking at the magnitude of tasks needed to induce breathlessness.

And Dr. Mahler and his colleagues believed that better clinical measurement required assessments of the functional impairment and magnitude of effort that causes dyspnea.

And he also wanted to obtain a measure that could be used by different interviewers, and which could produce consistent results among interviewers. Now, these are the four year studies.

The first two, 114 and 115, were conducted in the United States; and the second two were conducted in Belgium, and I believe in The Netherlands. The applicant explored the data, and saw that patients were classified as responders, with an improvement of plus one.

And then tiotropium was statistically different from placebo. Responders were defined as those who had a score of at least one, and when you consider the results for the responders in the one year studies, it is important to realize that I think around only 55 percent actually had a TDI reported at one year.

The rest of the information has been

imputed by different roles. So the applicant came and met with FDA as you heard in July of 2000 to discuss their intent to elevate TDI from a secondary end point to a co-primary end point in the six month studies.

The studies had been completed, but they were still blinded. The major change was to the study hypothesis, which in their amendment now reads, "The proportion of patients with a TDI focal score greater than or equal to one unit is different than those treated with tiotropium, compared to those treated with placebo.

So what went on at that meeting, the FDA and the applicant agreed that TDI could be promoted to a co-primary end point. However, the following conditions needed to be met, and the applicant needed to justify the clinical significance, but when it needed change in the TDI, both for the comparison of the mean scores, and for the comparison of the responders.

Again, the responder was a one unit change, and so he wanted to see validation data for this one unit change as being clinically meaningful,

and this would include showing the TDI correlates to the clinical improvement of subjects.

It is important to note, however, that we did not agree that the Studies 130 and 137, the six month studies were adequate to support an indication, and that would be a review issue.

So now that we agreed that TDI could be a co-primary end-point, our next step on the FDA's part was to review the NDA, and to set the context for the rest of my presentation, I just want to give you a very, very brief overview of some of the issues that I look at, and that my colleagues look at when we review an NDA.

The first thing that I look at are the study protocols. I read them and assess their clarity, their completeness, and scientific merit, and then I look at the conduct and the analysis of the studies. And I will compare what was actually done in the studies with what was stated in the protocols.

And I also assess the quality of the conduct and look at the issues related to patient discontinuations. So in terms of many of these

issues, as I discussed TDI as it was used in these clinical studies submitted to the NDA.

So these are going to be the five major areas that I am going to focus on and that we identified in a review as being of concern to us, and they include the clinical trial design, and the development of the instrument, and the validation of the instrument, and the implementation of the instrument.

And the definition of a clinically meaningful difference, and I am going to discuss these first in general, and then I am going to turn to specifics.

So as you know now, TDI was originally a secondary end point. It became a primary end point after the studies were completed, but before the data were blinded. And I believe that the studies may have been designed and conducted differently if TDI had been defined from the outset as a prospectively defined end point.

And this has important implications, as you will see during the rest of my talk. For example,

there is lots of issues resulting in the implementation of the instrument. For example, there is issues regarding the training and blinding of the interviewers, and the issue of recall to baseline.

The second major area of our concern was the development of the index, and the goals were to improvement clinical assessments of dyspnea and to obtain a scale that could be used by different interviewers.

And the goal was not to develop the TDI for use in clinical trials for new drugs. The clinical nature of the TDIs were reflected by these next bullets. There is no evidence that patients were involved in generating items reflecting aspects of dyspnea that are of concern to them, and it appears that clinicial judgment was used.

And also that the TDI uses non-standardized questions. The population used to develop the TDI was from the United States, and there were no international settings used, and this is extremely important, because the six month studies were conducted in 18 countries.

The third area of our concern is the validation of the TDI. The big issue is the lack of validation of the TDI when it was translated into other languages, and for use in other cultures.

And the validation needs to be specific to the format and wording of the instrument. Every change in format and wording requires validation, and this was not done, and I will show you some examples shortly.

Also in the clinical studies, TDI was administered immediately following the SGRQ, and the ordering of the tests also requires validation, which was not done, particularly in this study where the interviewers were instructed to look at the SGRQ before administering the TDI.

The fourth area is the implementation of the TDI, and there is more evidence in the NDA that interviewers were trained or were blinded to the patient clinical status. This is a major concern of ours, because it could lead to bias in the TDI assessments.

There is also much ambiguity in whoever

completed the form. Was it the interviewer or was it the patient who completed the TDI? Here again we see the issues regarding the ordering of the instruments, the SGRQ, and the TDI, and the multi-national location populations in the six month studies.

The fifth area is the clinically meaningful difference of one unit. Again, there is no evidence of patient involvement, and there is no evidence of a pre-specified plan.

Ideally, the building plan for the TDI or any patient reported outcome that is going to be used in a clinical study should have been prospectively addressed as part of the development program.

Now I want to turn -- and this is just a repeat of the slide I had on earlier. I am going to go over the specifics of the slide that I had on earlier. I am going to turn to the specifics of each of these five major concerns; the clinical trial design, and the validation and implementation of the instrument, and the definition of a clinically meaningful difference.

And I know that my presentation is going

to sound redundant at times, but it is because many of the specifics, cross many of these five basic areas. Okay. We are going to the major area of clinical trial design issues.

The TDI was interviewer driven, but by this I mean that the interviewers were instructed to review the SGRQ before administering the TDI. They then asked open-ended questions, and both of these questions could lead to bias in the results of the TDI.

Another major area was the blinding of the interviewers. There is no indication that interviewers were blinded to the clinical status of the subjects, their treatment status, and their adverse events.

For example, if a patient reported dry mouth, this might have led the interviewer to believe the patient was receiving tiotropium because tiotropium is an anti-cholinergic.

And as you know the SGRQ was administered before the TDI, and so the interviewers were sensitized to the patient's reports of dyspnea. For

the training of the interviewers, we have no assurance that interviewers were trained, and this is particular important because the questions to patients are not standardized.

And here is the description of the open-ended questions and the intent of the questions, and this comes from Dr. Mahler's article in 1984. Open-ended questions concerning the patient's breathlessness, the intent was to allow the observer an individual's dyspnea as part of the usual or standard questions asked of a patient when taking a history of respiratory disease.

And the applicant, as you heard earlier, is placed in your background, and in a 1995 article by Eakin to support the validation of the BDI and TDI, and she points out the need for creating in both indices as you will see in my next two slides.

For the BDI, she says, in our experience to use this instrument reliably, it was necessary for our four raters to discuss some standardized questions, and to come to some consensus as to how ratings should be made on each one of the three

scales.

Ongoing assessments of inter-rater reliability to check for tendencies of each rater to stray from initial standardization was also needed. So here we see the need for the interviewers to come together to reach consensus on grading the three aspects of the BDI, and the need for ongoing assessment for inter-rater reliability.

And for TDI, she says that the TDI may be affected by bias on the part of the patient and interviewer, because it asks both individuals to make judgments about improvement, versus deterioration in the patient status and space line.

And like the BDI, the TDI lacked standardized questions for raters. So she highlights the potential for bias because of the requirements for making judgment about the past patient status relative to baseline.

And she also points out the lack of standardized questions. The ordering of the instruments is also critical. The TDI immediately followed the SGRQ, and the SGRQ may have influenced

both the patient's responses to the TDI, and the interviewer's questions to the patients.

The recall for baseline is another issue because these studies were one year and six months observation. For example, I personally would have trouble remembering my health status six months ago, or even one year ago, and Dr. Mahler's 1984 article and other studies in the literature, to be very much shorter in duration, where a baseline may be much easier to remember.

I will now turn to some of the issues regarding the development of the TDI. First and critically, there is really no indication of patient involvement. So key issues that have not been addressed include the reading level, the comprehension, and the interpretability and recall of the baseline on the part of the patients.

And we don't know if the three items in the TDI and their wording captured aspects of dyspnea that are important to patients.

The responses appear to be equally spaced, but they are not -- and I won't take up time here, but

if you look at some of the gradings for the three parts of the TDI, you will see that they are not equally spaced.

What is also very important is that the three items are simply added together without any rationale being provided in the NDA, and so we don't know if this is optimal or if there are items that are so highly related that they are being double-counted.

When I initially looked at the data, it does appear that two of the items are related. I think there is about 45 percent of patients who would answer the same for two of the components.

Regarding the validation of the TDI, here are some general comments. Again, there is on pre-specified plan, and most of the validation information that you heard this morning was really for the BDI and not the TDI.

I think there is about six slides for the BDI and two for the TDI. One report for the TDI this morning was a rehab study, and another is information that we have seen only for the first time this morning, and it wasn't submitted to the NDA, and it is

not in your backgrounder, and so we have not been able to look at it.

And a few of the validation studies that are referenced by the NDA are actually drug intervention trials. Again, there is the issue of the order of the administration of the TDI and the SGRQ.

And in the paper that is in your backgrounder regarding or by Witek and Mahler, I think it is going to appear sometime this year, the applicant supports the validity of the TDI, but describing statistically significant correlations between TDI and other outcomes.

And this again is reported both in the NDA and in this article. It is important to realize that this information in this paper is from the one year studies and most of the correlation given to you this morning was for the BDI and not the TDI, and rather than focusing on P-values -- and by the way it is not all that difficult to get significantly significant correlation co-efficients.

It is really more important to look at the co-efficients themselves, and they range from .22, or

minus .2, to minus .35 in the one year studies. The correlations are an indication of the linear relationship between two variables.

Another way of looking at them are the R-squared value, and simply you square the correlation co-efficients, and the amount of variation explained ranged from 5 percent to 12 percent, and this says that the amount of variation explained by fitting a straight line between TDI and other outcomes wasn't related that much at all.

Now, turning to the validation that relates to the multi-national studies, and there are quite a few problems here, all of the studies referenced in the NDA supporting the validation of the TDI were conducted in the United States.

The indication for dyspnea rests on the six months studies which were conducted in 18 countries, and there were approximately 600 subjects per study. In Study 130, 12-1/2 percent came from the U.S., and in Study 137, only 5 percent came from the United States.

There are numerous issues that we don't

know about. We don't know about the process that was used to translate the TDI, and we don't know the background of the translators, the quality of the translators of the translations, and there is no information on whether the translated versions were actually validated in the language and culture that they were being used.

So ideally we would like a translator who is fluent in both English and the target language, and a translator who is knowledgeable about dyspnea and is aware of cultural differences, and how that might impact the wording of the TDI.

Also, when translating indexes, it is important to translate them back to English so we can compare the translated version with the original version. If the back translated version is much different from the original, then it most likely needs to be retranslated again, and all translated versions need to be validated.

There is a memo about he content validity, only because the patients weren't involved. We don't know to what degree the TDI represents the three areas

of interest. Their functional impairment, the magnitude of tasks, and magnitude of effort.

The validation also needs to be specific to the version, including the wording and the format used in the clinical study. The formatting and wording of the TDI that was used in these studies really is not the same as was described in the 1984 paper.

The next two slides, I will show you some of the differences. The differences that you have seen may be very subtle, but even very subtle changes in the appearance of the index could be important.

And the best practice is to use the same format that has been validated. Okay. You are going to look at this slide and the next and say, well, what is the difference. But, moreover, if you are able to read this, I have only selected out as you can see three scores from one of the components.

And in this case it is the change in magnitude of task. It is important to notice where the italics are used for the name of the component and for each category, major deterioration, moderate, and

minor.

And it is also important to notice that a line is preceding each score. I have never been clear on whether interviewers were supposed to check that, put an X, circle a number, and that wasn't discussed in the NDA.

Now, if you look at the next side, this is what is in the case report form, and again this is just part of the case report form. So it loses some of the visual impact. Each of the components now has a number preceding it. Here we see number two preceding the change in magnitude of task.

The line preceding each score has now been eliminated, and so the intent was probably for the interviewer to circle the numbers, and now there is also a box around each component.

So when you look at the case report form, you see these boxes popping out at you. The font that is used is also different, and in a little while I will show you an additional important difference regarding instructions.

So who actually completed the TDI? There

is a lot of confusion. The answer is did the subjects, did the interviewers, and the answer is that in some cases the patient did, and in other cases the interviewer did. This is inconsistent with the proper way to administer the TDI.

And what led to this confusion is that the protocols are internally inconsistent. One part of the protocol says the observer should ask open-ended questions concerning the patient's shortness of breath and how it affects their daily life.

The observer will rate the patient based on the responses to these questions. And here the protocol indicates that an interviewer will complete the form. And elsewhere the protocols indicate that the patient will complete the TDI, and we see that patients will perform the shuttle walking test, and complete the questionnaires; and if SGRQ, the Mahler Dyspnea Indices.

The Division of Scientific Investigations audited two clinical centers, and this is standard practice for the division to go out and look at clinical centers, and they found that at one center

that the patients themselves read the questionnaires and completed the form.

And keep in mind that there are approximately 80 centers that remain unaudited, and you heard that the applicant went to the U.K. and found another center there where the patients had completed the TDI, but that still leaves unanswered the question of what went on at these other 80 centers in the six month studies.

Another source of confusion about who is completing the TDI is the instructions in the protocols. Here the protocol correctly suggests that the interviewer does the TDI, and it says for the magnitude of task, review the activities that cause breathlessness, ask the patient which activities now cause breathlessness, and is there a change from baseline in the selected rate.

But the instructions on the CRS suggest that the patients completed the TDI. And here at the top, you see that it says to circle one answer which describes best how your daily activities are influenced by your respiratory disease.

And notice that this instruction does not appear in any way on the original TDI described in the 1984 paper. I think it is also interesting to note does the subject know what daily activities mean, and do they really know what it means to be influenced by your respiratory disease.

I think all of us are probably comfortable with that, but in the general population I am not so sure, and there is no evidence that was presented to look at that. As I mentioned, bias may have been introduced because interviewers were possibly unblinded to the patient's status.

Again, this is an ordering of the SGRQ and in the TDI there is the issue of the recall of the baseline, and ideally we want an independent interviewer who is unaware of SGRQ, and the FEV1, and other spirometry data, adverse events, and other available patient status information.

Now, turning to the clinically meaningful difference, again there is no piece specified plan in the development process. The Witek and Mahler paper simply states a one unit improvement is likely quite

meaningful to the individual patient. There is no evidence of patient involvement in determining a meaningful change.

And this morning the applicant put up a quote, and I just thought it would be interesting to refer back to that from Guyatt. That a clinically meaningful difference is the smallest difference in score which patients perceive as beneficial.

Now, I am going to summarize my comments in a way that is slightly different from the way that I presented them, and so here are some of the issues that we have identified regarding this at the patient level.

There is an unknown level of involvement and this is important regarding the importance to the patient of aspects of dyspnea and the magnitude of the one unit change. There is the issue of their reading level, comprehension, and interpretability of the TDI, and they may not be able to recall to baseline at 6 months and 12 months.

At the interviewer level, we have this issue, and the blind indication status, the trainings,

nor assurance of the training, open-ended questions, non-standardized questions, recall the baseline, reviews the SGRQ, and possibly new other clinical data before administering the TDI.

So is a one unit change meaningful to patients, and we really can't be sure, primarily because of the lack of patient's involvement, and the absence of a pre-specified plan.

We don't know who completed the form, and in some cases it was the patient, and in some places it was the interviewer. The issue of multi-national populations in the six month studies showed up in several areas that I have gone over.

There is the impact on the development and validation, and interpretation of the results, and what I also want to emphasize is that the linguistic and cultural issues, and the quality of the translations, and the absence of validation studies and languages other than American English, because British English and American English are actually quite different.

The development was interviewer based, and

was not patient-based. Patients weren't involved in generating items that were important to them, and the TDI was not developed for use in multi-national populations.

The validation has not addressed the order of the administration, the formatting used in the studies, and its use in multi-national studies. So that completes my comments, and I would thank you for your time, and now Dr. Sullivan will address the clinical aspects of the NDA.

DR. SULLIVAN: Good morning. My name is Gene Sullivan, and I am a pulmonologist, and I am a medical officer in the Division of Pulmonary and Allergy Drug Products. I am also the primary medical reviewer for NDA 21-395, and I am going to spend the next hour or so summarizing the findings of the agency's medical review of the application.

Before I begin, I want to be sure to acknowledge the contributions of the reviewers from both the Division of Biometrics, and the Office of Clinical Pharmacology, Clinical Biopharmacology, because some of the points that I am going to make in

my presentation were generated from their reviews of the application.

This slide provides the structure of my presentation, and I am going to begin with some background remarks, and in that section I am going to highlight some of the division's thinking in regard to the labeling of drugs for COPD, and I will touch on how labeling considerations may sometimes impact the choice of clinical endpoints in the study of these drugs.

Next, I will briefly touch on what I think are the clinically pertinent pharmacokinetic and pharmacodynamic characteristics of the drug, and then I will move to an overview of the Phase III clinical program, and I recognize that you have seen a lot of this material already, and so I can be fairly brief there.

Next I will address the most notable safety findings that came out of our review. Now, in that section, I am going to focus primarily on the one year placebo controlled trials, because I think that in general the longer trials and trials that include a

placebo control are the most likely to provide interpretable data in regard to observed adverse events.

I will, however, touch on some of the observations from the remaining studies. Then I am going to move to efficacy findings, and following the same pattern that the applicant chose, I am going to divide my comments into the data which addressed the bronchodilator efficacy, and then the data which address the purported efficacy on the symptom of dyspnea, and then I will round it out with some remaining remarks about additional efficacy variables that were examined.

Finally, I will summarize the most salient aspects of my talk, and then after my talk, there will be time for the panel to ask any questions to clarify any issues that I may have raised.

So as you have heard the applicant has proposed this indication for the drug tiotropium. It would be to treat bronchospasm and dyspnea associated with COPD, and as has been mentioned, no drugs that are currently approved in the U.S. for COPD carry an

indication for the treatment of specific symptoms of COPD, or for the treatment of the disease itself, and then in the next few slides, I will get to what I mean by that.

Before I go on, I do want to comment sort of parenthetically that the drug theophylline is somewhat of an anomaly in this regard. The indications section of the labels for theophylline states that they are indicated for the treatment of symptoms and reversible air flow obstruction associated with chronic asthma, and other chronic lung diseases, e.g., emphysema and chronic bronchitis.

I did want to point that out, but as you know, theophylline is a very old drug, and the contents of the label for theophylline don't reflect the current standards and practices.

So the currently approved drugs for COPD are all bronchodilators, and probably for that reason the indications sections and the labels for these drugs read that they are indicated for the treatment of bronchospasm associated with COPD.

And that language is chosen specifically

to create a distinction between the treatment of bronchospasm in the setting of COPD, versus the treatment of the disease itself. So the bronchodilators have been shown to relax airways in the muscle, and relieve bronchospasm, but they have not been shown to treat the disease.

And what I mean by that is bronchospasm, airway smooth muscle contraction leading to lumenal narrowing, is only one component of the very complex disease of COPD, and while we are very comfortable that these approved drugs do treat the bronchospasm component, they have not been shown to treat other important aspects of the disease, such as mucous production, and such as structural changes in the lungs.

And certainly they have not been shown to effect the natural history of the disease. So therefore we approve these drugs with the indications stating that they relieve bronchospasm in the setting of COPD, and stay away from saying that they are indicated for the treatment of the overall disease.

And in order to establish that efficacy in

regard to bronchospasm, we generally use spirometric measures of bronchospasm, particularly the FEV1, and we are fairly comfortable that the FEV1 can be considered a direct measure of that degree of bronchospasm.

But if you start talking about treating the whole disease, meaning this constellation of physical science and symptoms, the various pathophysiolgic processes, and histopathologic features, then FEV1 quickly becomes more of a surrogate endpoint, and it is a direct endpoint of bronchospasm.

Now, I just mentioned that FEV1 is generally considered a direct measure of bronchospasm. But I want to emphasize the fact that the agency generally would not approve a drug if its sole benefit, its only benefit, were on some physiologic parameter, such as FEV1.

