regression approach. All of those imputations, every single one, including a request by the Agency to do a PROC MI, in which we used the propensity score option under PROC MI, every one confirmed results, and that is there was a strong, statistically significant difference in FEV_1 and six-minute walk at six months within all of those constructs.

So to exhaust the sensitivity analysis realm, I think we've done that. I think we've done a tremendous amount of work in that area.

To address your question about the similarities of the treatment group to the missing patients, we did an analysis of baseline characteristics of patients who completed the study and those who were missing. And, in fact, there were no statistically significant differences in any parameter.

DR. DOMINIK: Are those available for us to take a look at? I mean I'm not just interested in statistical significance, but what is the difference in characteristics?

DR. CHIACCHIERINI: We can develop a table or a set of tables that we can show you, yes.

DR. BIRNBACH: Dr. Halabi.

DR. HALABI: Well, Dr. Dominik has really

covered all my issues. Thank you.

DR. BIRNBACH: Dr. Domino.

DR. DOMINO: I was curious about the materials in this device. You had mentioned silicone and nickel, and I was just curious about whether there were allergic reactions to nickel in particular and what the prevalence of allergies to nickel is in people, anything particular about the device that might be problematic.

DR. ERNST: In our study, we did not see anybody with any allergic reactions throughout, nor to my knowledge are there any reported allergic reactions outside the United States in the use.

Just to clarify, this is not nickel, but nickel titanium alloy, which is really very frequently used in medical devices like stents, for example. They're frequently made out of that and, you know, silicone also in the airways is a very frequently used material, for example, silicone stent.

DR. DOMINO: My other question deals with whether this research has been published and, if not, has it been submitted to journals and been rejected? And if it has been, what are some of the comments from peer review?

DR. SCIURBA: This has been presented at national meetings. It is in its final stages of submission. It has not been submitted. The trial design data was published in an online -- BMC online but the results of this trial have not been finally submitted.

DR. DOMINO: And then the last question --

DR. SCIURBA: It has never been rejected.

DR. DOMINO: The last question deals with long-term follow-up and preclinical data. I mean you don't have long-term follow-up in your patient population, but are there animal species who you might have caused similar type of, you know, chronic obstructive pulmonary disease, put these devices in and then followed for a number of years that might be a number of years in a person's life span?

DR. SCIURBA: So with regards to the animals, there's been no studies beyond a year specifically with respect to that. But also in -- no long-term catastrophes in patients that are outside of one-year windows, but the trial again was prespecified to go one year duration, and so it was not in the purview for rigorous data collection beyond that point.

DR. BIRNBACH: Okay. It is now 10:10, and

we will take a 10-minute break. We will resume at 10:20. I would like to remind the Panel members that there should be no discussion of the PMA during the break amongst yourselves, with the Sponsor, the FDA, or with the public.

DR. CHIACCHIERINI: Mr. Chairman, may I get a clarification from Dr. Dominik before we go off and run the tables?

DR. BIRNBACH: I would rather that happen when we come back if that's okay.

(Off the record.)

(On the record.)

DR. BIRNBACH: If we could all take our seats please.

We will now hear the FDA's presentation. The first FDA presenter is Melanie Choe, Ph.D., the review team leader for this PMA. Dr. Choe.

DR. CHOE: Good morning. I'd like to welcome everyone to the FDA's presentation of the Anesthesiology and Respiratory Therapy Devices Panel.

The presentation today will focus on the premarket approval application, P070025, for a first of a kind device, the Emphasys Zephyr Endobronchial Valve System.

My name is Melanie Choe, and I'm the lead

reviewer for this PMA application.

The following topics will be discussed today in the presentation. I will provide a device description followed by a brief introduction to the clinical study along with the preclinical evaluation status for the device. The statistical evaluation will be presented by Mr. Alvin Van Orden, followed by clinical evaluation by Dr. Douglas Shure and post-approval study proposals by Dr. Jiping Chen if the device is determined to be approvable. Concluding FDA's presentation will be our specific questions to the Panel which will be discussed in the afternoon session.

As the Sponsor has already presented the device description, I will just briefly mention that the EBV system is a sterile single-use system consisting of three components, the Zephyr Endobronchial Valve for implant in the bronchial lumen. This valve is packaged within a loader system for compressed loading into the housing of the distal tip of a delivery catheter. The valve is then delivered in the bronchial lumen via a bronchoscope as shown in pictures 2 and 3.

Once implanted in the bronchial lumen, the one-way value is intended to prevent airflow into the

hyperinflated regions of the lung distal to the valve while allowing airflow out of the hyperinflated region

The Zephyr EBV system is a first of a kind Class III device. In order to gain premarket approval, the Sponsor must provide data that clearly demonstrates reasonable assurance of safety and effectiveness in determining the safety and effectiveness of a premarket approval device. The following relevant factors must be considered: the patient population for which the device is intended; the conditions for use for the device as suggested in the labeling or advertisement of the device; whether or not the probable benefit of the device outweighs the probable harm it may cause; and the reliability of the device.

As previously stated, the device is intended to improve forced expiratory volume in one second and six-minute walk test distance in patients with severe heterogeneous emphysema who have received optimal medical management.

The U.S. clinical study was conducted under the following investigational device exemption. The pivotal study was referred to as the endobronchial Valve for Emphysema Palliation Trial or VENT. It was

an unblinded, prospective, randomized, multicenter trial of the Zephyr EBV treatment group compared to optimal medical management control.

In total, the Sponsor enrolled 321 subjects into the VENT trial at 31 investigational sites in the U.S. between December of 2004 to April of 2006, using the Zephyr EBV system, which represents the main dataset in support of this PMA application.

Two co-primary endpoints, mean percent change of FEV_1 and six-minute walk from baseline to six months, were to be tested using a one-sided superiority test with a significance level .025. The primary safety endpoint was the major complication composite through 6 and 12 months, which included a variety of respiratory-related events that the FDA clinical reviewer will present more in detail.

At the time of the pivotal study approval, the Sponsor proposed a 30 percent MCC delta between the treatment and control groups. However, this was not agreed upon by the FDA. The Sponsor was informed at the time of the pivotal study conditional approval in April of 2003 that FDA intended to evaluate the complication rates for the Zephyr EBV and the control groups based on demonstration of benefit.

The Sponsor also conducted in vitro and --

performance and characterization studies of the Zephyr EBV system. Animal tests were conducted in sheep to assess the ease of EBV delivery, ease of removal, migration resistance, inversion resistance and atelectasis, which were all determined to be satisfactory.

Several engineering questions remain regarding the Zephyr EBV fatigue test that FDA is working interactively with the Sponsor to resolve.

Test results demonstrated that the device is compliant with FDA recognized international standards for biocompatibility.

Packaging and sterilization processes were also validated according to FDA recognized international standards.

Due to the complexity of the device, a wide variety of specialists were consulted to review this application. The FDA review team consisted of clinicians, statisticians, engineers and biologists from different offices, and their names are listed here for their recognition.

I would now like to introduce Mr. Alvin Van Orden to begin our statistical presentation.

MR. VAN ORDEN: Good morning. My name is Alvin Van Orden, and I will be presenting a

statistical review of the VENT clinical trial.

I will discuss the following topics, study design, subject accountability and protocol violations, primary and secondary effectiveness results, statistical significance and the estimation of the treatment effect, additional analyses, safety results, European data, and a summary.

Both the control and treatment groups received optimal medical management, which is the standard of care. Multiple endobronchial valves were placed in the target lobe of the treatment patients. The patients and investigators were not blinded to the treatment received.

Patients were randomized in the two to one fashion, treatment to control, and the randomization was stratified by target lobe and exercise capacity.

To co-primary endpoints were chosen representing physiological and functional assessment. The physiological co-primary endpoint is the percent change from baseline in FEV_1 at 180 days. The prespecified window for the 6-month visit was plus or minus 14 days.

The functional co-primary endpoint is the percent change from baseline in the 6-minute walk test at 180 days.

The primary safety endpoint was Major Complications Composite, or MCC, which combines important major complications.

For the study to be successful, the Sponsor needed to reach both effectiveness endpoints, and the clinical benefit needs to outweigh the safety risks.

The Sponsor prespecified three effectiveness analysis populations and one safety population.

The Intention to Treat, or ITT, was prespecified as the primary effectiveness analysis population. The ITT population as defined by the Sponsor includes all patients that were randomized. Analysis for this population was done using an agreed-upon multiple imputation method to predict what we would have observed in the missing patients if they had come in for a visit.

The Completed Cases, or CC, population includes all patients that came in for a visit. The number of Completed Cases listed here is as reported by the Sponsor, but not all patients that came in for a visit performed all of the scheduled tests. Thus, the actual sample size for the Completed Cases is smaller than the numbers given here and varies from endpoint to endpoint. Patients that died, for

example, were not included in any of the Completed Cases analyses.