In order for a drug to be approved, there has to be some clinically meaningful benefit to the patient. So implicit in our use of the FEV1 in approving these bronchodilators has always been the

assumption that improvements in FEV1 for a COPD patient do result in something clinically meaningful for the patient.

And I think that is borne out every day in clinical practice, and in particular I would point out that the way that we use data regarding the as needed use of bronchodilators in clinical trials.

So we look at the as needed use of albuterol in clinical trials as some index of efficacy, and we do that because we know what patients know, which is that when their symptoms worsen, they reach for their albuterol, and they reach for their albuterol even though it was approved because of a spirometric improvement, they reach for it because it is going to improve their symptoms.

So what this means taken together is that, first of all, bronchodilators, are bronchodilators only, and they relieve the airways from the muscle contraction, and they don't claim to alter the other pathophysiologic processes in COPD.

And, two, that although we have used FEV1 in the approval process, we have always assumed that

is not the only benefit to the patient, that there is a real clinically meaningful benefit to the patient.

And in that context it is not clear that symptoms can be demonstrated on the basis of a bronchodilator activity, merit or represents unique specific indications for a bronchodilator drug other than what we would normally expect for a bronchodilator.

This slide reviews some of the more common efficacy variables that we see in the study of COPD drugs. It is not meant to be a background. As I mentioned the drugs that we have now for treatment of COPD are bronchodilators, and therefore the primary efficacy end point has usually been some measure of bronchodilation and far and away the most common and most accepted measure of that is the FEV1, because COPD is a chronic disease, and these drugs are intended frequently for maintenance therapy.

And we generally like to see the primary analysis of that end point be performed after chronic use. Now, FEV1 can be examined in different ways or illustrated in different ways. You can look at the

peak FEV1 soon after administration, when the effect reaches its maximum.

Or often we see an area under the curve type analyses of FEV1 time curves, meaning that on a particular test day a patient undergoes serial spirometry at several time points, and the FEV1 is then illustrated along a curve according to the time, and that area under the curve is compared between the drug and its comparator.

Then there are numerous secondary end points which are often used to help support the efficacy of these drugs, and they include other spirometry variables, such as the forced vital capacity.

As I mentioned we look at rescue albuterol use as a measure of efficacy. We are seeking peak flow measurements used more and more in COPD studies, and their primary use has been asthma studies, but they are often included in COPD studies now, and they are usually self-administered twice daily by the patient, and recorded in a diary, and then analyzed in some way.

We are often also seeing some measure of expertise capacity of the patients, and frequently something like the six minute walk test, and as was mentioned, the shuttle walk test was used in some of these trials.

And then you can look at various ways to express the occurrence of COPD exacerbations, and you can look at the number of exacerbations, and you can look at the number of patients with at least one exacerbation, and you can look at the time to the first exacerbation and so forth, and all of those are usually included as secondary end points.

And then we see the inclusion of various so-called patient reported outcomes, including the symptom scales, and the health related quality of life type instruments. Moving to the Phase III program for tiotropium, in all studies the applicant looked at a bronchodilator measure, particularly the FEV1, as the primary, or at least as the co-primary efficacy variable.

And as has been mentioned, the applicant chose to express or to look at the FEV1 rather than at

the peak at the trough, which is a predose measurement. It is a very good idea in drug development programs to include some measure of efficacy at the end of the dosing interval, because that justifies the dosing regimen that is proposed.

If you lose efficacy by the end of the interval, perhaps the drug should be dosed more frequently. And so we often see some measure of end of dosing interval activity as a component of these studies.

It is less common for us to see it as a primary end point, although certainly acceptable. The one potential problem with using the trough variable as the primary efficacy variable is that in general we have a little bit of less consensus regarding what magnitude of efficacy we would expect of a drug at that time point, at the end of the dosing interval.

So as I mentioned, you want to see continued efficacy throughout the dosing interval, but exactly how much, we don't really have a consensus on that. We have a much better feel for what constitutes a clinically significant acute bronchodilator

response.

Often a change of 200 mls, or 12 percent in the FEV1 is applies as a minimal acute bronchodilator response. So it is a little bit hard. When we look at a primary efficacy endpoint, we want to see whether it was statistically significant, and really was it clinically significant, and we have a little less experience assessing what we would require or expect at that trough time point.

Now, as has also been mentioned, after four of the studies had been completed and analyzed, the sponsor examined the data, and realized that they might be able to detect a statistically significant drug effect if they looked at one of the secondary end points, the TDI.

And in particular that in those four studies, the specific TDI analysis was a mean value analysis, and so comparing the mean value in the treated group to the mean value in the placebo group.

But they analyzed the data, and in those exploratory analyses realized that if they defined a threshold of one as a responder, and applied a

responder analysis, they may be able to show a difference between their drug and the comparator.

As you know, responder analysis is where we pre-specify some threshold above which you will call the patient a responder, and below which you will call the patient a non-responder.

So there were two studies that had been completed, but the blind had not been broken, and the sponsor chose to amend the protocols to include both the FEV1 co-primary and a responder analysis of the TDI as co-primary analysis.

And as Dr. Kammerman has emphasized, this decision to elevate when the protocol was written a secondary endpoint to a primary endpoint may be important, because it seems that the protocol paid less attention to the implementation of the TDI than it might have otherwise if it were originally a primary endpoint.

So when you design a protocol and you have a primary end point, the collection of the data that is going to go into that analysis is very carefully guarded, and you want to be very clear and very sure

that the data is collected perfectly, but you may pay less attention when it is one of numerous secondary end points.

Now I am going to spend a few minutes on the PK and PD characteristics of tiotropium. The systemic bioavailability of tiotropium was explored both after oral ingestion and after oral inhalation, and as you can see, after oral ingestion, very little of the drug ends up in the circulation. But after oral inhalation, a more substantial portion ends up in the blood stream.

Now, ideally for a locally active pulmonary inhalation drug, you would want to minimize oral inhalation bioavailability, and that way you can dose the drug at a sufficient level to achieve your efficacy goals without worrying about systemic absorption that could potentially be associated with adverse effects.

Of course, that is not a consideration if the mechanism of efficacy is a systemic delivery. After single dose administration oral inhalation, the drug reaches its maximum blood concentration at five

minutes.

That is often the first test or the first sample that was taken in these studies. So in the first sample at five minutes, that is the Cmax. And it falls away quickly, but it is detectable in the blood for about 2 to 4 hours using the assays that are available.

What is interesting is that the urinary excretion is quite prolonged, meaning that if you administer a single dose of 108 micrograms -- and that is more than the proposed dose of 18 micrograms. But if you administer a single dose of 108 micrograms, you can detect the drug in the urine for 25 days after that single dose.

The last point on this slide is with regard to volume and distribution, and the drug seems to distribute widely wide to the tissues, with a very large volume of distribution of 32 liters per kilogram.

The kidney is very important in the elimination of tiotropium, and 74 percent of the drug is eliminated in the urine as the parent unmetabolized

compound, and initially that happens fairly quickly.

By four hours, 44 percent of the administered dose has been eliminated, but then that subsequently slows down so that by 24 hours, only half of the administered dose has been eliminated.

And when you go up to four days, still only 61 percent of the administered dose has been eliminated. One other observation about the renal handling of this drug is that it has been observed that the renal clearance of the drug exceeds the creatinine clearance, and what that means is that there is some sort of active renal secretion going on, and you are likely using a transporter.

Now, I mentioned that three-quarters of the drug goes out in the urine as the parent compound, and the fate of the remaining 26 percent has not been very well established. It is apparent that it has metabolized either through non-enzymatic hydrolysis and also a component through the liver, using the cytochrome P450 system, specifically CYP 2D6, and to a lesser extent, 3A4.

Using the urinary excretion data, the

terminal elimination half-life of tiotropium was determined to be 5 to 6 days. Now, there is a little discrepancy between the terminal elimination half-life as determined by that urinary data, and the apparent effective half-life.

And by that I mean that if you have a drug whose true effective half-life was 5 to 6 days, and you administered it on a once daily basis, you would expect an accumulation factor of approximately 8 to 9-fold.

The clinical studies with tiotropium instead showed an accumulation factor of 2 to 3-fold, and what that suggests to us is that the true effective half-life may be closer to 24 to 36 hours.

So those are two expressions of half-life; one, the terminal elimination half-life, and one what we are calling the effective half-life. And probably both of those have some clinical significance.

And at least for a systemically active drug, it would be the effective half-life that you would use to help design a rational dosing interval, and less so for a locally acting pulmonary inhalation

drug, whose efficacy may not mirror its pharmacokinetics.

But the terminal elimination half-life may become clinically important, for instance, in the setting of an adverse drug reaction, in a drug where the terminal elimination half-life is quite long, and if a patient suffers an adverse drug reaction, it may take quite a long time for the drug to be eliminated from the body.

The last point is that the pharmacokinetic characteristics that I have described -- and particularly I mean this very large volume of distribution, and the long terminal elimination half-life, suggests to us that what is going on is that the drug is distributed extensively and binds tightly to the tissues in the body, and then is very slowly released back into the circulation.

One pharmacodynamic characteristic, and I am been covering the pharmacokinetics, that I thought was worth mentioning and has been touched on by the sponsor, is worth mentioning because it differs from the other orally inhaled bronchodilators that we have

now.

And that is that the pharmacodynamic effect increases with multiple daily dosing. So we have two sources of data to illustrate this point. One source of data comes from the spirometry data in the Phase III studies, and the other comes from a substudy which was performed in a subset of patients who participated in the year long ipratropium-controlled study, which was performed in Europe.

And in that substudy, 28 patients underwent more extensive spirometry monitoring instead of what was specified for the remainder of the patients, and they underwent six hours serial spirometry, and they underwent it more frequently; at days 1, 2, 3, 8 and 50.

And I will show you the data from these in a second, but the interpretation of this data is that he maximum effect is achieved by day eight, and the sponsor has used the phrase steady state to indicate this maximum effect which is achieved after multiple daily dosing.

So this slide shows the data, the FEV1

data from the two, one year placebo controlled trials. These are the U.S. trials, Studies 114 and 115, and the FEV1 is expressed as the average value over the 3 hour serial spirometry, and as the peak value that was achieved during that 3 hour serial spirometry for each day that it was measured, for tiotropium and for placebo, for each study.

And the message on this slide is that the effects seen on the first day in regard to the average or to the peak is not as large as the effect that was seen after multiple daily dosing. The first time it was checked here was eight hours or eight days.

Now, I did want to point out that at first glance it may look that the pharmacodynamic effect begins to wane after day 50, but I don't think we should over-interpret that observation, particularly in light of the fact that the same type of pattern goes on in the placebo patients.

This slide is the data fro that substudy that I mentioned, and it was called Study 129, and it was a substudy of one of the larger ones, and here the FEV1 data is expressed both as trough, and as peak,

and as average.

The trough on day one is in fact the baseline, and it is before dosing, and the remainder of the values are responses, meaning change from that baseline value. And what this data indicate are that it is not really until day eight that we start to see the maximum effect.

In addition, there is other data from this substudy where they looked at daily morning peak flows, and found that the maximum effect was reached at day six.

Now we will move on to the Phase III program again, and I know that the applicant has already discussed this topic and so I will be fairly brief. These tables show the six pivotal trials grouped according to -- they were replicates or almost replicates. There were some subtle differences between each of these.

The first group, 114 and 115, were performed in the United States, and they lasted a year, and they compared tiotropium to placebo, about 450 to 470 patients in a 3-to-2 randomization, and as

I mentioned the primary end point was trough FEV1, and it was analyzed primarily at 13 weeks.

The second set of studies were European studies, and these studies did not include a placebo control, but rather an active control, ipratropium, which was administered QID. There were fewer patients here, 280 and 247, and they were randomized in a 2-to-1 fashion. The same primary end point analyzed at the same time point.

And the final set of two studies are the six month multi-national studies, in which there were three arms; tiotropium salmeterol, an active comparator, and placebo in a 1-to-1 randomization, and there were approximately 600 patients per study.

Again, as I mentioned, there were two co-primary end points, and they were applied primarily at six months according to the protocol.

And as Dr. Kammerman mentioned, thee were multi-national studies, with a very small fraction of patients coming from the U.S., 5 percent in one study, and about 12-1/2 percent in the other.

You have seen the inclusion and exclusion

criteria, and they are essentially what we see customarily, with a couple of exceptions in the COPD Phrase III trials, there are two things that I want to point out.

One is that baseline bronchodilator responscivity is sometimes measured in studies, COPD studies, and that was not measured and was not a criterion for exclusion or inclusion into the study.

In regards to the exclusion criteria, some patients with certain conditions that I think may be fairly common in the COPD population were excluded from the study. For instance, symptomatic prostate hypertrophy, or bladder outlet obstruction, narrow angle glaucoma, and evidence of some degree of active cardiac disease, such as having had a heart attack in the last year, and having any cardiac arrhythmia which requires drug treatment, or having been hospitalized for heart failure in the last three years.

So I think it will be important to recall these exclusion criteria when we are discussing and analyzing the safety data from these studies.

This table provides the baseline

demographic features of the patients who participated in each of the studies, and again these are the two long U.S. studies, and these are the two year long European studies, and these are the multi-national six month studies.

And what you can see here is that the studies primarily involved men, particularly in Europe, and the patients were all Caucasian. Very few studies or none had a percent Caucasian of less than 90 percent.

The average age of the patients was in the early 60s, and their smoking history ranged from 33 to 34 pack years in Europe, to around 60 pack years in the United States; and the multi-national studies were similar and between, and they had a duration of COPD for about 10 years, and FEV1 was a little lower in the U.S., about a liter, and about 1.22 or 1.23 liters in the European studies, and the FEV1 to FVC ratio was in the low to mid-40s.

So one of the messages from this slide is that there are in fact some differences between the populations studied in Europe and the U.S. in regard

to the pack years of smoking, and the FEV1 impairment.

Now I am going to move on to some of the salient safety findings. As has been mentioned, a total of 13 hundred patients were exposed to tiotropium in Phase III, and the safety evaluations that were performed were what we commonly see for these studies; adverse events, vital signs, examination, labs, and ECGs.

One comment about the ECGs is that normally the way that we like to see the ECGs is that you check the ECG after the first dose to look for an acute effect, and periodically after chronic dosing to look for acute and chronic effects.

And you specify in the protocol that the ECGs be performed at or near the time of the Cmax of the drugs, and so you want to know the maximum concentration in the blood, and check the ECG around that time.

Very rarely the cardiac pharmacodynamics of a drug differ from the pharmacokinetics of the drug, and if you know that, you time your ECGs to the cardiac pharmacodynamics. But for the most part, we

ask that the ECGs be performed at the Cmax.

And that was not the case in these studies. The ECGs -- the protocols did not specify when the ECGs would be performed, and so they could be performed at the individual center before or after, or so many hours after the dosing.

We don't know, and that was not specified, and we couldn't find that information. The other point about the Phase III studies is that none of the Phase III studies included Holter monitoring, and that was done in Phase II as I will talk about in a moment.

Now, I just mentioned a couple of relative deficiencies in the Phase III safety data. I will say that in Phase II they did have some timed ECGs, and that was in a multiple dose-ranging study, which examined doses up to 44 micrograms.

So that the dose is higher than what are proposed for clinical use. These were 29 day studies, and so we have only chronic exposure up to 29 days in regard to the timed ECGs, and the ECGs as has been mentioned were performed at 1, 3, and 5 hours.

So the first ECG was beyond the time of

the Cmax. A separate study in Phase II did include Holter monitors in 72 patients before and on treatment, and I will speak to that in a few moments.

Now, as I mentioned, when I discussed the safety database findings, I am focusing primarily on the one year placebo controlled trials, primarily because the longer duration, one year as opposed to six months, and the presence of a placebo control helps us to more rationally attribute adverse events as a drug effect.

Now, one other introductory comment is that sometimes when you are looking at placebo controlled trials, the occurrence of adverse events can be affected by the duration of exposure.

So if in a placebo-controlled trial more placebo patients are dropping out of the study, perhaps due to lack of effect, then the occurrences of certain adverse events may look lower than placebo simply on the basis of the duration of exposure.

I say that to say that I don't think that potential bias as a compounding factor is operative here because the median exposure was similar in the

two groups. The category of adverse events that were most common were gastrointestinal and as has been mentioned the frequency of dry mouth far exceeded that in the placebo group.

And in this slide, and in my subsequent slides, I will follow the convention of providing the data for the tiotropium, and then followed by the comparators. So this is the list of gastrointestinal -- specific gastrointestinal adverse events that were seen more frequently.

I will point out that constipation in particular because I am going to address that in a subsequent slide as well.

In these year long studies, it was not uncommon for patients to develop upper respiratory tract infection. However, the occurrence of upper respiratory infection in the tiotropium group was greater than that in the placebo group, and we will see that in other studies.

And these are the remaining respiratory system adverse events that occurred more frequently in the tiotropium group. They may or may not reflect the

effects of drying on the mucous membranes of the upper airway.

So we saw chest pain more frequently and rash more frequently, and finally urinary track infection, and I want to point that out specifically because again I will have further slides that will address urinary tract infection, and also because there is at least a plausible mechanism by which tiotropium could increase the risk of urinary track infection.

And by that I mean if there is a systemic anticholnergic effect, it could result in some degree of urinary status and put the patient at increased risk of urinary tract infection.

This slide addresses the six month studies, and what we saw in the six month studies is that there were actually fewer differences between tiotropium and placebo. These were the adverse events which were more common in the tiotropium group, as compared with placebo, and what I have done is in yellow text indicate the adverse events signals that we saw in the year long placebo controlled trials.

So in the year long trials, we saw dry mouth and we see it again here in the six month trials, and in the year long trials we saw upper respiratory tract infection, and we see it again here; pharyngitis and sinusitis.

One side comment is that the overall occurrence of --- you may notice the overall occurrence of adverse events is lower in the six month studies than they were in the one year studies likely just related to the duration of exposure.

Now, I should mention that there were some data shown this morning by Dr. Kesten in which all of the placebo controlled data was pooled, and that is data that we have not seen before, and so I can't really comment on it.

I would comment that p-values were included in the slides, and I don't think that applying p-values to this type of data is relevant. The other is that the data were presented in patient years, according to patient years exposure, and there are certain assumptions that go into that type of explanation of the data.

It assumes that the risk of that adverse event is constant over time, and I am not sure that that can be assumed. So I will say that I can't really comment further again because I have not seen that type of analysis before today.

Now, for all new drug applications, we asked that the sponsor examine both the safety and the efficacy data for any evidence of interaction with certain demographic features. And so what this slide shows is the safety interactions that were discovered in the one year placebo controlled trials.

And we saw safety interactions in regard to age and gender. We were really not able to perform interaction studies based on race because there were so few non-caucasians.

So in order to assess for an age interaction for these adverse events the populations were divided into patients who were less than 60, patients who were between 61 and 70 years of age, and patients who were more than 71 years of age, or 71.

And there were three adverse events that showed an interaction; dry mouth, constipation, and

urinary tract infection. So in the youngest group of patients the occurrence of dry mouth was 11 percent, but it increased as the patients got older, and the occurrence was 21 percent in the oldest patients.

Likewise, for constipation, it was two percent in the youngest, and rose to six percent in the oldest patients. And urinary tract infection rose from 3.3 percent in the youngest to 12 percent of the patients in the oldest group.

And we didn't see that type of interaction at all for the dry mouth or for constipation. There was some evidence of a age interaction for a urinary tract infection, likely meaning that in this population of patients, as you get older that you are at an increased risk for developing a urinary tract infection, but it appeared to us that the interaction was stronger in the patients on drugs, suggesting a true drug effect.

And in regards to gender, what we saw is that women develops dry mouth much more frequently than men, and that is not something that was seen in the placebo group. A few other safety observations.

Regarding urinary retention, there were four patients in these one year placebo controlled studies who developed significant urinary retention and all of those four patients were treated with tiotropium.

And what I mean is that all of these patients required a full catheter, and in fact three were subsequently started on medication for BPH, benign prostatic hypertrophy, following the event.

Keep in mind that patients with symptoms of benign prostatic hypertrophy, or bladder outlet obstruction, were in fact excluded from participating in these trials. Nonetheless, four patients developed obstruction requiring a full catheter.

Then finally under a micturition disorder or micturition frequency, the observation is that there was a greater frequency of patients in the tiotropium group, as opposed to placebo patients developing adverse events characterized by either of those two terms.

In regard to constipation, one other observation I mentioned was the age interaction, and

the other observation is that in fact there was one patient who was treated with tiotropium, who in fact was hospitalized with a fecal impaction.

The last observation here is of uncertain significance, because we don't at this time have any mechanism to explain it. But the observation from the data, and these are the one year placebo controlled data, is that the adverse events characterized as diabetes or aggravated diabetes, or hyperglycemia, were more frequent and occurred in 14 or 2-1/2 percent of the tiotropium patients, versus one or .3 percent of placebo patients.

And as has been mentioned, we pay particular attention to potential cardiovascular effects, both because of the mechanism of the action of the drug, and because of the patient population which I will go into, and we know very well that cardiovascular disease is quite common as a concomitant disease in the COPD population.

And what we observed is that under cardiovascular effects, in the category of heart rate and rhythm disorders, there seem to be a possible

signal of drug effects, meaning that adverse events in this category were more frequent in tiotropium, as compared to placebo, and serious adverse events.

So these are adverse events that reached the threshold for being declared serious, and were also more frequent in the tiotropium patients. I will point out that this signal was not seen in the ipratropium controlled studies, and we have no data from that to suggest an effect.

And as has been mentioned, we did not detect a safety signal on the ECGs that we have available given their limitations.

In regards to death, the first and most important observation is that the incidence of death was similar in all groups. However, there is one observation that may be important, and probably is worth pointing out. In the placebo controlled one year studies, 5 of the 7 deaths that occurred in the tiotropium group were attributable to cardiac ischemia, or arrhythmia.

And that compares with one out of the seven deaths that occurred in the placebo groups. In

the ipratropium controlled trials, there were -- the deaths due to MI were three of the nine tiotropium deaths, and none of the three ipratropium tests.

I mentioned that there was Phase II data to support the cardiovascular profile, and we did not see any safety signal on the Holter monitors, which were performed in 72 patients before and on treatment.

There was one subject who developed a four-fold increase in ventricular ectopy on tiotropium, but that needs to be taken into context, because a number of other subjects actually exhibited decreased ventricular ectopy.

I will point out that a number of patients exposed or that underwent Holter monitoring is somewhat low. If you look at the label for Serevent, they describe 284 patients who underwent five, 24 hour Holters. These are COPD patients.

And although I have emphasized the placebo controlled trials, because it is much easier to attribute a drug effect, you may be interested in seeing how the adverse event profile compares in the ipratropium controlled trials.

So these were the European trials, and we don't have a placebo arm for a comparison. What these represent are adverse events that were more common with tiotropium than with ipratropium, and they are only included on the table if they were also more common in tiotropium than in placebo in the year long placebo trials.

So we saw chest pain in the placebo controlled trials, and we see it here again, and again we saw dry mouth, and we see it here again.

Perhaps worth noting is that the degree of dry mouth seems to be, or the occurrence is more frequent certainly than in the ipratropium. And there are some others here that might relate again to the drying effects in the airway that are not clear.

Again in the placebo controlled trilas we saw upper-respiratory tract infections more frequently, and here upper-respiratory tract infections occurred in 43 percent in the tiotropium group, compared with 34.6 percent in the ipratropium group. And finally again we see urinary tract infection, 3.9 versus 2.2.

Now I will move on to the efficacy data. Again, I have divided the efficacy data into the bronchodilator efficacy, the dyspena, and miscellaneous others. So this slide shows the results from the U.S. studies, the year long, one year U.S. studies, 114 and 115, and as has been mentioned, the primary end point in these studies was the trough FEV1 response at week 13.

And the table shows that tiotropium was statistically significantly superior to placebo in both trials, with a treatment effect size of about 140cc's generated by an improvement in the tiotropium group and a slight decline in the placebo group.

And if you look at the same variable, the trough FEV1 at the other clinic visits, tiotropium was also statistically superior to placebo at all of the other visits, and the effect sizes at this point were 110cc's to 160cc's.

Now, that is the trough, and I mention the distinction between the trough, looking at the trough FEV1, versus some measure of peak, and here tiotropium was also statistically superior to placebo on the peak

FEV1, and on the average FEV1 during those three hour serial spirometries performed at each clinic visit.

The FEV1 data may be worth a little closer look. The mean peak FEV1 response at day one was about 240cc's, and on subsequent clinical visits, as I mentioned, it increased to about 250 to 310cc's.

Now, although the mean peak at day one was 240, this should say the mean FEV1 response at each individual time point on day one. So, at a half-an-hour, two hours, three hours.

You look at each one of those, the mean response was always less than 200cc's, and we want to investigate why there was an apparent discrepancy, and the reason is that the individual patients reached their peak at different times during that spiral spirometry.

So that at any particular time, about a third or less of the patients were actually reaching their peak, and the reason that I point that out is that that could potentially have some impact on how we describe the onset of action of the drug.

To round out these year long studies,

tiotropium was also statistically superior to placebo on the forced vital capacity response, whether it was looked at the trough, average, or peak, and also for the peak flow measurements, and again those were home measurements, and the mean over each week was examined.

And that tiotropium was superior for most weeks, with effect sizes that ranged from eight early in the course of the study to around 31 liters in the morning, and 13 to 40 liters in the evening, liters per minute.

These are the European ipratropium controlled trials, and again the same primary efficacy end point was used. I should make a note regarding this primary efficacy variable. We know based upon the pharmacodynamics of ipratropium.

That at the trough value after a previous evening dosing and then coming into the clinic and measuring trough values, you are unlikely to detect an effect of ipratropium based simply on its known pharmcodynamics.

So it would not be surprising that the

tiotropium would show a similar effect size against ipratropium as it did against placebo. And having said that, tiotropium was superior to ipratropium on this variable in all clinic visits, and the effect size again were around 110 to 180cc's.

This slide shows the data from the six month multi-national studies, and focusing on the co-primary end point, which was again the trough FEV1 response. And again this slide shows that tiotropium was statistically superior to placebo in both studies at week 24, with a treatment effect size of about 110 to 140cc's, again because of an improvement in the tiotropium group, and a slight decline in the placebo group.

Again, looking at the trough FEV1, the same variable. At all other clinic visits, tiotropium was statistically superior, and the effect sizes were similar, 110 to 150cc's.

Again, tiotropium was statistically superior to placebo on the peak FEV1, and the average FEV1, during what was either a 3 hour or a 12 hour serial spirometry, depending on the study.

And then finally as seen in the other studies tiotropium was superior to placebo in regard to forced vital capacity looked at in several ways, and in regard to the peak flow.

Now I will move on to discuss the dyspena findings, and Dr. Kammerman has already reviewed some of the issues concerning the instrument itself, the instrument that was used to establish efficacy in regard to dyspena, both in the instrument and how it was validated and developed, and so forth, and how it was implemented in these particular studies.

So I am not going to go into that further, but instead will just present the data. This is the data from the six month studies that were used primarily to support the dyspena claim.

And this is the responder analysis, again defining a responder as a TDI score greater than or equal to one, applied at six months, and what we see from this table that tiotropium was statistically superior to placebo in regard to the percentage of patients who showed any improvement on the TDI.

I phase it specifically in that way to

emphasize the fact that because of the instrument, and because of the way the applicant defined a minimal clinically important difference, there was no degree of improvement that a patient could indicate that would not be considered to be a clinically meaningful response.

And that is again built into the instrument, and then in how it was applied using the minimally important difference of one. So you could score zero, but if you wanted there to be any positive improvement, that is a clinically meaningful response.

Two other points that I wanted to make on this slide. One is regarding the actual effect size that was shown. It is relatively small or modest. In one study, 16 percent more of patients who were treated with tiotropium achieved this TDI responder; and in the other study, 12 percent more of the patients received their responder.

So by giving tiotropium rather than placebo to these patients, you achieved 16 percent more of them that became responders based on the definition, and here 12 percent more. And the other

point that I wanted to make from this slide is that salmeterol, the active comparator, and again a bronchodilator approved on the basis of FEV1, and a drug that does not have an indication for dyspena, also faired fairly well on this end point.

In Study 130, the difference between placebo and salmeterol was not statistically significant. In Study 137, it was, and in fact in Study 137 the percentage of patients who were responders was numerically greater than that with the tiotropium, and that is reflected in the p-values here, where superiority over placebo met a p-value of .01, and here the placebo value was .05.

One other comment about the analysis of the TDI at 6 months, is that the datasets used for those six month analyses necessarily included fewer patients that were randomized to treatment. So this slide shows the numbers of patients who were randomized, versus the number of patients who could be included in that statistical analysis.

And there really was no way of avoiding it for a few reasons. One is that in the statistical

analysis of the TDI, one of the co-variants in the statistical plan was the BDI score. It had to be included.

So if a patient for some reason did not have a BDI score, they couldn't be used in the TDI analysis. And likewise if the BDI score was scored in a way that said amount uncertain or unknown, or short of breath for -- or limited for reasons other than shortness of breath, they could not be included in an analysis.

And the other reason why there is sort of a fall off in the number of patients is that the first time the TDI was administered was at week eight, and so that any patient who dropped out before week eight had no TDI data that could be carried forward in a statistical analysis.

So the numbers aren't that dramatic, although in this placebo group about 25 percent of the randomized patients couldn't be included in the analysis. And I just would point that out because at some point in some studies, when the number of patients who can be included in the statistical

analysis falls to some degree, it impacts the ability to arise at firm conclusions based upon those statistical analyses.

Again, there is no way of avoiding it. That is how it had to happen. But at some point when the numbers get too low, you start to wonder what you are really learning from the data. And then finally regarding the primary analysis, or primary efficacy variable or co-primary, is this slide that looks at a number needed to treat analysis.

It is a different way of understanding the treatment effect size with this drug, and according to the number needed to treat analysis, either in the individual studies or the combined data, you would have to treat approximately eight patients with tiotropium to achieve one patient over than what would be expected with the placebo, who was a responder based on this definition.

Now, of course, the TDI was administered on days or on visits other than six months. It was administered at 8 and 16 weeks, and this slide goes over the data from those studies, and the message is

very similar.

Again, in each of the studies, at both 8 and 16 weeks, the percentage of responders based upon the value of one, was superior, statistically superior in tiotropium, as compared to placebo, and the same pattern was seen with salmeterol, where statistical superiority was not achieved in Study 130, but was achieved in Study 1137, with a low p-value, and in both the 8 weeks and in the 16 weeks, again the percentage of responders was greater in the salmeterol group than it was in the tiotropium group.

And then as has been mentioned, you can also look at TDI as mean values, comparing the mean value in the treated, versus the mean value in the comparator, and in fact as I mentioned, that was the specified analysis for the four year long studies.

And this slide shows for each study -- and remember that these four are placebo controlled, and these four are actually active controlled with ipratropium. In this column, you see the visits at which the TDI mean score was statistically superior in the tiotropium group.

And in this column, you see the weeks at which that difference between treatment groups exceeded one, and again I am using one as the sponsor's proposed definition of what would be a minimally clinically important difference.

And what you see is that it is very frequent to achieve statistical significance from placebo, but less frequent to achieve a difference that exceeds one. Now, the next few slides provide some additional data that reflect on the efficacy of the drug in regard to the symptom of dyspnea.

We have talked then about he primary efficacy variables, and let's look at some of the secondaries. Studies 130 and 137, these are the same studies that the TDI was used as a co-primary, including this post-dose shuttle walk test.

So that was administered on day one, post-dose, and at weeks 8, 16, and 25, the same intervals at which the TDI was administered. The shuttle walk test is a standardized test in which patients are told to walk back and forth at a steady pace on a 10 meter course until they are unable to maintain their

required speed without becoming unduly breathless.

So this is the distance that they are able to walk and which is limited by their breathlessness or dyspena. In conjunction with the shuttle walk test the Borg Dyspnea Scale was applied both before and after each shuttle walk test.

Many of you are familiar with the Borg scale. It ranges from zero, which means nothing at all, to 10, which means maximal. It is a little bit unusual in that when you get to five on this scale, you are already at severe dyspena, and scores from 6 to 9 reflect very severe dyspena, and then very, very severe dyspena, until you get to maximum.

So the data from those examinations are that in regard to the walking distance, the distance that patients were able to walk without becoming unduly breathless, there was actually no difference between groups in either of the studies.

In fact, in one of the studies the placebo group was numerically, although I emphasize not statistically, but numerically superior to tiotropium in one study. And the walking distance did not

increase during the study in any of the groups.

So I think that this may impact your deliberations about the strength of the dyspena signal. In regard to the Borg scale, with one exception, there was no difference between tiotropium and placebo on that scale.

The only exception was week eight, when a statistical difference was noted both pre-and-post exercise, and the value on this zero to 10 scales was -- the effect size was 0.24 and 0.32, again on a zero to 10 scales.

The one other way to address dyspena would be this so-called COPD symptoms score. I think the applicant showed you some of the data. That was applied in several of the studies and the COPD symptom score is the investigator's assessment of the patient, and their status over the prior week in regard to several COPD symptoms.

And the investigator scored them on a four point scale, zero to three. And the results showed that tiotropium was statistically superior to placebo if you looked at the component shortness of breath.

If you just pulled that out and looked at shortness of breath, it was statistically superior at most visits.

The effect size on the four point scale was 0.13 to 0.36, and I put it in here, but I'm really not sure how to interpret this data, because we don't know how well validated it is, and I suspect that it has not been validated, this symptom score.

Nor do we know if it is reasonable to pull out a component of the symptom score and look at it. Nor do we know how to interpret this effect size in regards to its clinical meaningfulness.

The next few slides will consider a few additional secondary end points going by the groups of studies. These again are the one year U.S. studies, and what was shown here in these studies in regard to the remaining efficacy variables was that tiotropium was statistically superior to placebo in regard to this physician's global evaluation.

Again, we don't have much information on its validation, nor do we know how to interpret an effect side of 0.25 to 0.59 on a 1 to 8 scale. Tiotropium in these studies was also superior to

placebo in regard to the as needed use of albuterol, with subjects required 5 to 6 fewer doses of albuterol per week in the year long placebo controlled trials.

We did not see any consistent meaningful difference in these studies in looking at COPD exacerbations, or COPD hospitalizations. We did not see a consistent meaningful difference shown in the St. George's Hospital Respiratory Questionnaire, or in the SF-36.

In regards to the European ipratropium controlled studies, we did not see an effect on the as-needed albuterol use, or on COPD exacerbations on hospitalizations. In the six month multi-national studies, tiotropium was again shown to be superior to placebo on this physician's global evaluation on all test days, except one, with effect sizes shown on a scale of 1 to 8.

Again, it is hard to know how to interpret that. We didn't see any consistent meaningful difference shown in as needed albuterol use surprisingly. There was statistical superiority in one of the studies, but in the other study,

statistical superiority was not obtained.

Nor did we see a consistent effect on COPD exacerbations or hospitalizations, or the SGRQ, or a patient satisfaction questionnaire. So to summarize, the pharmacokinetic features of tiotropium are somewhat unique among inhaled bronchodilators, particularly the very large volume of distribution.

And a very long terminal elimination half-life, and the apparent tight tissue binding with slow release back into the circulation. On the safety side, dry mouth is common, and we saw both an age and an gender interaction, and we observed in the year long ipratropium trials that in fact the occurrence of dry mouth is more frequent with tiotropium than with ipratropium.

There were several adverse events that occurred more frequently with tiotropium than with placebo, and they may be reflections -- some of them may be reflections of the drying of the airways, and some could reflect a systemic anticholinergic effect.

And then again we observed a possible effect in regard to heart rate and rhythm, which may

merit some further evaluation.

In regard to efficacy, tiotropium appears to provide clinically meaningful bronchodilation, and its duration of action seems to support once daily dosing. The maximum bronchodilator effect isn't reached until after multiple daily doses.

And there is a demonstrable, at least statistically demonstrable, effect on the TDI. However, the clinical significance of this effect is not known. First of all, as Dr. Kammerman went into extensively, there are issues with the instrument, and its implementation in these studies.

And then other issues about how to interpret the effect side and the minimally important clinical difference and so forth. One other point which I wanted to include is that the package didn't address either the safety or the efficacy of concurrent as needed ipratropium, which may occur in the clinical setting.

So with that, I will conclude my remarks, and invite any clarifications that you may need.

CHAIRMAN DYKEWICZ: Thank you.

DR. SULLIVAN: Mark, I just wanted -- I'm sorry, but I wanted to point out that Dr. Kammerman is going to have to leave, and if there are biostatistical questions that may be directed to Dr. Kammerman, it is better to do those early. Thank you.

CHAIRMAN DYKEWICZ: All right. We are open to questions from the committee about the FDA presentation. Dr. Chinchilli.

DR. CHINCHILLI: Yes. When Dr. Kammerman said the sponsor was blinded when they decided that they wanted to make TDI a primary outcome in the two shorter term studies, the 6 month studies, does that mean that hey were blinded to the data, or does that mean that they could see the data, but were blinded to the treatment identity? So I was not clear.

DR. SULLIVAN: I think it may be best to have the applicant address exactly what was known at the time.

DR. MENJOGE: This is Shailendra Menjoge, the biostatistian on the project. We had the data in-house; however, we did not know any treatment codes.

DR. CHINCHILLI: That's what I mean. So

you saw the data, and you saw there were differences in groups. You just did not know which group was which?

DR. MENJOGE: No, we didn't. There was no way to find any differences or anything. Basically, the data was collected and it was brought in-house, some of the data, but there was absolutely no knowledge of any treatment at all.

There were no analyses done or anything like that either.

DR. CHINCHILLI: Oh, okay.

CHAIRMAN DYKEWICZ: Dr. Swenson.

DR. SWENSON: Yes. A question for Dr. Sullivan. You presented your interpretation of this COPD exacerbation rate somewhat differently from what we heard from the company Can you share or address that issue, because they came out with an indication or a suggestion that they decreased the rate of exacerbation, and you told us otherwise.

DR. SULLIVAN: Sure. Right. It is our practice to look at individual studies alone, and in the analyses that the applicant provided, there were -

- the studies were grouped together, and so they met analysis if you will. So what I have said is that we did not see a consistent finding.

In other words, a statistical significance was not shown in either study. If you group a bunch of the studies together, I believe that is where the data from the applicant came from.

CHAIRMAN DYKEWICZ: Dr. Patrick.

DR. PATRICK: Dr. Kammerman, you mentioned that there was -- that you observed a fair amount of overlap or co-linearity between the three components of the TDI. Is it possible that that co-linearity would then drive the responders' analysis?

DR. KAMMERMAN: Well, I am not sure that it would actually drive the analysis. If somebody had a positive response in one of the components, they were likely to also have positive responses in the other two components. There were very few instances where the positive on one of them would overcome a couple of negatives on the other two.

Where it could make a difference those is that if you started changing the clinically meaningful

difference thresholds, and let's say from 1 percent to 2 percent, or 3 percent -- I'm sorry, the unit of change went to a three, then if they are related, the change of two many not really mean that much more than a chance of one.

DR. PATRICK: Just one real quick follow-up. Wouldn't all f this depend on where you started? So if you had dyspnea at rest, a one unit change to eliminating dyspena when you could dress might be very much different than going from walking on level ground to walking on a hill?

DR. KAMMERMAN: Yes.

CHAIRMAN DYKEWICZ: Dr. Sullivan.

DR. SULLIVAN: I just wanted to comment further on your question about the exacerbations. Some of the difference between the presentations may reflect the fact that I believe the data presented by the applicant had to do with time to first exacerbation, and the analyses that I looked at were the numbers of exacerbations.

So some of the differences may be explained in that way.

CHAIRMAN DYKEWICZ: Dr. Joad.

DR. JOAD: Are there any other implications of dry mouth besides your concern about sinusitis, like dental problems, for instance?