The Per Protocol, or PP, population includes all patients that came in for a visit within the extended window and did not have any major protocol violations.

The primary safety population, which the Sponsor named the Modified Intent To Treat, or mITT, includes all patients that returned for at least one visit post-randomization.

The following changes were made to the statistical analysis plan after the last patient had been enrolled in the study for six months.

The European arm of the study, which had an identical protocol to the U.S. arm of the study, was not pooled.

The original list of nine secondary endpoints was changed to four, quality of life measure, St. George's Respiratory Questionnaire (SGRQ), the modified Medical Research Council which measures dyspnea, a measure of the exercise capacity, cycle ergometry, and the amount of supplemental oxygen used by subjects.

The other major change in the study design was the creation of an extended window, which for the

primary endpoint changed the window from plus or minus 14 days to minus 30 to plus 45 days.

101 control subjects and 220 EBV subjects were enrolled in the study, and this table shows the subject accountability.

Only about 60 percent of patient came in for a visit within the prespecified 14-day window.

Patients seen within the post-hoc extended window were not treated as missing in any analysis.

Still, over 20 percent of patients did not have a 6-month visit within this extended window. Some of these patients died, others formally withdrew from the study, and others returned after the extended window or not at all.

Note that the control group had a higher percentage of patients that withdrew or never returned for a visit.

The Sponsor reported 2,492 protocol violations. The Sponsor determined that about 70 of these violations were clinically important violations. Sixty-two patients did not meet the inclusion/exclusion criteria and should not have entered the study. Nine patients took medicines that should have excluded them from the study.

Between the two violations, 49 patients in

the EBV group and of the control patients were excluded from the Per Protocol population due to clinically important violations.

In the primary effectiveness analysis, the difference between the treatment and control groups in average change from baseline was estimated as 6.8 percent for FEV_1 and 5.8 percent for the 6-minute walk test. Both of these differences achieved statistical significance as the one-sided p-values are less than 0.025.

This table shows the estimates of the difference between the treatment and control groups in all three effectiveness populations. The size of the treatment effect changes very little across the different effectiveness populations, though the six-minute walk test is not statistically significant in the Per Protocol population, partly due to a smaller sample size and partly due to a smaller treatment effect.

At 12 months, the difference between the treatment and control in FEV_1 remained constant or increased from the 6-month difference, and the difference in the 6-minute walk test decreased and is not statistically significant in any of the patient populations.

This table presents the results for the secondary endpoints at six months as reported by the Sponsor. The Sponsor prespecified using the Hochberg adjustment for multiplicity, which looks at the secondary endpoints in a hierarchical fashion. If the largest p-value of the 4 secondary endpoints is less than 0.025, then no adjustment is made, but if the largest p-value is above 0.025, then the second largest p-value is compared to 0.025 divided by 2 and so on.

In the Sponsor's analysis, the four secondary endpoints ultimately chosen all had one-sided p-values less than 0.025 in the ITT population. For supplemental oxygen, statistical significance, as was determined by the Sponsor, seems to contradict the confidence interval, also calculated by the Sponsor, because the confidence interval includes zero or no difference between the treatment and control. If supplemental oxygen is not significant in the ITT population, then after using Hochberg's adjustment, none of the secondary endpoints would be statistically significant in the ITT population.

The differences between the treatment and control were fairly consistent across the patient populations, though most of the secondary endpoints

were not statistically significant in the Completed Cases and Per Protocol populations after applying the Hochberg's adjustment for multiplicity.

Statistical significance was achieved for both primary endpoints in the primary ITT population and in the Completed Cases population.

Statistical significance does not imply clinical significance as any size difference could be judged to be statistically significant if the sample size is large enough.

The primary endpoints should achieve both statistical and clinical significance, and the estimated treatment effect must be large enough to justify the associated risks.

Statistical significance was achieved for the four secondary endpoints in the ITT population after Hochberg's adjustment for multiplicity assuming supplemental oxygen is significant.

If the same multiplicity adjustment had been made for the nine original secondary endpoints, none of the secondary endpoints would have been statistically significant.

In the Per Protocol population, the six-minute walk test and all secondary endpoints are not statistically significant.

In the six-minute walk test, the four secondary endpoints are not statistically significant at 3 or 12 months in any population.

There are four important factors that may impact the estimation of the treatment effect.

First, lack of blinding. Because the patients knew if they were in the treatment or control group, they may be susceptible to the placebo effect. Also, the unblinded investigators may unintentionally exhibit treatment or assessment bias.