DR. SULLIVAN: We didn't see that. I think that one of the considerations about the frequent occurrence of dry mouth has to do with the blinding of the study as well. But as far as more serious adverse events related to drying of the oral mucosa, I don't want to raise that.

CHAIRMAN DYKEWICZ: Dr. Parsons.

DR. PARSONS: Just following up on that. Since the incidence of dry mouth when you adjusted for gender appeared to be a lot greater in women and there weren't that many women in the larger trials, are there any issues about other effects in women, in terms of cardiac effects ultimately?

Were there enough women studied? I just worry when one variable appears to be significantly increased? Is there reason to suspect that there might be more problems?

DR. SULLIVAN: I think that is a very

reasonable question. I can say that I didn't see any gender difference in regard to the cardiac effects. Again, it is hampered by the fact that a few women were exposed, particularly in the European studies.

In the United States studies, it was a little bit more balanced.

CHAIRMAN DYKEWICZ: Dr. Swenson.

DR. SWENSON: To the question of the dry mouth, and maybe this is a question to someone in the company. Do you really consider that a systemic effect or is that possibly a combined systemic and local effect?

DR. DISSE: I would like to address this from the systemic absorption by inhalation, which I think every health drug has, and from the pattern of onset, we believe that it is a systemic effect.

And also from animal pharmacology, you can follow up that the dryness of salivary secretion is always the most sensitive anticholinergic signal which appears first.

CHAIRMAN DYKEWICZ: Dr. Apter.

DR. APTER: I am concerned about the

demographic distribution of the population tested. Dr. Kammerman mentioned in the questionnaire that there might be cultural differences, and language differences, but also adverse effect differences.

And we know from the experience of ACE inhibitors, for example, that African-Americans are much more likely to experience angio-edema than Caucasians. So I am concerned when the study was set up and negotiated between the FDA and the company that there were not more measures instituted to ensure that there would be a broad range of minorities, such as was seen in this country -- African-Americans, latinos -- and you mentioned maybe there is some data about Asians. I don't know.

And the other issue, too, is that minorities have poor health across all diseases than Caucasians. So they would be -- and I don't know of Dr. Menjoge's demographics, but these patients would be more likely to be exposed to these medications.

DR. SULLIVAN: I think we are certainly sensitive to representing all populations in these clinical studies. I can't speak to the discussions

that went on now several years ago before these pivotal studies were being planned.

I know that it is our current practice to advise responders in Phase II to be sure to include adequate representation. I should say that in that CDCMMWR report, it is apparent that the occurrence of COPD is more frequent in whites than in African-Americans.

So to some extent the disparity is explained by the burden of disease, but I don't think it is entirely explained.

CHAIRMAN DYKEWICZ: Dr. Chowdhury, did you want to make a comment?

DR CHOWDHURY: I just wanted to make the same point here, that when a study is planned and conducted, typically one would make an attempt to have adequate representation of both the genders, and the way they show racial distributions, and that is what is expected.

However, the fact is that the with the data that you have, that is the data that you have, and I would ask you to comment on the overall data,

and that may be one of the considerations that you want to recommend making to us.

CHAIRMAN DYKEWICZ: Dr. Parsons.

DR. PARSONS: One thing that I couldn't clearly determine from the literature provided to us is what other medications could these patients be on when they enrolled in the trial?

It wasn't clear to me that there were specific exclusion criteria for inhaled steroids, for example, which granted may not be approved, but certainly that a lot of patients are on.

And my question is could patients be on alternate medications, and if so, was the frequency of distribution the same between placebo and the trial participants?

DR. DISSE: So as you can see here, this was a baseline pulmonary medication on entering into the trials, and many patients were on inhaled anticholinergics. Of course, these had to be withdrawn. Beta-agonists were inhaled almost entirely in everybody, and these of course could be continued on demand.

All the beta-agonists had to be withdrawn, and inhaled steroids could continue on a steady level, and oral steroids could continue, and theophylline oral could be continued, except in one set of our replicate trials, and a few patients on oxygen.

DR. PARSONS: There is a bit of a difference in inhaled steroid use. Is that statistically significant?

DR. DISSE: No, it is not statistically significant. There is some variability also with our studies, and so this is the studies conducted in the United States, and European studies in proportion on steroids was at about 70 percent. So a lot higher.

DR. PARSONS: Is there any association between the concomitant use of inhaled steroids and the change in TDI scores? And were the percent of responders more likely to be on inhaled steroids?

DR. DISSE: We can show the subgroup analysis for FEV1, as well as for TDI, and we have not seen an interaction here. So tiotropium was effective no matter there is co-administration of inhaled steroids or not.

CHAIRMAN DYKEWICZ: Dr. Morris, did you have a question?

DR. MORRIS: Yes. This is a question for Dr. Sullivan, and possibly for Dr. Kesten. In regards to the cardiac Aes, and the data that was presented, is there any clustering of the AEs, cardiac-wise, on drug versus placebo in regards to time on drug?

DR. SULLIVAN: In our dataset, we weren't able to -- the dataset that we had available, we weren't able to look for that type of a pattern. Perhaps Dr. Kesten has looked at it.

DR. KESTEN: We did look for that, and there was no clustering in this specific time frame from cardiac AEs.

CHAIRWAN DYKEWICZ: Dr. Stoller.

DR. STOLLER: My question regards the agency's level of confidence in the minimally important clinical difference. And in particular one of the issues brought up in the applicant's briefing document regards -- I mean, one of the other ways to examine this is recognizing that there weren't a priori definitions of minimally important clinical

difference as, for example, has been shown in some of the other available indices, CRQ and SGRQ.

One on the supportive arguments appears in the table on page 38, Table 3.2:3, which dichotomizes the TDI transitions, and then looks at in those dichotomizes greater than or less than one TDI differences in the SGRQ, for example.

And my question perhaps is really to Dr. Sullivan or Dr. Kammerman, and it may be an invitation to the applicant, is did you have an opportunity to look at the actual scatter of those data.

The actual distribution of the SGRQ -- do you see where I am? Whether you had an opportunity to look at the actual distribution of those data as a way of either strengthening the idea that there is a relationship with other a priori defined minimally important clinical differences, and if you didn't, whether there is an opportunity to look at the distribution of those data with regard to whether this is outlined based or not.

You know, you have raised several concerns, some of which I share with regard to the

actual administration of the instrument, which is perhaps separate from this issue. The other issue is how confident are we in the minimally important clinical difference.

DR. KAMMERMAN: Well, I have not looked at this table in a while, but when I first look at it, my impression is if we want to use this as evidence to support a clinically meaningful difference in an evaluative instrument, then patients would need to be classified in a different way based on, for example, to three groups of patients whose clinical status remains stable over time, and improved over time, or decreased over time.

And then look at the responders in that fashion, and as for this, there is still the problem that the SGRQ was administered right before the TDI, and so we will see all patients with improved breathlessness had a mean score of minus 6, and all patients with no change or worsening of breathlessness, had a mean change of .74.

And there is still the issue here of the bias, but moreover, personally speaking, I am not so

confident about one unit as being clinically meaningful. There is the degree that Dr. Patrick raised about the overlap among the three items.

The scoring of each item was not consistent, and they were just simply added together, and I still haven't seen really good evidence to support the one unit change as being meaningful.

CHAIRMAN DYKEWICZ: Does the Sponsor wish to respond?

` DR. WITEK: I would first like to answer that point, and then if you would allow us to just address the issues that have come up in some of the biases. If you can just please pull up Slide 2763.

To your point, Dr. Stoller, about -- and you can display that, but the analysis that was done by the agency, I think we respect that, but we would like to point out some of the issues where we don't believe that these biases have manifested, such as the dry mouth.

But if we look at a more objective measure to your point about just dichotomizing and if you

responded or not. Here you are just looking at taking the entire cohort from the one year study. If you responded on the TDI, you do see less albuterol than when you didn't.

And let me just show this as another way to correlate that these measures are associated. So even with the point that the SGRQ was looked at right before, this is a little bit more an objective measure. You can take the slide off.

DR. KAMMERMAN: I have one question. Could you please address the issue of missing data, and how that affected the TDI in the one year studies?

DR. WITEK: Could we just address them in the order and we will make sure that we come back to that.

DR. KAMMERMAN: Okay.

DR. WITEK: Then just let me address one point of the biases, and then I would like to have Dr. Jones to comment on some of his general experiences. If we can just put up slide 2748, please.

It was discussed about correlation coefficients being low and explaining very little

variability. If we can just display the slide. These are data from the multi-national studies, and I have to acknowledge that this data were not presented to you.

But this is just looking at the associations between the BDI and the change in SGRQ, and the dyspnea score, and the global evaluation, and the FEV1 that was mentioned.

And one of the ways that we look at the question of the multi-national biases is that we have made a dichotomy of these correlations, for example, in the countries, and in the multi-national countries that were native-English speaking, and then non-native English speaking.

Perhaps an indirect, but one of the ways that we can look at that, and we see that the correlations here at the BDI, whether you looked at native-English or non-native Engish speaking countries, are similar.

And the degree of these correlations are exactly what one would expect, and less of a correlation between an objective measure, and a

stronger correlation here as you see between two dyspena measures.

And although they are different, we wouldn't expect a very high correlation on things that would be measuring different things. For completeness, if we can just go to the next slide.

CHAIRMAN DYKEWICZ: Let me just interject. You can present that, but really because this data has not been presented to the FDA, or the committee before this point, it should not be considered in our deliberations.

DR. WITEK: Okay. So this is just showing for the deltas and we see the same thing. I would just like with respect to the manifestation of some of the potential biases, that Dr. Jones could discuss that and his experience with his instruments specifically to ours.

And then we will come back to the question on the missing data.

DR. JONES: Dr. Kammerman makes a very important point, that bias in clinical trials tends to with the unblinding observer of patient, tends to lead

to an over-estimation of the treatment effect.

One of the interesting things is that in other studies using our instrument, we found that other agents are associated with a higher number of side-effects have been associated with a decrement in the observed treatment effect.

That has been with entirely different data, and long acting (inaudible), but that is a phenomenon that has been observed. And in these particular studies, we found that the TGI response in those who were reporting a dry mouth was lower than in those who didn't.

Now, that is important because that is one of the indications whereby a patient or the clinician may have judged the treatment that the patient was receiving, the active treatment, because dry mouth is a symptom of the active treatment.

If you could show me Slide 3, I think. No, the next. Thank you. This one. These are data from the 6-month tio, albuterol and placebo studies, and looking at the percentage of responders with the presence or absence of dry mouth. So the y-axis is

the percent of responders, and the pink is the patient with no dry mouth, and blue is the patient with dry mouth.

And we can see for each of the treatment groups the patient who had no dry mouth had a higher response rate in the TDI than people who had the dry mouth. So I think that is one concern that can be settled in the specific context of this study. And if I could go back two slides, please.

DR. KAMMERMAN: Could you put up that slide again, please? I am just trying to absorb it.

DR. JONES: Shall I talk it through again or would you like to look at it?

CHAIRMAN DYKEWICZ: Dr. Parsons, you had a question on that slide?

DR. PARSONS: On that slide, could you just walk us through what the n's are, please?

DR. JONES: The n's at the bottom here, there were 32 patients with dry mouth, and 316 with no dry mouth, and albuterol, seven patients with dry mouth, and 333 with no dry mouth. Placebo, seven with dry mouth, and 302 with no dry mouth.

So it is very much a minority of patients who had a side effect signal, whether or not they were receiving the treatment.

DR. KAMMERMAN: I want to make sure that I understand. We just look at the patients who had the dry mouth, and the bar is on the left, and that there was an effect, I think, because in the tiotropium, there were 30 percent responders than 10, or 15 and 15. What am I missing?

DR. JONES: Among the 32 patients who had a dry mouth, the response rate on the TDI was 28 percent, and among the 316 patients who did not have a dry mouth, the response rate on the TDI was 44 percent.

DR. KAMMERMAN: Well, just looking at it, there appears to be an interaction because no dry mouth clearly is going down almost linearly, but those with a dry mouth have an increase for tiotropium, and for the other two arms, they level out. I think if you did an analysis with contingency tables, you would see a correlation of some sort.

You may, but of course within there is a

treatment effect as well. The placebo presumably didn't have a treatment effect, and in some of the studies, and in the TDI studies, salmeterol had a smaller treatment effect than with Tiotropium.

So one would expect an interaction with the treatment, because it is an active treatment.

DR. SULLIVAN: I wonder if you have a rationale for this observation. It seems paradoxical to me if the dry mouth is a systemic manifestation of exposure, then those with dry mouth likely had more drug delivered to their lung, because that is where the absorption comes from.

And yet those patients with more drug delivered to their lung seem to respond not as frequently. Is there a rationale for this observation?

DR. JONES: I think there are two rationales. One I see that Dr. Disse would like to answer from the pharmacological perspective. I think that there is a psychometric perspective; that we know that patient's pre-inspection of breathlessness can be altered by blowing cold air on to their face, or

blowing air up their nose with no change in alterial blind gases.

So sensations around the face alter patient's perceptions of breathlessness. So one explanation of this is that a dry mouth makes people feel less or more breathless, or they don't perceive a symptomatic gain compared to when they do have a dry mouth.

CHAIRMAN DYKEWICZ: Yes, please proceed.

DR. DISSE: I think we have to take into account that dry mouth reflects two things. One is of course sensitivity of the individual patient, and the second may be elevated systemic levels.

But this has not necessarily to do with drug levels in the lungs. In fact, we have analyzed patients with dry mouth and those without for the FEV1 response and there is no difference.

DR. KAMMERMAN: I just want to say that -- and I have just been thinking about this, but that if there truly is no relationship between the outcome on TDI and whether or not a patient was experiencing dry mouth, you would see the same slope from tiotropium,

to salmeterol, to placebo, and that isn't want is being shown here.

DR. JONES: I think the basic factor in that start date, in that analysis, is that patients with dry mouth had a smaller response rate than patients who did not have a dry mouth across all treatment groups.

DR. KAMMERMAN: And I agree with that.

DR. JONES: But if we start looking at the end, and if I could have that slide up again, please. The ends are now getting very small down here. It is 7 out of 300, and so the power of any direct comparison is going to be small.

But there is nothing in this data to suggest that patients with dry mouth had a higher response rate. That is the only point that we can make of it.

DR. KAMMERMAN: Well, it isn't so much that they had a higher response rate. It is whether the response rates differ according to whether they had a dry mouth. Among patients who had a dry mouth, their response rates -- am I explaining this

correctly?

The question is if somebody is a responder, or has dry mouth, is the difference between responses in tiotropium and placebo the same as the difference between those who don't have dry mouth on tiotropium, versus placebo.

And from the picture that you have drawn here, and I don't have the numbers in front of me, it looks like that isn't the case.

DR. JONES: I think it is a reasonable hypothesis. The point is that we would never be able to test it with the numbers, because as I pointed out, they are too small. But as I said, there will be an interaction because we would expect on the basis of the other data that the tiotropium treated patients would have a higher response.

But I think we can take this higher response in the salmeterol treated patients, but there are more patients with tiotropium who have dry mouth and no dry mouth, compared to salmeterol and placebo, but we are putting this slide up to show that there is nothing in this data that we can see to suggest that

people with dry mouth have more -- were responders rather than those that did not.

And I think that we would make no further point than that. May I go on to --

DR. KAMMERMAN: I'm sorry, but this is my last comment. The question is not whether people with dry mouth had different response rates than those without dry mouth. The question is are those who are on tiotropium and had dry mouth, did they have different response rates than placebo patients who were on or had dry mouth, compared to those who didn't have dry mouth at all?

DR. JONES: If I could have that slide back again. I think the -- I think I will need some notice of your question to fully interpret it. There are a greater percentage of patients who have -- we would need to do a statistical analysis to see the size of that interaction. I think that is all that we can say. I think that is all we can say. Could I go on to another point?

CHAIRMAN DYKEWICZ: Please go on to the next point.

DR. JONES: May I have this slide, please. Another point is that ipratropium and tiotropium both cause dry mouth. So that we -- that if the assumption is that there is a signal coming through about active treatment, and what the cause is to responder bias in favor of tiotropium, we should see that.

And we should not see so much of a difference between tiotropium and ipratropium. May I show the third slide, please.

You have seen this slide before, and initially tiotropium had a bigger improvement in breathlessness compared to ipratropium, and that certainly we could not exclude the possibility of there being some bias being introduced.

But if we then look at what happened to the ipratropium treated patients during the study, they became worse. There was obviously some underlying, other biological factor that was going on unrelated to the treatment perhaps, but we see that the change in tiotropium treated patients track that change in a very similar way.

And I think it is reasonable to postulate

that if there had been observer bias in terms of treatment effect, that that effect would have at least have been sustained in some way, and we would not have seen this tracking of what happened in the ipratropium treated patients.

And I would just remind you that these patients, also that some of them had dry mouth as well, and in the other study, we see a similar picture. There isn't quite as much fall-off in the tiotropium treated patients, but again the patterns are very similar.

And I would argue that if there was consistent bias here that we wouldn't have seen this pattern.

CHAIRMAN DYKEWICZ: What I would like to do is to have questions specifically directed on this point, Dr. Jones, because then we will break for lunch thereafter. All right. We will resume at 1:00 p.m.

(Whereupon, at 12:01 p.m., the committee meeting was recessed.)

A-F-T-E-R-N-O-O-N S-E-S-S-I-O-N

(1:03 p.m.)

CHAIRMAN DYKEWICZ: Okay. Let's reconvene. Welcome back. What we are going to do organizationally is first give the session for open-public hearing, which I think will be relatively brief, and then we will give an opportunity to the sponsor to address some issues that were unresolved prior to the break.

I would also say that the committee has received 55 e-mails discussing the topic of discussion today. The Chair recognizes Dr. Wlodzimierz Rozenbaum. Please identify your affiliations and any conflict of interests, and your comments are limited to five minutes.

DR. ROZENBAUM: Good afternoon, and my name is Wlodzimierz Rozenbaum, and I am the owner-moderator of COPD-ALERT, a non-profit, internet-based, support and advocacy group for COPD patients, caregivers, and medical professionals.

COPD-ALERT is a member driven organization, and we do not receive any funds from any

private organization or government agency. I also have a personal stake in your hearing. I have severe COPD, and I was forced to retire on disability more than two years ago.

On behalf of COPD-ALERT, and many thousands of COPD patients in the United States, I wish to thank the agency and the committee for holding a hearing devoted to Spiriva, and for making it possible for the patients and the advocies to participate in and contribute in your deliberations.

The name Spiriva evokes strong emotions among COPD patients. Medical reports about successful clinical trials conducted around the world, as well as comments about it, have been proliferating exponentially.

There is also quite a bit of anecdotal data from individual COPD patients which adds the human dimension to the formal clinical reports.

This excitement is quite understandable. To this day, there is hardly any COPD-specific drug available. This is despite the fact that COPD is the fourth major cause of death in the United States, and

that the morbidity and mortality figures continue to climb.

There is a real danger that within the next decade that COPD will move to the third place, if not higher. It is our hope that medical research accelerates the development of COPD-specific drugs, like Spiriva, which in addition to its proven therapeutic efficacy, causes no major side-effects.

We must at least slow down the COPD deadly spiral, if we cannot stop it. But COPD is not only about death. This is a crippling, debilitating disease, tying patients to breathing support machines, and mercilessly destroying their lives, breath by breath.

The Work Bank study suggests that some 25 percent of COPD patients will die during their productive middle age, losing 20 to 25 years of life. At the same time, millions of COPD patients who continue to struggle with their disease are disabled and unable to work.

Now, the American Lung Association has described COPD as the second most disabling disease

for American workers. It is a small wonder that the economic costs are enormous.

According to the Centers for Disease Control, more than $50 billion per year, a conservative estimate, is spent on COPD-related medical expenditures, with an additional $50 billion in indirect costs.

The primary source of medical expenses for COPD patients are extended hospital stays and expensive medications. The University of Washington's alarming study shows that while COPD patients constitute 10 percent of the patient population, they account for more than 70 percent of all medical care costs, and these costs continue to escalate.