Second, the post-hoc extension of the window, it may not be appropriate to treat the 16 percent of patients seen in the extended window the same as patients seen within the prespecified window. The results may be biased due to the post-hoc definition of the extended window.

The third factor is missing data. Over 20 percent of patients did not have observed 6-month outcomes in the extended window. The underlying assumption in the imputation of missing data is that missing patients would have had similar results to those patients whose results were actually observed. This assumption, also known as the missing at random assumption, is unverifiable.

Fourth, protocol violations, about 21

percent of patients had clinically important protocol violations. In both primary endpoints, inclusion of these patients increases the size of the difference between the treatment and control. This is evidenced by the fact that the Completed Cases population shows a greater treatment effect than the Per Protocol population.

In what was prespecified as an additional analysis, the Sponsor also presented a responder analysis in which any patient that showed a 15 percent improvement in one of the primary endpoints was called a responder. The proportion of responders was then compared for both primary endpoints. No ITT responder analysis was performed.

The Sponsor shows a statistically significant increase in responders for the FEV_1 endpoint in the Completed Cases population, but the six-minute walk test was not statistically significant in any effectiveness population.

In this FDA responder analysis, a responder is defined as showing a 15 percent improvement on both co-primary endpoints. There is a higher proportion of responders in the treatment group, but the difference is not statistically significant.

The Sponsor also performed analyses on the

additional variables seen here; residual volume, diffusion capacity, and quality of well-being were all secondary endpoints in the original protocol. The BODE index was never prespecified as anything other than an additional analysis. These p-values have not been adjusted for multiplicity. So claims of statistical significance are unverifiable. Note that even for the BODE index, the median change from base for both groups was zero.

In a prespecified fashion, the Sponsor screened over 40 variables and their interactions in order to find a subgroup of the population in which the device might show greater effectiveness. It is not intended that the Panel be able to read all of the variable names, only that you see the number of variables that were screened.

Among the variables screened were four different continuous variables all measuring heterogeneity in different ways. Both ipsilateral and thorax DS heterogeneity were measured at total lung capacity and residual volume.

The Sponsor has defined a subgroup using the variable baseline ipsilateral DS heterogeneity at total lung capacity, as this variable appeared to be related to increased effectiveness in both primary

endpoints at six months.

The complete fissure subgroup did not show a significant improvement in both primary endpoints.

A sufficient adjustment for multiplicity was not made for the testing of these 40 variables and interactions. So the claim of statistical significance is unverifiable.

Additionally, the definition of this subgroup is unclear as the cutoff used by the Sponsor to define high versus low heterogeneity has changed throughout the course of the review process from greater than 10 percent to greater than 15 percent.

Finally, the Sponsor performed an analysis that showed there was a statistically significant relationship between high heterogeneity treatment interaction and death and LVRS, meaning that in the high heterogeneity subgroup, treatment patients may be more likely to die or have LVRS than patients in the control group.

In the Sponsor's presentation, they have shown significantly better results in each of the two primary endpoints in this high heterogeneity subgroup.

However, in the FDA responder analysis, that looks at both primary endpoints together, the

high heterogeneity subgroup does not show significantly higher response rates for the treatment group, and in this analysis, the subgroup does not appear to be much better than the larger study population.

While the differences are not statistically significant, it should be noted that the percentage of responders was about three to five times higher in the treatment group than in the control.

This is a post-hoc analysis, and the study was not powered to show differences in these responder rates.

This table presents the primary safety endpoint, major complication composite, which is a combination of the following major adverse events: death, empyema, massive hemoptysis, pneumonia distal to valve, pneumothorax, and respiratory failure.

As in the previous slide, the difference between the treatment and control is not statistically significant, though there is a strong trend towards higher rates of major complications in the EBV treatment group.

The percentage of patients that experience major complications was over five times higher in the treatment group than in the control, but as in the

previous slide, the study was not powered to show differences in these rates.

The EBV treatment group had statistically significantly higher rates of the following adverse events: hemoptysis, other pulmonary infection, increased shortness of breath, hypoxemia, non-cardiac chest pain, nausea or vomiting, and all valve-related adverse events. The control group did not have significantly higher rates of any adverse event.

The EBV treatment group also had statistically significantly higher rates of the following serious adverse events: COPD exacerbation as a category of events, hemoptysis, and all valve-related serious adverse events.

Rehospitalization was categorized as a secondary safety endpoint and not a serious adverse event, though it is often the result of a serious adverse event.

The control group did not have significantly higher rates of any serious adverse event.