COPD is a neglected disease. Insufficient attention is being paid to the fact that there is an extreme shortage of viable treatment options. Physicians have only two choices: to experiment with medications developed for asthma, or to consider surgery.

Asthma medications relieve symptoms, but their effectiveness diminishes over time, and they

often have undesirable side effects. Surgery is an option for very few patients. This is why Spiriva has evoked so much interest and hope among COPD patients.

After all, tiotropium bromide is not a mysterious new substance. Both asthmatics and COPD patients have been using its variation, Ipratropium bromide, for many years. Altrovent, unlike many other bronchodilators, is well tolerated and does not cause worrisome side effects.

As the clinical trials in this and other countries have shown, Spiriva is well tolerated and provides a kajor relef for shortness of breath for as much as 14 hours without causing any harm to patient's other organs and systems.

It is my understanding that this Committee has received credible and uplifting testimonials from individual COPD patients, who take Spiriva under the supervision of their doctors.

COPD patients expect that this Committee and the FDA will move fast forward towards the approval of Sprivia. We urge you to do so. Thank you very much.

CHAIRMAN DYKEWICZ: Thank you. We will now proceed with the opportunity for the sponsor to respond to questions about the instrument methodology, and issues about the clinically meaningful response on the TDI, and then also permit Dr. Jones to give some further clarification.

DR. WITEK: Thank you very much, Dr. Dykewicz, for this opportunity, because it is very important that we put some of the comments into perspective for better understanding.

There were several points raised regarding issues of training and the lack of us documentating inter-rater reliability, et cetera, and the reason why this is important to us in clinical development programs is that these things must be guaranteed in order for us to show an effect, because if they are not manifesting, we lose sensitivity.

And the fact that we have shown, as I have shown you consistently in these studies the effect, we believe that those issues are acceptable here. The other point before we get to the points of bias, and we will let Dr. Jones finish his question, just a

little bit about perspective with respect to the differences, let's say, of 15 percent.

There are other drugs that are widely used that have used symptomatic benefits in their clinical development program, and here we have seen, for example, with antihistomines, for rhinitis, and NSAIDS for osteoarthritis, and our own drug, Flomax, for BPH.

There in those studies, we are looking at responder rates to symptomatic benefit. The range that is seen is in the range from an 8 percent difference to a 15 percent difference between drug and placebo.

So that also gives you a little bit of a perspective regarding the differences that we have observed here in our responder rates. What I would like to do now is have Professor Jones finish his discussion around the issues of bias, and then we will certainly be available to answer any questions regarding my comments that were just made.

CHAIRMAN DYKEWICZ: Dr. Jones.

DR. JONES: Thank you for giving us the opportunity to respond, because I am very sorry that

Dr. Kammerman is no longer here, because she has raised some very important issues. I think we were about two-thirds of the way through. She raised concerns that -- two concerns.

One is that the SGRQ, which is a health status instrument that addresses issues around disturbances of activity, among other things, before the patients responded to the questions about the TDI, she was also worried that the clinician would know the patient's FEV1 response, and that may have conditioned the way in which they scored the TDI.

I think there are two points about this. First, if we deal with the SGRQ. The SGRQ and the TDI in some respects address very similar issues. The TDI, or the SGRQ has got items such as being breathless, and walking upstairs. That is the type of thing that is addressed by the TDI.

So one would expect concordance there. And it is very difficult to imagine a circumstance whereby the information in the SGRQ should be different from the information used for the TDI. They are very similar.

The point about the SGRQ is that it is a point estimate. The patient has no idea what their previous score was. They are not given it, and they are not given their previous questionnaires.

And as Dr. Kammerman pointed out, it is actually very difficult to remember what your health status was in the past, which is why the TDI refers to the patient's baseline index, and each time the measurement is made, they refer back to the baseline index.

And there is no way that they know how they previously administered the SGRQ. So I do not believe that there is any way that the SGRQ responders should have contaminated the TDI response.

The other point that she made was about the FEV1. It is perfectly feasible that if a patient has a big change in FEV1 that any reasonable observer will think, okay, I can see a big change in the FEV1, and there must have been a big symptomatic improvement.

If that were the case, one would have seen a tight correlation between the TDI score and the

change in FEV1, but it wasn't. It was at 0.21, which is exactly at the level that we have seen in other clinical trials and in other studies using similar instruments, and indeed with the TDI.

So I don't think it is a very real concern that she has had, but I don't think that there is any evidence from this data that there has been contamination of the observer by either the SGRQ or the FEV1.

CHAIRMAN DYKEWICZ: Thank you. Question from Dr. Schatz.

DR. SCHATZ: When you mention -- the issue of recollection. Are patients actually shown what their BDI is, and then asked to respond to that? Is that the way it is done?

DR. JONES: Correct. That is the methodology.

DR. SCHATZ: And is there any particular reason -- in other health related quality of life instruments, the same instrument is just administered, and sensitivity is looked at over time. Has that been down with the BDI?

Is there any reason to think that the BDI administered, which doesn't require any recollection, would have been an alternate way to do this?

DR. JONES: That is a good point. The original version of the Chronic Respiratory Questionnaire was designed to be administered in the same way as the TDI. The patients were given their first score, and then they were asked to score the subsequent ones in relationship to the original one.

Gordan Guyatt has not changed that and said that the patients don't or aren't given their previous score, and a number of us have felt that that was not necessary. Our instrument isn't referred to at baseline state.

And I was discussing with Dr. Mahler yesterday abou why not just administer the BDI as a point estimate at each time, and we both agree that that is a very sensible way forward.

We should understand though that at the time that the CRQ and the TDI were developed that psychometricians -- and Professor Feinstein was one of them -- believed quite strongly that one needed to

anchor a state to get sensitivity to change.

I think the science has developed since then and we know more.

CHAIRMAN DYKEWICZ: Dr. Patrick.

DR. PATRICK: If the content of the SGRQ is similar to the TDI, then I believe that we passed by a slide. Wouldn't you see very high correlations between the changes in the SGRQ and the changes in the TDI?

DR. JONES: In the briefing pack, there are data showing the correlations obtained in the tiotropium studies, and if I remember correctly, the correlation between change in SGRQ and TDI is 4.5, which is lower than the cross-section of correlation between the BDI and the SGRQ at baseline.

DR. PATRICK: And is that what you would expect?

DR. JONES: Yes.

DR. PATRICK: Wouldn't you expect higher?

DR. JONES: No, I wouldn't, because the range of changes that you obtain -- and as you know, with all the longitudinal studies, the correlation

between two measures is nearly always lower than any cross-sectional studies.

And the reason for that is the range of variation the data is generally speaking smaller, and so one ends up with a weaker correlation.

DR. PATRICK: Okay. Just one last question on this. If we know the minimally important difference in the SGRQ, wouldn't one way to do this to anchor the TDI would be to anchor the changes in the TDI to the SGRQ, and did you try that?

DR. JONES: It has been done, and it does then raise the question about the validity of the four unit change in the SGRQ.

DR. PATRICK: Right.

DR. JONES: And there is an analysis showing that they are in fact really quite closely related. But I think one gets -- it is a peace of evidence that supports the threshold for the TDI. It doesn't confirm it.

As you know testing the validity of instruments such as this is brought up through a body of evidence that shows consistency, and it is one

piece of consistent evidence.

CHAIRMAN DYKEWICZ: Thank you. Dr. Stoller.

DR. STOLLER: I have two questions, and one is a follow-on for that, and it is really a follow-up to the question that I posed earlier with regard to the minimally important clinical difference.

Recognizing the difficulty of identifying a gold standard for minimally important clinical change, and the somewhat arbitrary nature of those definitions, however well respected, I still -- and leaving aside the methodologic issues, I am still interested in the data distribution on Table 3.3.2 with regard to when the data word dichotomized is greater or equal to one DTI unit, there are mean values about those responders versus non-responders.

And the table provides meaning and standard deviation data, but not distributions. This paper was advanced as validation of the minimally important clinical difference in the paper in press.

And so it becomes germane, recognizing that one is not anchored necessarily on the other, and

it is nonetheless advanced as a criterion of further support of the relevance of a one unit change.

And I wonder if that distribution data are available so that one could ascertain whether this mean value is due to a few outliers, or whether it truly reflects some more homogeneous clustering of a greater than four point unit as a correlate of greater than one unit. Does that make sense?

That was the question that I asked before, and it got lost in the flurry of other issues.

DR. JONES: No.

DR. STOLLER: Okay. My other question is to Dr. Jones, and it regards some of the methodologic issues. You know, given the attention given to the SGRQ with regard to British and American translation, and the subtleties of the index and recognizing that it has been shown to be reasonably good in that context in the one study of which I am aware, I wonder if there is any concern about the very issues that we were talking about before.

That is to say that the presence of correlation in non-English speaking and English

speaking is not quite the same level of precision of attention to the reproducibility of the instrument as one would have in a head-to-head comparison in as subtle a difference as British and American English.

And so it gets to again this substantive concern that I think has been raised about how one would approach the methodology of being convincing if one designed this a priori as the primary outcome measure, as opposed to the kind of methodologic afterthought of using this as a co-primary outcome measure after the actual administration and training, and so on.

It gets to your level of concern, having studied this with the St. George's about the -- you know, about how much of an issue in your mind, and how to explain the disparity between the level of attention given to some other indices, in terms of minimally important clinical difference, and the relative absence of that with regard to the index used as the co-primary outcome here, the BDI and TDI.

And I ask that question as someone who has been very interested in the Baseline Dyspnea Index and

someone who has published, like Don, having worked with Alvan on this very index. So I would be interested in hearing your thoughts about that.

DR. JONES: You raise a whole host of very interesting points, and I will try and keep my responses brief, although I would like to make them longer. The first point is that I share your concern about adequacy of translation, and I have written about validation in different countries, and it is a very different process and difficult process.

These questionnaires I find remarkably robust in our hands. They are much more robust than people thought they would be, but it is very much dependent on having good translation, back translation, processes, and that was done in this case.

So I think -- and in fact I have written as an editorial saying that there are now enough studies validating different translations of our questionnaire, because we know that if the translation and back-translation process is done properly, we can be sure that questionnaires behave similarly in

different countries.

And so with that first slide, we were not able to show the second slide, which was that from the data from the Tiotropium studies, the correlation analysis shows that correlation between the TDI and the BDI, and the reference measures is very similar between English and non-English speaking countries, as good as I could have possibly expected.

The other important point about that is that these data are remarkably consistent across clinical trials, and across continents, and across languages. The size and effect of tiotropium compared to placebo in the U.S. was really very similar to the size and effect seen between tiotropium and ipratropium, an active drug in The Netherlands.

Another point about the translation is that one of the advantages of the BDI and TDI is that they are interviewer administered. So that you have to train fewer people. For example, there are fewer opportunities for misunderstandings as a result of the translation process.

When one does this translation, back

translation, process and have focus groups, you find that you get the best possible cultural and linguistic validation. Just one antedote. When the American version of the SGRQ was created, the focus groups could not agree on one particular aspect of it. So we incorporated both.

So even focus groups don't always get it right. But I am confident that the translation and back translation process that was done here was adequate. That the training of the interviewers was adequate.

As you know, if you don't get the interviewers to use the instrument properly, it results in poor psychometric properties. It increases the noise and reduces its sensitivity.

So quite clearly the agency's concern is going to be that somehow the company has exaggerated the treatment effect, but really all of Dr. Kammerman's concerns about the validity of the instrument in different countries, and the way that it was applied -- you know, I really want these instruments to be trained and used properly.

I think they would work towards reducing the effect size and not increasing it. I know of no study where bad technique, unless it is unblinding leads to an exaggeration of the effect size.

So just as an independent observer, I believe that the methodology was sound enough. I am sure -- and I was not involved in the change to the placebo, but I am sure that if this was going to be the co-primary end point, more effort would have been put into it, which would have tightened up the results yet further. It would not have reduced them.

CHAIRMAN DYKEWICZ: All right. Thank you. Dr. Apter.

DR. APTER: I am still confused. We are supposed to distinguish between relief of bronchospasm and relief of dyspnea, and bronchospasm has a physiologically accepted measure, the FEV1.

Nevertheless, if you relieve bronchospasm and you administer the TDI, I am sure that patients would say that they could get dressed better, dress breathlessly, walk up hills better.

So I am not sure -- and we have no good

physiologic measure of dyspnea. We have the pO2, but that wasn't measured here and we are not really talking about that in these patients.

I am not sure that we are able to distinguish between relief of bronchospasm and dyspnea at the clinical level.

DR. JONES: My colleagues are looking to me to respond if you would like. I think you raise a very important point. Dyspnea is a sensation, and like pain, but far more complex than pain. It is a result of a number of different pathways.

And we know that there are a number of different measurable physiological variables that contribute to breathlessness. It is largely related to the work of breathing, and the work of breathing depends to some extent on the compliance, the stiffness of the lungs, and the lung volumes.

So there are a lot of different factors that will influence the overall perception of breathlessness, and a pharmacological agent, this is a very simple pharmacological agent. All it does is that it dilates up the airways.

But in fact probably more important in terms of breathlessness is that it allows the lung volumes to reduce so that the work of breathing is less, and so the patients feel less breathless, and there have been various studies done not using tiotropium, but using other bronchodilators, showing that the improvement in breathlessness correlates better with the improvement in lung volumes and the work of breathing, than the changing in FEV1.

So there is a link between bronchospasm and breathlessness, but it probably is mediated through another, or two or three other physiological mechanisms as well. I don't know whether that has answered your question a little too tutorial.

DR. MAHLER: May I also address that question?

CHAIRMAN DYKEWICZ: Yes, you may.

DR. MAHLER: Your question hits a key area in our pulmonary community, and that is that we have had an over reliance over the years, decades, on FEV1 as a primary outcome measure, and I think as we have done studies looking at dyspnea measures, whichever

one you want to use, health status measures, we see very modest correlations between FEV1 and dyspnea, and health status.

And I think it means at least to me and to many of the people that I interact with, that they are really measuring different constructs, different components of the overall disease COPD.

So I think we can say, hey, FEV1, bronchodilation, dyspnea, a subjective sensation that relates to air flow obstruction, that relates to hyperinflation, and that relates to psychological issues, and that relates to deconditioning, and all we are trying to do is say let's get a global score for dyspnea, and let's get a global score for health status.

And let's elevate that to comparable levels in looking at what we do in treatment wise. And I think the goal guidelines that we are aware of and that were published last year, illustrate what we are supposed to do in COPD, and they say strictly that all of our evidence indicates that we are treating the symptoms of COPD because none of our other

interventions treat any of the other major outcomes -- survival or change in FEV1-- other than smoking sensation, in oxygen therapy. So at least that is my perspective on your question.

CHAIRMAN DYKEWICZ: Dr. Patrick.

DR. PATRICK: I would just like to follow that up while you are up there, because both of your opinions would be -- and I think we are at one of the hearts of the matter here, which is what makes the BDI and the TDI a measure of dyspnea.

And having this as a measure of dyspnea, that is in some cases responded to or recorded by the patients, and in other cases it is recorded by the interviewers. According to the protocol, it was supposed to be by the interviewers.

If this is subjective sensation, what makes this system, this system of measurement, an idea system for measuring dyspnea, and is it a measure of dyspnea, or is it a measure of the impact of dyspnea.

And I am very confused when I read the instrument. To me it is an impact of dyspnea, because a patient could sit at home and do nothing, and not

have dyspnea.

DR. JONES: Could I first comment, and then let Dr. Mahler respond. I very much understand your perspective, and it is related to that slide that I showed showing the relationship between dyspnea and exercise is complex.

There were levels of exercise that caused dyspnea, and there were levels of exercise that can't be done because of dyspnea. I think the important thing is that as I said the dyspnea measurements are grounded in the metabolic costs of the activity.

Ideally, we would measure it in a laboratory, but we can't. We use these reports of daily life. And we are assuming -- well, we know that there is a graded level of activity, and that that is the stimulus to breathlessness.

But you are absolutely right. That if patients get breathless, they will stop the activity. The two are inter-related and it is impossible to deny that. However, the development of this instrument was very much from a clinical perspective that the driver for the breathlessness was the

activity, rather than from my perspective, and to some extent your perspective, the patient's view of the impact of the disease.

DR. MAHLER: Yes. Again, I would agree with your statement and what Paul said. I mean, how do we measure pain? We say here is a visual analog scale, and mark the intensity of your pain that you are having.

Is that really measuring paid? Not really. It is measuring the impact or the perception of pain by that individual, and I think we are stuck with that same circumstance in breathlessness. And the people who said, hey, let's development some instruments for quantifying the sensation, because we think it is important.

Yet, we don't have these perfect ways to measure it, but we are developing more and more ways to understand it. And without going through a lot of detail, we have got all kinds of validity, and reliability, and responsiveness, captured around the BDI and TDI, that at least have convinced a lot of people that it is a reasonable instrument to use to

look at outcomes when interventions are applied in COPD and other chronic respiratory diseases.

CHAIRMAN DYKEWICZ: Dr. Schatz.

DR. SCHATZ: Just another methodology question that was brought up. The concern that it is not clear that the three additive or the three factors that are added are in fact added, and are in fact independent, and that that is the best way to score that.

I wondered if you had any additional comments on that.

DR. JONES: Perhaps I should let Dr. Mahler comment first, and then --

DR. MAHLER: Well, we set it up that conceptually functional impairment, magnitude of task, and magnitude of effort, are distinct components or contributions to the severity of breathlessness.

We have not done any formal testing saying, well, should we weight one, versus another, and we have not done that. And I think how would you do it? Well, there are statistical ways to go about it. On the other hand, as was pointed out in this

data, as well as in other data materials.

It is very seldom that the person gives a positive response on the TDI in one component, and a negative response in the other component. And I believe that emphasizes that everything is moving in the same direction because I can do things easier, and I don't have to pause as often, and all those things have enabled me to do my work outside the home or inside the home, and they kind of parallel each other.

I can't say from a statistical point of view that they shouldn't be weighted or there is no absolute overlap completely.

DR. JONES: May I add that I believe that redundancy of this type is actually valuable, because it increases the precision of the estimate. It is like triangulation or making duplicate estimates in our bioessay.

So in fact I learned when developing our instruments, I learned from this approach, and I do believe that redundancy is actually valuable in this instrument, because it does increase the precision.

CHAIRMAN DYKEWICZ: Dr. Sullivan.

DR. SULLIVAN: I just wanted a chance to follow up on one of your comments regarding the call to baseline, and the extended duration of these studies, as compared to perhaps the validation studies, and that many of them are published on the TDI.

As you mentioned many of the validation studies are -- some are interventions, and most of them aren't drug interventions, and the drug interventions tend to be shorter studies.

So one of our concerns has been how well the patient can think back, and something that seems to comfort you is the fact that they are presented with their BDI score, and then asked to say how they changed.

But I wondered if you could comment. You are allowed to show an improvement of plus one if you discern a change within a grade. So the BDI is -- you are assigned a grade, and so presumably six months later you are told what grade you were in before.

Now you are able to report a plus one change if you are better within that grade. So

doesn't that still mean that you have to recall quite well how you were doing in that time past?

DR. MAHLER: Basically, the baseline information is given back to the patient, and rather not the absolute grade, but here is what you told me before. You have difficulty in certain tasks, and you have difficulty in certain efforts.

That would be the intent in providing the information, and not saying, oh, you were a grade one on magnitude of task, and have you changed a half-a-grade here. That would be impossible.

DR. SULLIVAN: Perhaps the company can respond. Then you are saying that there is more information available to the interviewer than just the grade. There is notes from a clinical history taking, and I can see why that would happen in the clinical setting, but I am not sure at a clinic visit for a study whether the interviewers had the information you are saying.

DR. MAHLER: Well, you would not have to necessarily have comments written on the side. You could simply read the information, the criteria, for

that grade back to the person and say, well, you told me that you had trouble walking up a hill, or whatever the specific criteria is.

DR. SULLIVAN: So you would read or describe the grade.

DR. MAHLER: This is what you told me.

DR. SULLIVAN: But then they are able to say I am still that grade, but I am better within that grade.

DR. MAHLER: And then the interviewer has these criteria for the TDI in front of him or her, and then through the interviewer process tries to tease out what the magnitude of change is, and that's why we think an experienced interviewer, someone with knowledge and experience about respiratory disease, should be an interviewer.

DR. SULLIVAN: It still seems to me that the patient will have to recall where they fit within that grade back six months ago, and that it is not --

DR. MAHLER: They are going to have to recall how they were doing at that time period.

And all I can say is that an observational

study in COPD over two years, we have seen a steady decline of .7 units over 2 years in our patients with COPD who have had, quote, optimal treatment at our institution.

So I think that component fits with clinical experience; that is, people's breathlessness tends to get worse, and at least on the TDI it is represented by comparing to their baseline state.

DR. MAHLER: And of course their memory, and their recollection of how they did two years ago may change it, and after a year or two, I may think that two years ago I was better than I really was.

DR. SULLIVAN: And I think that is a potential limitation of the instrument.

DR. JONES: Could I just answer that, Dr. Sullivan? It is a good point, and in other areas in this field it has been known as response shift.

The point is that it leads to insensitivity, rather than increased sensitivity, and if I were to design a measure for a one year study, I wouldn't base it on this, because I would be concerned that there may be a response shift and the failure to

remember correctly would increase the noise relative to the second.

DR. SULLIVAN: I understand that argument, and I guess sometimes it is periless to determine what might have been shown if it had been done more rigorously, and I can see why theoretically you would think it would decrease the noise, but we have to address what was actually done and what the data are.

DR. MAHLER: Can I also point out that if any intervention shows no change over a period of time, and if you accept that the natural history of the disease over that same time period is a negative direction, no change or maintenance of your severity of breathlessness is actually an improvement.

CHAIRMAN DYKEWICZ: Dr. Stoller, did you have a question?

DR. STOLLER: Again, with regard to the kind of methodology of the administration of the instrument, recognizing that these studies were conducted obviously in many countries, and in many centers, the question is who were the actual interviewers?

What was their skill set, and who were they? You know, characteristically, when this was developed by Dr. Mahler, this was administered by lung doctors, and so on, and leaving aside the issues of training, simply the skill set of the individuals administering it.

DR. WITEK: Yes, to get to your question, Dr. Stoller, I don't have the exact education level or training level of the coordinators and the people that were interviewing the patients, but I could say in general that these are nurses, respiratory therapists, or lung function technicians, to give you the range of those patients.

CHAIRMAN DYKEWICZ: Dr. Parsons.

DR. PARSONS: Just to follow up on that. Was there any specific training or was there a guide written for these people to follow? In other words, if they didn't have experience administering this questionnaire before, which is likely, were there some guidelines that they were taught, or was there some attempt over 80 centers to make sure that everybody was doing it right the same way?

DR. WITEK: Yes, and in our investigator meetings, as part of the process of reviewing the protocol and learning how to use the centralized spirometry, that is where we have the centralized training for all of those individuals that participated in the study. So it was really limited to that investigator meeting.

DR. PARSONS: But those weren't the people administering the questionnaires?

DR WITEK: For the most part. I think I can't give you the exact number, but the study staff that reports to the investigator meetings are typically the ones responsible.

CHAIRMAN DYKEWICZ: Dr. Stoller, did you have a follow-up question?

DR. STOLLER: Just a clarification on Dr. Witek's comment. So do I understand your response to be that every one of the study coordinators was either a pulmonary function technician, respiratory therapist, or nurse?

DR. WITEK: No, I can't give the hard data of the background, but in general those are the types

of individuals that are conducting the studies, yes.

DR. STOLLER: Absolutely. I understand.

CHAIRMAN DYKEWICZ: Dr. Patrick.

DR. PATRICK: So if we go to the actual BDI and look at things like usual activities, did the interviewer at BDI list those usual activities?

DR. WITEK: That I am not certain, but that would be on the SGRQ, the last page.

DR. PATRICK: I know, but when they get to the follow-up, because what Dr. Kammerman showed was very interesting, was some examples of things like able to do things more rapidly, and able to do things with more vigor.

Now, vigor and doing things rapidly are words that are not necessarily on the BDI, but would be an interpretation by the interviewer having discussed with the patient, I'm assuming, what tasks they do do. And I would imagine that we are blind to what is actually the content of what those changes are. Is that correct?

DR. WITEK: That last point I am not sure of.

DR. PATRICK: Let's say that you --

CHAIRMAN DYEWICZ: Please speak into the microphone.

DR. PATRICK: Let's say that you were -- that your usual activity was to go grocery shopping, and it was terribly difficult for you to go grocery shopping at base line. So your BDI was that I had given up grocery shopping, grade one.

So when you came back, and the interviewer talked to you, it would be up to you to talk about grocery shopping, or would the interviewer say at base line that you told us that you had given up grocery shopping. Are you grocery shopping now?

DR. WITEK: I think the specific comments are not necessarily always documented.

DR. MAHLER: A good interviewer would say, just like a physician taking a medical history, what kind of activities are you doing now, and how are you able to go grocery shopping now compared to six months ago.

Let's say the person stopped going grocery shopping and is just hanging out at home. As part of

the questions, you should also ask are there any activities that you stopped doing or have abandoned, and if so, why. Is that because of breathlessness

Now, again, you could say that is an advantage of this interviewer approach, and you can get subtleties out of it, or you could say it is a disadvantage because it depends on someone probing.

But as opposed to a self-administered, you simply have a few boxes to choose, and you can lose that subtlety, and we believe that is important in the responsiveness of the instrument.

CHAIRMAN DYKEWICZ: Dr. Sullivan.

DR. SULLIVAN: Dr. Mahler, I think that gets to maybe clarifying between you and the applicant, but when you were discussing a good interviewer, and that is the way that you designed the instrument, and so the good interviewer would have the clinic notes from the last time.

And it says here the last time that I talked to you about grocery shopping, and you have given it up. In the clinical trial, the interviewer is going to have the case report forms.

DR. MAHLER: You would not necessarily have those comments, but --

DR. SULLIVAN: But there would be no way to know about grocery shopping unless the patient brought that up. You could ask generally on --

DR. MAHLER: You could ask generally and then zero in on what activities you are doing, and have you stopped doing anything, or are there some new things that you are doing because you can breath better.

DR. SULLIVAN: I think that brings out an important difference between the way the study was designed and is used in certain circumstances, compared to the way that it is used in a clinical trial.

And where in the clinical setting, you have your notes. It says here grocery shopping in my handwriting from six months ago. This is now six months later, and I have nothing, and unless the patient offers that, I ask the general questions, and perhaps the patient will remember that six months ago I had given up grocery shopping and I am no longer

doing that.

Or grocery shopping used to be difficult and it is still difficult, but I am a little better at grocery shopping. But I wanted to clarify that point, because it is a point of concern that we have regarding how it was implemented in the trial.

DR. MAHLER: I can't comment on how other sites or study coordinators apply it. But certainly at our site, people frequently will scribble things down on that sheet of paper as part of the form, and include those activities, and whether that is done in other sites, I have no idea.

But you should be able to in the interview process be able to pull those things out relatively quickly.

CHAIRMAN DYKEWICZ: Thank you. Are there any further questions of the sponsor or the FDA from the committee? Dr. Chinchilli.

DR. CHINCHILLI: Yes. This morning, Dr. Kammerman alluded to the fact that there was some data imputation with the TDI, and I was wondering if the sponsor could elaborate when the analysis was done,

what form of data imputation was there?

DR. MENJOGE: You know, there is no perfect solution for the missing data. However, what we did was actually we used the last observation carried forward method, and only in the case of the worsening of the disease, which is about less than 5 to 10 percent of the patients, and we used the last observation carried forward, and that is the techniques that we used.

And we did the analysis with and without imputation, and they basically showed the same results.

CHAIRMAN DYKEWICZ: Dr. Atkinson.

DR. ATKINSON: Yes. This morning, I believe they mentioned, or the company mentioned that there were four patients that had had urinary obstruction requiring catherization, and I was wondering how long, and if they had any clinical information on how long that condition had persisted, and how long it took to resolve.

DR. KESTEN: Those events were generally 24 hours to several days, and there were one and two

patients who had follow-ups with either medication for BPH, and some subsequently had trans-retheral reception of the prostate. But the period of catherization was temporary.

CHAIRMAN DYKEWICZ: Any final questions for the sponsor, or the FDA? All right. What I am going to do now is move to the phase of the meeting where we have discussion amongst the committee on the various topics.

And I am going to actually change the order a little bit, because we have been having so much discussion since the return from the break about dyspnea, it would be logical I think to continue on with that discussion.

And so I would like to focus the committee though on several different issues. First of all, and maybe because we have been talking so much about it just recently, what do you think about the TDI as an instrument for assessing dyspnea, and then following that, what do you think about the execution of the administration of the TDI instrument in the studies that are being presented for this new drug

application.

I will open up things generally. Dr. Apter.

DR. APTER: I think for clinical use the TDI certainly is very useful. I think it needs to be altered for clinical trials. I think there have to be ways to write in what the patient said, and activities like the grocery stopping.

For example, what activities in particular or even a set of activities, like a group of activities that equal moderate activity, so that it can be more formalized.

CHAIRMAN DYKEWICZ: Dr. Patrick.

DR. PATRICK: I think we are in the same realm here, and I guess it is because of the age of the patients, and that we are somehow moving towards interviewer administration. As we are in the field of psychiatry, where we often do rating scales based on interviewer questions and observations.

Even if we don't have the specific items, like grocery shopping, in something like the brief psychiatric rating scales, all interviewers would need

to be trained at the same level of standardization.

And so that the inter-rater reliability was documented prior to using it as an outcome measure in a clinical trial. It is my understanding that was not done in this case, and therefore we don't know what was done. Dr. Jones has been pretty convincing that if it was really terrible, we might have seen much more noise and much more difference.

However, this is based primarily on the responder analysis, and not on the mean changes, and the other methodological issues surrounding the statistical analysis of the measure. So I think as the TDI, I would agree with Dr. Apter that it is perfectly adequate as a clinical measure as a staging measure, and for use in clinical practice.

For the use in clinical trials, the rigor of such an instrument needs to be maintained at a very high level in order to be able to interpret the findings, and we have not a clear demonstration that it was administered consistently across the different sites, the translation questions, nor the standardization of the interviewer training.

CHAIRMAN DYKEWICZ: Dr. Stoller.

DR. STOLLER: I would hold the view that as someone who has been interested in the BDI and TDI that this a very clinically sensible measure, and I think is applicable in clinical practice, and in research.

In fact, its very strength, as I think Dr. Mahler pointed out, is that it rests on the kind of information that clinicians would elicit from patients that really go beyond the subtleties of filling out particular boxes, and may escape the opportunity to capture that clinically subtle information in the context of a somewhat more rigorously defined instrument.

And in fact in conversations with Alvin Feinstein about it, in fact the very strength of it was that it was clinically sensible. Now, that said, the appeal of the instrument, therefore, requires the ability to suddenly capture the information.

And so my concerns are not so much around the instrument itself, which I think has the advantages as we have heard very eloquently stated by

the various conversants. But my concern is, and as I think I heard Dr. Jones say, and echo, was that if one were to use this a priori as the outcome measure in a clinical study, one would pay attention to validating its ability to capture those subtleties in ways that were not possible given the evolution of this as a co-primary outcome, with the data admittedly still blinded, but already captured.

That understanding requires faith in the notion that a methodologically, sub-optimally captured measure, would bias in the direction other than the one that we see.

And I guess I am not willing to make that leap of faith in the context of a clinical trial, in which the indication rests on the methodologic solidarity of the instrument to capture that measurement.

So I would say that in response to your two-tier question, and I have great faith in the instrument, and I believe that the instrument can be very carefully calibrated, and I am sure that if Dr. Feinstein were here, he would echo that strongly.

That was the impetus to develop a clinically sensible instrument at that time. But I think he would also say that were he reviewing data in advance of a rigorous conclusion around an outcome measure anchored on this.

And he would say that the methodology needs to be more rigorous around demonstrating the impact of this particular intervention on the outcome measure. And I guess while respecting the breath of experience about the way that bias goes with methodologically sub-optimally captured information, I myself am not willing to make that leap of faith in regard to this, and to the indication with regard to dyspena.

CHAIRMAN DYKEWICZ: Thank you. Ms. Schell.

MS. SCHELL: Thank you. I just have some concerns that I wanted to bring up. I agree that the instrument is a valid instrument, and it has to do with the skill of the interviewer, and not so much as the result, and what I am trying to say is that sometimes the interviewer has to be standardized all

across, because as we know with many of our patients that are being interviewed, their mental state has also deteriorated along with their disease state.

And it is difficult to get answers from them, or correlate those answers, and if the interviewer isn't trained or skilled in interviewing, and getting those probing questions, it is difficult to get a direct answer from the patient.

And so I think it is important that there is a standardization of the interviewer for this process.

CHAIRMAN DYKEWICZ: Thank you. Dr. Parsons.

DR. PARSONS: The only other point that I would like to make is that even -- and I agree with all of the comments that have been made about the TDI, but I think the one other part we have not discussed, or has not come out quite as much is when there are subtle changes in the TDI, in terms of numerical changes, what do those really mean, and it is not clear to me that those really have been tightly correlated with and going out to a group of patients

and saying does this make a difference.

So, yes, indeed, your score may have changed, or you may go through the grocery store a little bit faster and not to denigrate that, but that may or may not make any difference ultimately to somebody's quality of life.

And I think that to use the instrument for research purposes, it would be tremendously helpful to understand what changes in those numbers really mean to patients, and what it means to the quality of life, and their ability to function.

So that you are not just looking at a raw number. You can actually then say this is what the impact is on that number, and what that number means.

CHAIRMAN DYKEWICZ: Thank you. Other comments on the TDI? If not, let's continue to focus on the TDI, but from the perspective of the results that were generated for the new drug application. Do you believe that focusing only on the TDI results, that the improvement that has been reported is clinically significant, clinically important. Dr. Apter.

DR. APTER: I guess I have to say because of all of the methodologic problems, I don't know what to say. I can't be convinced, although it may very well be a good drug.

CHAIRMAN DYKEWICZ: Dr. Meyer.

DR. MEYER: I hope I'm not overstepping my boundaries here, but I would suggest that this question might be helped by saying that if there were no methodologic concerns, and if we had a perfect institution of this, or incorporation of this into the clinical trials, and we saw these results, what would people think of those.

CHAIRMAN DYKEWICZ: Dr. Patrick.

DR. PATRICK: I think Dr. Parsons answered this all before in my view, and without knowing across these patients whether this was a minimally important change to them, it is difficult to say that we have defined this one unit as the MID.

In addition, there is only one possibility here, in the sense that one unit is the minimum amount of change, in terms of the grading. And that issue, in response to would a patient say that this was an

important change, or the smallest important change, is missing information for us. So that is a pretty important piece in the MID.

CHAIRMAN DYKEWICZ: Dr. Chinchilli.

DR. CHINCHILLI: Yes, I agree. Just because there is statistical significance, it doesn't translate into clinical significance, and from what I gather, my clinical colleagues on the panel are struggling with that, as to whether this is clinically meaningful.

So my interpretation of this is that I would say it is not condusive of evidence to say that it is effective based on looking at dyspnea.

CHAIRMAN DYKEWICZ: I will add my own comment. I think that Dr. Sullivan presented some very important analysis on this data, and that was where he was looking at the dyspnea efficacy analysis and mean values in the six studies that were being presented.

And on the question of whether there was a difference of greater than one, which has been proposed as something that would be clinically

important, and if you look at what I would count up to be 27 or 28 time points in these various studies, and bits of data, there were approximately only 12 that there was achievement of either a difference of one or greater than one.

So if you look at it one way, you could say that half the time or more it really is, as Dr. Sullivan has indicated, is not supporting the idea that there is a clinically important difference.

Now, the question also then, and it begs the question as to whether a difference of one is a clinically important difference, and do you potentially have to have even a higher threshold than that.

I think that Dr. Parsons' comments have already addressed that, and I just simply don't know, and whether you achieve a clinical difference of one, whether that is going to represent a significant clinical change or an important change shall we say in the patient outcome.

CHAIRMAN DYKEWICZ: Other comments on that point? Dr. Joad.

DR. JOAD: I would just say that it seems to me that if a change of one, at least in the two of the categories, would be possible, and still you would be within one of the categories within the basic test, and that they have to remember six months back. It just doesn't seem possible for me that that would be a clinically important difference.

CHAIRMAN DYKEWICZ: Dr. Stoller.

DR. STOLLER: Dr. Meyer put on the table the question of if this were a methodologically ideal would we put credence in a difference of one, and it really gets to what are the criteria, and how do we ascertain what is a clinically important difference.

Others have put on the table feeding that back to patients in some feedback and say do you regard this as clinically important. I would regard much of the literature about establishing minimally important differences has to do with the parallelism of other kind of clinical anchors, other subjective measures and other objective measures that in aggregate point towards establishing some threshold that we would regard as minimally importantly

different.

And in my own view, in some ways -- and in fact the validation paper that is in press, of course comes from this dataset, and in some ways the establishment of minimally important difference comes from correlations of lots of outcome variables from lots of different studies that say that these things all move in the same direction or not.

And in that regard, I think leaving aside the methodologic shortcomings, because that is the premise of the question, I would say that I find that the datasets are somewhat convincing in helping me believe that a difference of one is important.

I wouldn't say that I am absolutely from the available information sold on the point, but it certainly moves that issue towards being more convincing to me. Again, the premise of the question being if it were ideally administered, and methodologically acceptable.

CHAIRMAN DYKEWICZ: Dr. Parsons.

DR. PARSONS: Jimmy, can I ask you a question just for clarification. If there was another

drug that was studied in this patient population, and the only outcome was TDI, greater than one, would you be comfortable at this point saying that it correlates well enough with changes in FEV1?

In the past, we have been told that there is a significant change which has been defined in FEV1, and that their assumptions that there are clinical changes that go along with that, based on a lot of information from the past.

So would you be comfortable flipping it at this point and saying if a study came through and all that was measured was TDI, that a TDI of one, a change of one, means that the physiologic variables occurred? I am just curious.

DR. STOLLER: I'm glad that you focused your question that way. I would say at this point, no.

CHAIRMAN DYKEWICZ: Other comments on the TDI instrument, and the data that has been presented? All right. Then going a bit broader, discussing any other end points that were presented to look at the question of dyspnea.

Is there anyone who would like to make some comments relative to an aggregate, and do you believe that there is other data that would be of sufficient enough validity or reliability to increase the assessment, or the confidence of the assessment, that there has been some clinically important change?

All right. Another point that I would like to have the committee discuss is the concept of dyspnea itself as an indication for treatment with a drug? As the FDA has pointed out to us today, this would be a departure from previous practice.

Dr. Apter addressed this to some degree earlier, and I would like some discussion about the indication of dyspnea. Is this something that is important to have, or is this something that is not really of relevance to the prescribing physician. Dr. Schatz.

DR. SCHATZ: To me the answer to that question has to do with the extent to which dyspnea represents something above and beyond the bronchodilator effect, and we have heard both some theoretical and I think some data to suggest that

dyspnea in fact represents more than just a bronchodilator effect.

But I don't hear us feeling that we have seen enough data to answer that question. So my answer would be that I think that dyspnea, to the extent that it does represent something different than a bronchodilator effect would be an important outcome.

Certainly it is an important patient center outcome, but we would need to have, I believe, the clinical tools to be able to sort that out.

CHAIRMAN DYKEWICZ: Dr. Joad.

DR. JOAD: As a general comment, it seems to me that I would like drugs like this to be for an indication other than bronchodilation. As a practicing physician, you don't give someone a drug because it changes their FEV1. You give them a drug because it makes them symptomatically better in some ways.