It was agreed that the European data would not be pooled with the U.S. data as part of the primary analyses for this PMA. However, it is instructive to look at the results of the European

arm of the clinical trial.

Note, that no ITT analysis was done for this data. The estimates of the difference in the six-minute walk test are lower than in the U.S. arm of the trial, though the FEV_1 estimates are similar.

Also there was statistically significantly more major complications in the treatment group of the European arm of the trial.

In summary, statistical significance was achieved in the primary effectiveness analysis.

However, the estimates of differences between the treatment and control may be impacted by the post-hoc definition of the extended window, the proportion of missing data, the inclusion of major protocol violators, and the lack of blinding.

There were also higher proportions of adverse events and serious adverse events in the treatment group.

Remember, statistical significance does not imply clinical significance. And now I would like to introduce my colleague, Dr. Shure, who will further discuss clinical issues related to this device.

Thank you.

DR. SHURE: Thank you, Alvin. By way of disclosure, I'd like to say that I have no financial

interest or conflict of interest, and by way of background, I've had over 30 years in academic medicine. I was Chair of the Pulmonary Division at the University of Mississippi, and I'm a past President of the American College of Chest Physicians.

My colleague, Dr. Julie Swain, my co-reviewer, is an internationally known cardiothoracic surgeon who also has no financial conflict of interest here. We are both clinical consultants to FDA. Dr. Swain also has over 30 years of experience in academic medicine, chaired several cardiothoracic surgery divisions, and is an experienced surgeon in lung volume reduction surgery.

These are the areas that I'm going to be covering and a lot of this has already been discussed. So I'm going to highlight specific issues within these areas that are clinically relevant for those of you on the Panel.

In terms of the procedure, these are some of the items to keep in mind. First, the VENT trial involved treatment of the single lobe in each patient. The number of valves placed in each lobe was left to the judgment of the bronchoscopist based on his or her assessment at the time of whether the

lobe had been entirely blocked off. The target lobe was selected according to an algorithm based on a software analysis of high resolution chest CTs by the Core Laboratory.

The software used was developed by the Core Laboratory and is not commercially available or FDA approved, and I think Dr. McLennan mentioned a recent FDA approval of a software system, but I do not believe it is the software system used in this study.

You already heard in detail about the overall study design. I want to focus on two aspects of the design for you to consider.

First, I'd like to mention the control group which was optimal medical management. The Sponsor has emphasized comparisons of the VENT trial to LVRS results in the NETT trial, and there are several issues related to this approach to consider.

During the study development, FDA suggested to the Sponsor that LVRS be used as a control. The Sponsor rejected this suggestion. FDA then advised the Sponsor that no comparison could be made to LVRS without a LVRS control.

Secondly, we need to keep in mind that similar entry criteria don't guarantee the same population and, in fact, the NETT population, while

similar, was slightly worse at baseline in terms of FEV_1, TLC, total lung capacity, residual volume, and diffusing capacity.

In addition, the method of patient selection involved a visual assessment of CT based heterogeneity in the NETT trial, which is different again from the software analysis method used in the VENT trial.

Also there can certainly be unknown covariates based on these and other factors. For example, differences could occur because the studies were conducted almost a half decade apart, and the subjects were willing to undergo surgery which might represent a different population in ways that we can't identify.

And finally, comparing VENT to NETT is essentially using historical controls, and you probably noticed that there were no side-to-side comparisons of the results in part because the data are differently reported in the two trials, and without line data, we can't make these comparisons.

Next I'd like to focus on potential ramifications of the fact that the study wasn't blinded.

These are some possible effects for you to

consider as consequences of the unblinded design. Almost all treatments have a placebo effect. In general, the magnitude of the effect has been shown to be related to the amount of ritual associated with the treatment. With devices, the placebo effect can be quite high.

We usually think of a positive placebo effect associated with treatment, but there can also be a negative placebo effect in the control group, which in an unblinded study knows that they are not receiving the treatment.

In addition, since the healthcare personnel aren't blinded, unconscious treatment and assessment biases can also occur.

The entry criteria have already been covered, and these are the major ones here. I just want to note that the Core Lab CT software to determine heterogeneity was an integral part of the major entry criteria, and again this was different from the visual assessment method used in that.

You've already heard about patient follow-up, and I just want to clarify here that the VENT protocol allowed windows for the assessment visits. They were prespecified by the Sponsor in the original protocol as you see on the bottom as plus or minus 14

days for 6 months and plus or minus 30 days for 12 months.