So I would very much like the indications to be based on a symptom, or on a word like dyspnea. As a pediatrician, I never used the word dyspnea, it just never comes up. Somehow I can take care of a lot

of pulmonary disease without that word.

And it is just a comment observing all of this, that it is such a complex thought, and it includes so many different things, is it useful. I just don't know if it is useful. I am just throwing that out as whether it helps or whether it is just functionally what you can do, and how much you try to do something, and how breathless you get, and your total lung capacity.

I mean, there are so much things that people throw into dyspnea, is a useful construct.

CHAIRMAN DYKEWICZ: Dr. Stoller.

DR. STOLLER: I should say that I applaud the attention to these subjective outcome measures, because I think as has been amply been stated, this is in fact what brings patients to our attention.

And so from a clinically relevant point of view, I deal with patients with dyspnea all the time, and what brings them as I think has been amply and eloquently stated, but what brings them to our attention is in fact this very symptom.

And we have struggled, you know,

clinically with whether these, as Dr. Sullivan pointed out, whether these are really surrogate measures and truly reflective of things that matter to patients.

And dyspnea is clearly that, and so there is no doubt in my mind that the attention to this is an indication for a drug is laudable, and I applaud the attempt to do so. The question in my mind is how convincing has been the ability to do so given the laudability of the goal.

But there is no doubt in my mind that that is absolutely essential and that more attention should be in fact given to these kinds of outcome measures in the assessment of clinical interventions.

CHAIRMAN DYKEWICZ: Dr. Parsons.

DR. PARSONS: I actually don't disagree with that at all, and obviously the reason that I gave any of my patients who have COPD is to decrease their symptoms. The one caveat that I just thought about was that I have a great drug to treat dyspnea, and it is morphine.

And it is not practical. Okay? It is not a good drug for dyspnea in a patient population that

we are talking about. So, yes, I would love to see dyspnea included as part of the evaluative process, but it can't stand alone, because then we can treat dyspnea.

And morphine is a terrible drug and so we need to be sure that we keep that in context. That as these more subjective measurements come along, I think we have to have more ground rules. We need to see other changes in a positive direction somehow related to physiology.

CHAIRMAN DYKEWICZ: Dr. Stoller.

DR. STOLLER: I take the point as to whether the outcome is clinically sensible, and I think that you and I would agree that the dyspnea benefits about morphine are at first pass not clinically sensible in the context of the other effects of this drug.

But in the context of drugs that have other physiologic benefits, but also by the way happen to improve a subjective measure, I don't think you and I would disagree at all about the importance of anchoring an indication for treating a patient, as

well as perhaps approving a drug on a symptom that brings people to our attention. There is no doubt about it in my mind, no doubt.

CHAIRMAN DYKEWICZ: Dr. Apter.

DR. APTER: So what would be ideal would be a combination of measures that get at the patient's perception and are tied in a physiologic benefit that the physician can measure. And of course that doesn't always happen, but that would be ideal.

CHAIRMAN DYKEWICZ: Dr. Swenson.

DR. SWENSON: Well, along those same lines, I wonder -- and I am not an epidemiologist, and I would throw this to those people that think about this all the time, but at some point, particularly using this index in the evaluation of drugs, would it be of some utility to throw the question ultimately back to the patient to ask would they be willing to sacrifice a certain amount of -- for example, their income, appropriately scaled to their own income.

But would they be willing to sacrifice for this benefit and this could conceivably add some value to knowing whether an index of one is a sufficient

improvement.

CHAIRMAN DYKEWICZ: All right. Back to the question about whether it is an important indication to state that dyspnea would be something that would be treatable by a drug. I am just trying to look at it in practical terms, as to whether the prescribing conduct, the prescribing decision making of a physician, would really be altered by statements about specific symptoms that are being relieved.

Symptoms such as symptoms of dyspnea, and maybe exercise tolerance, and wheezing. I think in practice a drug that would state, or a drug insert that would state that the product is for the indication of bronchospasm related to COPD, in essence it is still going to end up being used for treating patients who are presenting as Dr. Stoller's has, with subjective complaints of dyspnea and potentially wheezing and so forth.

So I am not convinced that it is absolutely necessary to position the appropriate use of this drug, and to have it listed dyspnea as an indication. Other comments on that point?

CHAIRMAN DYKEWICZ: Ms. Schell.

MS. SCHELL: I also have some concern. I agree that an indication is a good reason to put the drug out there for that, but when a patient looks at the label and reads this is going to help my dyspnea, and they are disappointed because their perception of their shortness of breath,or their breathlessness is not improved, then we are putting out kind of a message that this is -- well, a hope for them that isn't being succeeded.

Do you see what I am saying? That we are putting out that this is for breathlessness, and I go to the doctor, and I say I want this drug because it is for dyspnea and I have dyspnea, and I come back in six months, and if I am not any better, then I have this perception that this drug wasn't any good.

CHAIRMAN DYKEWICZ: Dr. Joad.

DR. JOAD: Just to follow up on the complexity of the word dyspnea in our conversation. When you mentioned that morphine would fix or might fix dyspnea, it would fix breathlessness and shortness of breath.

You still couldn't walk up a hill and you still couldn't take a shower, or whatever the problem is. So it just strikes me as such a complex inability to do exercises is one thing, and the feeling of the shortness of breath, or breathlessness is another thing.

And throwing them all together into one concept, and that then can be carefully analyzed and given a number to, strikes me as a very hard thing to decide to do.

CHAIRMAN DYKEWICZ: Other comments from the committee on the indication for dyspnea just theoretically from any drug.

DR. SCHATZ: I would just agree with you that I think from the standpoint of getting what appears to be an effective drug by the usual indicators to the people who need it, whether dyspnea is listed as an indication or not doesn't appear to be a major difference.

CHAIRMAN DYKEWICZ: Dr. Parsons.

DR. PARSONS: I was just going to note that if Ms. Schell is concerned that a number of patients

who had hoped that their dyspnea would be relieved, and then be disappointed when it wasn't based on the intent to treat analysis that was done by Dr. Sullivan, approximately 6 out of 7 patients would be disappointed, and based on if they are responding to that indication alone.

CHAIRMAN DYKEWICZ: Any further discussion on dyspnea or the tools, or the instruments used to measure it? All right. Then organizationally I would like to call the question, which is actually numbered as four on our agenda.

And that is do the data provide substantial and convincing evidence that tiotropium bromide inhalation powder, and that provides a clinically meaningful effect for the symptom of dyspnea in patients with COPD.

And this will be a yes or no answer format, and what I will do is take a poll of the members of the committee, and then at the end give an opportunity for any qualifications or final comments that individual members of the committee may have about the question. Dr. Kennedy. Okay. He doesn't

vote. Dr. Schatz. He doesn't vote. Dr. Patrick.

DR. PATRICK: No.

CHAIRMAN DYKEWICZ: Dr. Parsons.

DR. PARSONS: No.

CHAIRMAN DYKEWICZ: Dr. Atkinson.

DR. ATKINSON: No.

CHAIRMAN DYKEWICZ: Dr. Morris.

DR. MORRIS: No.

CHAIRMAN DYKEWICZ: Dr. Joad.

DR. JOAD: No.

CHAIRMAN DYKEWICZ: Dr. Stoller.

DR. STOLLER: No.

CHAIRMAN DYKEWICZ: And I vote no. Dr. Swenson.

DR. SWENSON: No.

CHAIRMAN DYKEWICZ: Dr. Apter.

DR. APTER: No.

CHAIRMAN DYKEWICZ: Dr. Chinchilli.

DR. CHINCHILLI: No.

CHAIRMAN DYKEWICZ: Ms. Schell.

MS. SCHELL: No.

CHAIRMAN DYKEWICZ: All right. Now,

having made those votes, I would like to give the opportunity for you to make any additional comments, but along those lines, question five is really addressing what might be some additional comments, and this might help focus additional comments in general.

What quality and quantity of data would constitute substantial and convincing evidence of a clinically meaningful benefit for the symptom of dyspnea in patients with COPD. To put it another way, if a study sponsor were to approach the FDA for the indication of dyspnea, what sort of data, and what caliber of data would you like to see in order to justify that indication. Dr. Apter.

DR. APTER: In addition to the things that have already been mentioned, I wanted to reiterate that it would be validated in diverse populations, ethnically, and gender.

CHAIRMAN DYKEWICZ: Dr. Parsons.

DR. PARSONS: Actually, Dr. Joad's comments made me realize that probably one of the first things that will need to be done is to define dyspnea. I think actually I know realize now that I

have thought about it a little bit more, that we are probably all talking about something different as we sit around the table, and we have never truly defined what it is.

And actually if you look at the TDI, it is more of a change in ability to do things perhaps. So I would say that like a lot of things in medicine, we often times think we are all talking about the same thing, and I don't think we are.

So I would suggest that we come up with or that a definition be developed for what is truly being measured and looked at, because as a clinician, I know what my patients look like, but it is clearly different than pediatric, and it may be very different than others.

CHAIRMAN DYKEWICZ: Dr. Patrick.

DR. PATRICK: I would like to qualify, and I think they have gone a long way, and so the word substantial might be -- I might have said that they provide substantial evidence. To me it was not convincing, because I still am not sure how the instrument was used.

I would think that the data that would be useful would be a study specifically on the minimally important difference, and that it included a separation between the reports of the sensation, and the activities that product that report.

And that it be a combination of a patient report and a clinician interview, and that we would need specificity if it is going to be based on a clinician interview of exactly what was the baseline.

I am not at all convinced, although I know Dr. Jones very well, that it is big mistake in a condition like this to give people their baseline activities. I am not sure that we can do it any other way.

This is an age-old thing, and so I would say that we need a test of that as well. So I am just going to suggest the evidence for that.

CHAIRMAN DYKEWICZ: Dr. Morris and then Dr. Stoller.

DR. MORRIS: I think in an ideal world tying such a hard to understand concept of dyspnea is something a little bit more concrete would be useful.

Something that is objective, and something that is reproducible, and some activity of daily living and reproducible testing might be something that we would strive for.

And something that could be tested in Europe or in the United States, and with a certain amount of work being expanded and applying that then also the rating of a dyspnea scale.

But some more concrete aspect of the test, rather than the subjective language part of the test.

CHAIRMAN DYKEWICZ: Thank you. Dr. Stoller.

DR. STOLLER: I would again preface my remarks by saying that as has been pointed out, there is no dyspnea meter. There is no gold standard so that the rigorous attempt to define this really uses the functional aspects of the symptom.

That said, these instruments about which we have heard much I think represent tremendous methodologic advances in our ability to place confidence in the measurement of clinically important outcomes.

Having said that, the kind of information that would be important to me to persuade me that dyspnea was a reliable and credible outcome measure in a clinical trial would be largely to address the methodologic issues that Dr. Kammerman summarized.

I have as I said before, I have belief in the clinical sensibility of actually several of the measures we have heard about through really the vigorous work of those who have discussed them; Dr. Jones, Dr. Mahler, and I am comfortable with either of those if ideally administered.

I would hope that there would be greater attention to the defining of the minimally important clinical difference. I agree with the comments made about demonstrating the reproducibility in different populations as we apply these drugs to populations other than those of the narrow clinical context of clinical trials, because in clinical practice that matters.

And one would need to know the conclusions around dyspnea and outcome measures are generalizeable, but I think most importantly my

reservations have to do with the methodologic shortcomings of applying the outcome measure in a kind of after the fact.

And that a rigorously designed prospective study in which attention to some of these methodologic details about the training, reproducibility, translatability, generalizeability of the measure, in the context of a reasonable demonstrated, minimally important difference, would certainly convince me of the utility of these measures as an indication for a drug.

CHAIRMAN DYKEWICZ: Dr. Schatz.

DR. SCHATZ: And as I alluded to before, in addition to all of this, it would seem to me that being very concerned about recall issues, that I would be in favor of seeing the longitudinal properties of the BDI in this validation process.

That it would be the BDI that would be done over time, and compared with other relevant clinical parameters.

CHAIRMAN DYKEWICZ: My own additional comments other than what have already been said is that I think you would want to have an instrument that

is the patient reported symptoms of dyspnea. We know that in other disease states that when there has been a physician assessment of patient improvement, compared to improvement in patient symptom scores, there can be some discordance.

And I think as much as possible should go right to the source, the patient, and if possible devise a questionnaire that asks them directly without filtering, even though it might be a learned intermediary, but without filtering, ask the patient symptoms that could be used to support whether there is actually an improvement in their symptoms. Dr. Patrick.

DR. PATRICK: I might add on to that, and that one of the reasons that the interviewer form is important is because is because of missing data, and therefore, the data is highly unlikely to have been missing at random, and so there needs to be an addition, an investigation of either the surrogate endpoint from the clinician, as well as different methods for imputation, in addition to last observation carried forward in the data analysis.

CHAIRMAN DYKEWICZ: Dr. Joad.

DR. JOAD: Well, if I am understanding you right, it seems like the crux of the word dyspnea is how much work can you do before you become dyspneic, and so it seems to me that it would be nice to validate it with one of those kind of tests like they discussed.

You do a certain amount of work, and then at which point along that work do you become -- and that is like a real test in a real laboratory. And then take the history, and see how valid and how.

CHAIRMAN DYKEWICZ: Thank you. Dr. Chowdhury.

DR. CHOWDHURY: Ye. Just a clarifying question on number five. We had in our discussions quite a bit of input regarding the quality of data, and I just wanted to also emphasize on the word quantity, and perhaps have a brief discussion on that.

And whether TDI itself is enough for one to (inaudible) indication or would somebody want some other measures to go along with it. I would like to have some input on that regard.

CHAIRMAN DYKEWICZ: Well, I will give you my response Echoing what I said just a few minutes ago, I think you would want to have both the TDI and another bit of data, which is directly getting reports of symptoms from the patient.

And I think there has to be a pairing of those two really for optimal assessment of that. Dr. Patrick.

DR. PATRICK: This is just an addendum to that. You also want the patient's global rating of the change that they have experienced, and whether it is minimally important to them.

CHAIRMAN DYKEWICZ: Dr. Stoller and Dr. Swenson.

DR. STOLLER: I would say that actually many of the data elements that would convince me of the efficacy of a drug that we have heard about in the context of this study. That to fantasize, that were we to have been shown a study that would have rigorously captured BDI and TDI data, SGRQ data, pulmonary function tests, physician global assessment, patient assessment, and that there were convincing

evidence that those -- you know, moved in a concordant direction, that would provide a weight of information from my point of view that would bolster and buttress the idea that these important measures would measure different things.

And I should emphasize in response to comments that these explicitly should and do measure different things. That the notion that we should validate a subjective instrument on a single physiologic measure is to my thinking clinically naïve, as it ignores the richness of clinical material, and clinical experience that forms patient's symptoms, and what brings them to our attention.

So if we really wanted to know what the VO2 max is, we should suspend interest in these clinically symptomatic measures, and simply measure VO2 max. We are explicitly interested in as clinicians, I believe, the richer experience of patients as they experience their illnesses, and these functional status measures in different dimensions, although there is some co-linearity of some of these measures as I think we have heard, they are designed

to capture that.

What is missing is the convincing evidence that they were captured in a way that would be -- you know, to say that I am not convinced is to not say that there is evidence that they don't improve dyspnea.

It's just that given the dataset, I am not convinced that these data as presented to us do that. In the ideal, these data elements, should the results be concordant in the way that I have described it, would persuade me if I am in a methodologically perfect way.

And if we could satisfy the premises that Dr. Meyer put on the table before, and if they were ideally captured in all of the close scrutiny about the methodology was addressed, I would find this quantity of data persuasive in my view.

CHAIRMAN DYKEWICZ: Dr. Meyer, did you want to comment specifically?

DR. MEYER: Actually, I wanted to ask Dr. Stoller just a follow-up for clarification of his points. With regard to the dataset that we saw today,

realizing that they are in fact measuring different things, what do you make of the fact that no effect was seen on something like a shuttle walk test, when you have an effect apparently on the TDI?

DR. STOLLER: You know, you bring up the issue of concordance and correlation between measures, and frankly I would ask or could ask the same question of what is a minimally important difference in a shuttle walk test, you know, that would define a basement threshold for what is important.

And I as a clinician would be much more content to accept someone's consistent reporting that they felt better could do more than if they could walk 10 meters further. I have this difficulty with six minute walk measures as outcome measures in studies that we read about different pulmonary illnesses, pulmonary hypertension among them.

So I am not bothered by some discordance in terms of the individual measures. If the weight of the evidence suggests that there is a general trend among multiple measures that are indirectly measuring similar, but not identical things, I would find that

persuasive.

And I am not sure that one could ever be more precise about -- you know, when one gets into the arena of if you are going to measure functional outcomes, one has to live with this non-complete concordance of measures, and I am personally comfortable with that.

CHAIRMAN DYKEWICZ: Dr. Swenson.

DR. SWENSON: Well, just to repeat this thought that what is ultimately most important is the patient and his or her evaluation of the effect of the drug.

And that's why I raised this point, that possibly at the very end it should be brought back to the patient as to is this meaningful to you, and could we come up with some way to pose that to the patient and only in a theoretical sense.

I don't mean to say that we should be advocating certain percentages of income or whatever to the cost of drugs, but to place it in a theoretical perspective. How much value is this to you, and would you be willing to sacrifice for this.

If some tool of that caliber could come up, I would be then willing to accept a value of one as being meaningful. Right now, we are still floating about is one a number that anybody can really hang their hat on.

CHAIRMAN DYKEWICZ: Thank you. Any other comments? Dr. Kennedy.

DR. KENNEDY: Thank you. I am sitting here thinking that there is probably a number of people in the audience and folks who are listening in, who are now planning dyspnea studies, and we are talking about it like it just fell out of the sky. That it was the first study that was ever done.

When in fact the Mahler paper was prepared because there was a need, and it has been in place for a while. And I am sure that every pulmonary drug that has been submitted to the agency in the last 15 years probably has some measure of this.

The question that I would pose to you as a committee is to give some consideration. If there are data on-hand now within the FDA that is able to measure these changes more discreetly than the one

unit change, and found out that all of the data in place were .3, this change of a full unit today would be something significant.

So what I am asking you to do is don't try and define the world on the basis of this one or two studies that you have seen today, but ask our colleagues at the FDA to help provide the industry with some input on all of the stuff that has been done up to this point.

CHAIRMAN DYKEWICZ: Thank you. Dr. Sullivan.

DR. SULLIVAN: Just to partially address that, is that in fact in these studies there was an active control of an approved drug for COPD, and the data was presented regarding how they responded on that end point.

So there is some information about how other drugs out there behave with this instrument.

CHAIRMAN DYKEWICZ: Any final comments on dyspnea or its assessment? Since we are talking about efficacy, let's go to what is numbered as Question Number 3 about bronchodilator effect of the drug, and

question 3 is do the data provide substantial and convincing evidence that tiotropium bromide inhalation powder provides a clinically meaningful bronchodilator effect when used in the chronic treatment of patients with COPD.

First, let's open this up to discussion. Well, I will say that I think that it has been established, and I don't know if anyone would take issue with that. Dr. Stoller.

DR. STOLLER: I would like to make one other point, and that is that once of the novel aspects as I think has been pointed out of this particular outcome measure is the trough or nadir level prior to dose.

And actually I applaud that as an end point, because although it has been less well filled out, in terms of being unconventional, and therefore not having the matrix of the magnitude of effect, it is from a clinical point of view, I think as has been pointed out, far more meaningful than a transcent peak effect.

And now it is convincing and reassuring to

me to know that in fact the peak and the trough outcome variables are the same with regard to the data that we have seen. But I think the notion of looking at trough effect, particularly in a long acting drug such as this, is a laudable and significant advance in the assessment of drug efficacy. So I would say that as a baseline.

CHAIRMAN DYKEWICZ: Thank you. Dr. Parsons.

DR. PARSONS: I just have a question. I totally agree that they have shown a significant bronchodilator effect, and I actually like the trough data as well.

In future studies will trough data be adequate? If we had not seen the greater than 200cc change in the acute, would we know what to do with the specific trough number?

DR. STOLLER: I think it gets to the issue what is the primary and secondary outcome measure. As a primary outcome measure, as is indicated here, I would favor the trough, but I would like you be absolutely very interested in looking at the

pharmacokinetics and the physiologic response over time, which I think we have been shown.

And so the answer is that if I were on a desert island, and had to pick one outcome measure, and would that suffice, I would say no. But of course in the richness of clinical investigation, we are often given a fuller dataset.

Now, if you were to ask me if the trough data were good, but there were no significant rise in the peak, what would I do with that, and I guess I would have to think about it. But I would actually find that more reassuring clinically to find the trough data over time sustained than even lower peaks.

So the answer I guess would be, yes, if I had to pick, in terms of primary outcome measures, I would favor the trough as was done here, and so I actually applaud that.

DR. PARSONS: Actually, my question was in terms of the agency, if they came back and said what is an appropriate trough change, and so we know the peak change is 200, and that is the number that we are using.

Do we now have a number for the change in trough level that we use?

CHAIRMAN DYKEWICZ: I don't see anyone volunteering.

DR. SHOLLER: Well, it gets to how the peak data were derived. I mean, in fairness, the 12 percent and 200 ml with people sitting around in the ATS spirometry statement saying what we in the FVC lab define as a significant BD response.

Prior to that it was 15 percent without an absolute volume. So I am not sure that 140 or 110 ml increment in a baseline population if the mean you want is 1.04 to 1.2 liters, is any less convincing than an arbitrarily embraced -- and furthermore, just to get to the arbitrariness of it, it is often accepted in the November 1995 ATS document.

It is often accepted actually as an outcome. I think I should correct that and I think it is in the spirometry statement and not in the November '95 COPD statement.

It is often accepted as an outcome measure for FVC, and yet we obviously understand that patients

with COPDs, and FVCs and not FEV6s are highly sensitive to exploratory time. So that it is not uncommon to see in the lab two successive blows; one at 12 seconds and one at 8 seconds.

There is obviously a 12 percent and 200 ml difference, which on that criterion would satisfy bronchodilator responsiveness, but is in fact not. It is simply related to the artifact of different durations of exploration, knowing that these patients can blow out for 15 or 20 seconds and still continue to exhale gas.

So my comments simply address the relative arbitrariness of the 12 percent and 200 ml. I personally find -- and to answer your question in regard to the data with regard to the magnitude of trough effect as building up, and an assembly of data that says this is a clinically significant trough effect of a long acting drug, yes.

CHAIRMAN DYKEWICZ: Dr. Parsons.

DR. PARSONS: I agree with you a hundred percent. I just am thinking that six months from now when the next drug comes in, are we going with 140 or

is a hundred okay? If I was in the audience, I would probably want to know if there is any recommendations.

And I have no clue as to what he right answer should be. I mean, I think the dataset shown today, in an aggregate, are convincing, but if you ask me specifically would the trough level alone be enough, I would say, boy, I have no guidance because I don't really know what the number should be, even based on the good people at the ATS telling me, but they haven't told me yet.

CHAIRMAN DYKEWICZ: Dr. Schatz.

DR. SCHATZ: What seems to me is that maybe you have answered your own question, which is that we can't do it as a single measure, and that we really have to take each case individually, and look at the aggregate.

But I would also say that knowing how this was part of an aggregate would help me be more comfortable with something like 120cc's in a future study, but I still would feel that I don't think we can answer your question right now.

CHAIRMAN DYKEWICZ: As good a question as

it was. Is there any further discussion? Ms. Schell.

MS. SCHELL: Yes. I just would like to add that I am excited that the dosing, the dosing on compliance on the patients that I care for. It is very difficult to take them out of medications now and for one dosing to get this result. It is exciting for me to see that, and I think it is a plus.

CHAIRMAN DYKEWICZ: Any further discussion on the bronchodilator effect? Then we will call for a formal vote. Again, do the data provide substantial and convincing evidence that tiotropium bromide inhalation powder provides a clinically meaningful bronchodilator effect when used in the chronic treatment of patients with COPD? Dr. Patrick.

DR. PATRICK: Yes.

CHAIRMAN DYKEWICZ: Dr. Parsons.

DR. PARSONS: Yes.

CHAIRMAN DYKEWICZ: Dr. Atkinson.

DR. ATKINSON: Yes.

CHAIRMAN DYKEWICZ: Dr. Morris.

DR. MORRIS: Yes.

CHAIRMAN DYKEWICZ: Dr. Joad.

DR. JOAD: Yes.

CHAIRMAN DYKEWICZ: Dr. Stoller.

DR. STOLLER: Yes.

CHAIRMAN DYKEWICZ: Dykewicz votes yes. Dr. Swenson.

DR. SWENSON: Yes.

CHAIRMAN DYKEWICZ: Dr. Apter.

DR. APTER: Yes.

CHAIRMAN DYKEWICZ: Dr. Chinchilli.

DR. CHINCHILLI: Yes.

CHAIRMAN DYKEWICZ: And Ms. Schell.

MS. SCHELL: Yes.

CHAIRMAN DYKEWICZ: Thank you. We will now turn our discussion to side effect profiles, and concerns about that. Because of a number of different issues were raised, I would like to focus the committee on several subtopics.

First of all, the issue of if you will anticholinergic side effects, including dry mouth and some of the GI side effects. If we are looking at obviously a drug that will be used in clinical practice, what is your assessment about the risk

benefit profile. Dr Schatz.

DR. SCHATZ: One of the things that impressed me was as common, and as much more common as it was in the patients taking the drug, it was very uncommon for patients to discontinue it because of that. So that makes me much more comfortable with accepting those side effects.

CHAIRMAN DYKEWICZ: Dr. Joad.

DR. JOAD: Could the FDA remind me what -- when the word adequate is on this question, what are the options. The options are to study more patients, and the other option is to do Phase IV follow-up of some sort. Could you just remind us if you choose to approve a drug, even though you have some concerns about side effects, what are the options for following that in the future?

DR. MEYER: Well, there are in fact options. I mean, generally the cut that we make internally is are there any gaps in the safety knowledge substantive enough that you wouldn't want to approve it. That you don't know enough about the risk to benefit ratio to put it out there.

There will always be some gaps in our knowledge, because no matter how many -- and this is a fairly large program that Boehringer Ingelheim did, but no matter how many patients they study, it is not until you get into several million patients that you begin to understand some of the more subtle signals, because a large trial such as in a database such as this, may give us a reasonable chance of finding something with a one in a thousand occurrence rate.

But if it gets into millions of patients, you are going to see some more subtle signals. But in any case, given the fact that you may have gaps that would preclude approving it, and given the fact that in the best scenario that you will never have a good complete knowledge of the safety, there is middle ground where there might be nagging questions that don't preclude approval, but do warrant some phase four studies.

Commitments from the company to further allucidate some area. I am no sure whether I have answered your question.

DR. JOAD: Yes, you did really well. I

just have one more part of that. A lot of the side effects are these ones that you would expect to happen in this age group anyway. So they are not going to be an adverse -- if somebody dies of an MI, or somebody has a fecal impaction or something, or a urinary tract problem, they wouldn't -- it would not come in as an adverse drug report probably. But that would be part of Phase IV to pick up those.

DR. MEYER: Well, those might be situations where a specific study could be warranted, because if it is something that occurs commonly in the population, even if it comes in as an adverse event report, it may be difficult to interpret that, because we don't really have a firm denominator for those kind of post-approval data.

So those are situations where it is a potential that you would want a Phase IV study, a rigorous study.

DR. JOAN: Can I ask one more question? The groups that were excluded due to side effects for these studies, when this gets marketed, will they -- will the part in the package inserts say this, that

those same groups should not get this drug at all, or be careful, or --

DR. MEYER: I am not going to answer that question because it is actually the basis of our question, too, that we are putting to you. So I don't want to put an answer into anybody's mouth.

CHAIRMAN DYKEWICZ: Dr. Swenson.

DR. SWENSON: Well, let me move to one concern that I have, and that is the issue of the renal excretion of this drug. Although the total absorbed dose is low, we have seen that the drug levels are measureable over time, and we are seeing a slightly greater rate of complications on the basis of anticholinergic effects with Spiriva as compared to the standard in the field; that is, Atrovent.

So I would be worried that if the drug is used more widely, and people would compromise renal function, that what may look just like going over the top of a dose response curve, and possibly just leveling out, or does that represent really an important steep portion of a dose response, and that we would expect to see a lot more problems in people

with renal insufficiency?

CHAIRMAN DYKEWICZ: Other comments? Dr. Stoller.

DR. STOLLER: I guess my response to this has to do with several things. One is what in my own mind is the magnitude of the risk, and what is the magnitude of the clinical benefit, and what is the functional performance of patients in the dataset with regard to the study, and discontinuing the study drug.

And then the general reliability issue; are there subsets that were explicitly not included in these studies for which the pharmacologic properties would pose particular problems.

So I am imagining some potential patient subsets, for example, and patients with significant co-morbidities of both, for example, kidney and heart disease that were explicitly excluded from these studies, and that as we heard before, might pose potential risks for a drug.

A patient with a creatine of four, and triple vessel coronary artery disease, and who happens to have COPD, and that is not by any means an

impossible scenario. These are patients for whom either there might be some language, cautionary language around the generalized ability of these conclusions to that patient population with regard to safety.

Or alternately -- and I am not sure how this is done, but some attention to the specific performance of this drug. Now, admittedly, those are patients not in this set. I have no concerns about the safety profile of the drug as presented in the populations to us, because I think that dry mouth is something as was pointed out that patients are willing to tolerate for the sake of the clinical benefit that they appreciate.

And I think that in the study population as we have seen it, large as it is, there was ample evidence that these are tolerable, and not life threatening, and not serious, and certainly not sufficient to deny people the opportunity to use this drug.

I just think that perhaps some attention -- and I am not sure what specific recommendation to

make, but some attention to these excluded subsets, in which the pharmacologic profile might provide particular concerns. Not that we have seen that, but on a theoretical basis, might require some more attention.

CHAIRMAN DYKEWICZ: Dr. Joan.

DR. JOAD: Just to answer the gaps. I would say that the only gap that I would want in an ideal world is Holter monitors on more patients. So whether it is really indicated or needed at this point or not I think is the issue, but it would have been really nice to have had that.

CHAIRMAN DYKEWICZ: Dr. Morris.

DR. MORRIS: I think in a broader sense the data presented in my mind do show safety within the population study. I wish I had numbers to present to show what frequency of the COPD population that will be interacting with as a physician would represent the group that were excluded from this study.

And in administering this product would we be introducing a potentially life limiting event, and

would their life not have had that life limiting event, even though we are talking about a person who has probably severe COPD and heart disease, and are relatively hypoxemic.

So I think that to me is an unanswered question; is how safety can I bring it to my patient now, who might have significant underlying heart disease, as well as COPD. I feel confident in the data on who were studied, and that does not represent an untoward risk of cardiac events from what was presented.

But there is still, I think, a significant population who were not studied.

CHAIRMAN DYKEWICZ: Thank you. Dr. Atkinson.

DR. ATKINSON: I would like to add that probably there should be some attention in marketing if this drug were approved that would make the primary care doctors aware that this just isn't another long acting ipratropium, but that it does have systemic absorption, and really emphasized the fact that people with perhaps unrecognized prostatic hypertrophy, and

mild renal disease, there may be side effects that may be unanticipated that you wouldn't see with ipratropium.

CHAIRMAN DYKEWICZ: Other comments? I might just say as a personal comment that one thing that we are obviously dealing with is if we have some populations that have been excluded from study, and those are going to be populations in real life that are going to be treated with this drug, there does need at some point to be some study of that population.

So I think there needs to be, even if it is post-marketing, some study done on patients who have, let's say, coronary artery disease, and significant cardiac disease, to assure the safety of the drug.

On the other hand, we are looking at a drug that is -- although it is a new entity, it is an anticholinergic agent. We do have a good amount of experience with another anticholinergic agent, namely ipratropium.

So I think we probably already have some

sense of any signals if there would be because of that drug class a significant adverse effect on cardiac status. So with the idea again, and with the reservation that I would have preferred to have seen some data about the safety of this agent in patients who have cardiac disease, I have some reassurance that this is of a drug class that does have a good track record of experience in the patient subsets that were excluded from this study. Dr. Morris.

DR. MORRIS: Just to play the converse of that. I think one of the reasons why we see a trough effect and not with this item today, and not with the ipratropiums, because they are different, and that because of that difference, we have to say that the drugs are different.

And that the potential for unsteadied events that are realistic, and potentially harmful, are out there, but yet we have not studied that population.

CHAIRMAN DYKEWICZ: Dr. Schatz.

DR. SCHATZ: But I was reassured to hear about the theoretical aspects of anticholinergics and

electrophysiology of the heart. So that what I think you said, Mark, is still correct based on that information.

CHAIRMAN DYKEWICZ: Dr. Joad.

DR. JOAD: What I would like then if it would be Phase IV studies for everybody on the effects of the drug, and the effects on urinary obstruction, impaction, and arrythmias. And as you mentioned, I think for the groups that haven't been studied, either they should be told on the product label that they shouldn't get it or that there should be studies on them for safety.

CHAIRMAN DYKEWICZ: Dr. Parsons.

DR. PARSONS: The only other areas that came up that we really didn't discuss was the increased incidence of hyperglycemia and diabetes. It was a little out of control as it was not that well-defined.

So it is not clear to me how big a problem that is, and if these are people who really go into DKA. I don't think so. Or if they have transient hyperglycemia. It seems like there was an increased

incidence, although totally unexplained based on what we know about the drug.

It should at least be monitored in some fashion because as it stands, there are certainly a number of patients with COPD, especially those with heart disease, who do have concomitant diabetes, and that could be a potential problem.

CHAIRMAN DYKEWICZ: I second the recommendations of both Dr. Joad and Dr. Parsons.

CHAIRMAN DYKEWICZ: Other comments? Does the FDA have any additional questions they wanted to pose before we take a vote on the question? No? All right. Then let's call the formal question, Question Number 1. Again, a yes or no response.

Is the safety database for tiotropium bromide inhalation powder for the treatment of COPD patients adequate for approval. Dr. Patrick.

DR. PATRICK: Yes, on the basis of the Phase IV recommendation.

CHAIRMAN DYKEWICZ: Well, we have to have an answer though. It can't be qualified. It has to be yes or no. If you believe that the data that

currently exists is sufficient to approve the drug, or whether you would defer approval, in which case you would say no. You would say no?

DR. PATRICK: No. Yes. Yes.

CHAIRMAN DYKEWICZ: You would say yes?

DR. PATRICK: Yes.

CHAIRMAN DYKEWICZ: Okay. Dr. Parsons.

DR. PARSONS: Yes.

CHAIRMAN DYKEWICZ: Dr. Atkinson.

DR. ATKINSON: Yes.

CHAIRMAN DYKEWICZ: Dr. Morris.

DR. MORRIS: No.

CHAIRMAN DYKEWICZ: Dr. Joad.

DR. JOAD: So if I don't know that there is going to be a Phase IV, I have to say no.

CHAIRMAN DYKEWICZ: All right. Dr. Stoller.

DR. STOLLER: Yes.

CHAIRMAN DYKEWICZ: Dykewicz, yes. Dr. Swenson.

DR. SWENSON: No.

CHAIRMAN DYKEWICZ: Dr. Apter.

DR. APTER: Yes.

CHAIRMAN DYKEWICZ: Dr. Chinchilli.

DR. CHINCHILLI: Yes.

CHAIRMAN DYKEWICZ: Ms. Schell.

MS. SCHELL: Yes.

CHAIRMAN DYKEWICZ: All right. Thank you. Now, just to clarify, the additional safety data that should be obtained, I think that we have already had some good discussion of that, but to kind of give a final opportunity to members of the committee, should the drug be approved, are there any additional Phase IV studies that you would like to see in different populations?

I would say one other thought would be looking at different demographic groups, in terms of African-Americans, Asian patients, and I think that would be important.

DR. APTER: Women, too.

CHAIRMAN DYKEWICZ: Women, yes. Dr. Chowdhury.

DR. CHOWDHURY: Just a comment. We had three notes here, and I was wondering if you are going

to ask the question what exactly would they want in terms of safety data prior to approval.

CHAIRMAN DYKEWICZ: People who voted no. Dr. Morris.

DR. MORRIS: I believe addressing documentation in patients with suspected heart disease or documented heart disease, dysrrythmias, that there is no increased dysrrythmia activity and/or deaths.

CHAIRMAN DYKEWICZ: Thank you. Dr. Joad.

DR. JOAD: Just what I spoke to. I think it could be approved now with prohibition of the groups who were excluded from the prior -- in the package insert to say they should not get this drug, the group that had heart failure in the last three years, or has arrhythmias, on medication, and have BPH, and to say that those people cannot have it now.

And then have a Phase IV to say that they can have it, or can't, and then also to follow long term the safety concerns, which I think are substantial given that it will be given to a lot of people, and it has a very long elimination half-life. I think you have to be very careful with this drug.

CHAIRMAN DYKEWICZ: Dr. Swenson.

DR. SWENSON: I think that the issue of renal insufficiency is important enough that this should be followed in Phase IV very closely. I think what we saw with the atrovent versus the tiotropium suggests about a two-fold increase in all of these anticholinergic potential problems, and therefore I don't know where we exist on the relationship between blood levels and these side effects.

I don't know whether we peaked out or whether we are on a steep dope response portion. So I think that issue should be followed closely. We certainly have -- this is an elderly group of patients by and large.

They get many drugs that affect renal function, and so they may start with normal renal function, but put on a drug such as a non-steroidal anti-inflammatory agent, or something of that nature, and their renal function will change. So I would be worried about that.

CHAIRMAN DYKEWICZ: Dr. Stoller.

DR. STOLLER: I would submit that the

Phase IV monitoring should in fact address all of the subsets not included in the dataset that we have seen, and that it also should address the specific concerns, albeit small, raised by the data that we have seen.

And in particular I would say that there ought to be monitoring with regard to women and non-caucasian groups, since those are not amply represented in the dataset that we have seen.

And that in addition the Phase IV monitoring should address some of the issues raised. As Dr. Parsons said, diabetes, and combinations of co-morbidities not represented here, particularly coronary artery disease. I am less concerned about arrhythmia based on the convincing data that we have heard.

But I am concerned about coronotropic effects in patients for whom that may be a significant concern, particularly coronary patients with significant coronary artery disease, recent MI and concomitant renal disfunction; as well as patients with known BPH, all of whom were excluded from these datasets, but for whom in clinical practice this might

pose morbidity not otherwise appreciated by the data that we have seen.

So Phase IV monitoring should be broad in its scope, but focused on these specific subsets in my view.

CHAIRMAN DYKEWICZ: Dr. Chowdhury, any additional questions to the committee

DR. CHOWDHURY: No.

CHAIRMAN DYKEWICZ: Fine. With that, we will adjourn, but did you have any final comments from the FDA?

DR. CHOWDHURY: Yes. I would first like to thank you for your participation and a thank you to the committee for their participation in this meeting. We really appreciate the time and effort that you have put into meeting.

CHAIRMAN DYKEWICZ: And I would like to add my personal thanks and wish everyone a safe trip home.

DR. CHOWDHURY: Just a couple of more small points that I want to make. Here as I said before in my opening statement, we would take this

into consideration from the clinical standpoint. However, we did not ask an overall approvability question in this meeting.

However, based on the questions that we had posed, what we have heard is all votes in favor, in terms of safety. I take that back. In terms of efficacy, and in terms of safety, the majority was again in favor for a yes.

So overall what we hear is a strong recommendation in favor of approving the drug from a clinical standpoint. I just wanted to reiterate that.

CHAIRMAN DYKEWICZ: I believe that is the overall consensus of the committee. Thank you very much.

DR. CHOWDHURY: Thank you very much, and have a safe trip back.

(Whereupon, at 2:59 p.m., the committee meeting was concluded.)