The actual data analyses that we received from the Sponsor and are reviewing here today are based on extensions to these prespecified windows that were developed by the Sponsor after the study was completed and after the data were available. The Sponsor didn't provide reasons for the extensions, but you can see that these new extended windows increased the allowable time for the 6-month visit by an additional 47 days, almost 7 weeks, and for the 12-month visit, by an additional 12 months. And this might have consequences for the determination of missing data as Mr. Van Orden has described.

You should also note since the Sponsor has compared their missing data to NETT, that the six-month window was extended asymmetrically after the data were available, and the interval isn't the same as NETT.

You already heard detailed descriptions of the study endpoints and statistical analysis. I want to highlight clinically that there are two components to the primary effectiveness endpoint, the FEV_1 and the six-minute walk. They are really two components of a co-primary endpoint and both have to be met

according to the goals of the trial, the intended use of the device in the protocol, and the indications for use of the device in the labeling which state that it is intended to improve FEV_1 and six-minute walk.

For those of you who might not be familiar with these tests, they represent two different forms of assessment. The FEV_1 is a physiologic measurement, and the six-minute walk is an assessment of performance or function. So the six-minute walk should provide evidence of an improvement, that an improvement in FEV_1 actually has an effect on patient function. Each, as has been noted, is performed according to American Thoracic Society guidelines in the study.

Now, I'd like to go over some aspects of the study size determinations. The sample size for FEV_1 was estimated based on considerations by the Sponsor that a "clinically significant difference in FEV_1 was 15 percent based on ATS, American Thoracic Society, bronchodilator response recommendations." The Sponsor used 17 percent improvement for the estimate of the 6-minute walk sample size because, as they stated in the protocol, it is between the clinically meaningful threshold of 15 percent and the

6-minute walk historical results of 20.4 percent. The Sponsor's references from the VENT protocol supporting these levels of clinically significant differences are listed below the yellow line on this slide.

Since the Sponsor's white paper, the supplement sent to the Panel, questions the reliability and significance of the six-minute walk test, I should perhaps point out that it is the Sponsor's own co-primary endpoint and the Sponsor uses the six-minute walk, not FEV_1, in evaluating image-based clinical success at six months. And they provided well-recognized references for its use in clinical levels of significance.

We should also remember that most tests have variability, including the FEV_1, which is why we need to see at least a 15 percent change. But both are well validated and have clearly recognized performance standards set by the American Thoracic Society and the European Respiratory Society in joint guidelines.

Again, with respect to the effectiveness endpoint, the Sponsor prespecified analysis at 6 months, and FDA requested collection of all effectiveness data through 12 months for this device,

that while removable, is intended to be permanently implanted in patients with a chronic disease.

Statistical significance for the primary endpoint was prespecified at a one-sided significance level of less than .025.

The non-statisticians on the Panel may want to note that you may be more used to seeing two-sided p-values with significant levels less than .05. So you'll to need make that adjustment to recognizing the one-sided testing and the corresponding .025 p-values in this study.

Turning to the secondary effectiveness endpoints, you've already heard from Mr. Van Orden about the changes to the prespecified secondary effective endpoints after the study was completed and the possible implications of those changes with respect to the final endpoints that were reported, and I won't repeat that analysis.

As we move to the discussion of results, I'm only going to review the results of the first three of these final four endpoints listed. The St. George's Respiratory Questionnaire, the mMRC score, and cycle ergometry.

Again, for those of you who are not familiar with these assessments, I want to point out

that the first two are questionnaires and convey quality of life information related to breathing and daily activities. They represent how a patient feels, and in this sense, they are soft endpoints and subject to placebo effects.

Cycle ergometry is a performance assessment and represents a harder endpoint.

I'm not going to discuss supplemental oxygen use because it was not regulated in the protocol and it could be biased in an unblinded study. Also, there's no recognized clinically meaningful difference for a change in supplemental oxygen use.

There were a large number of additional effective analyses in this study. Some were prespecified and some were post-hoc. Rather than trying to address all of them, I'll focus on the responder analyses that I've listed here.

Beside each one I've listed what levels the Sponsor has specified in the protocol are clinically important differences, 15 percent for FEV_1 and 6-minute walk, and 8 point decrease in the St. George's Respiratory Questionnaire, a 1 point decrease in the mMRC score, and a 10 watt increase in cycle ergometry. These levels are supported in the

literature, and FDA agrees with these levels as clinically important or clinically meaningful differences.

As I mentioned, there are a number of other analyses provided by the Sponsor, and they've highlighted the BODE index in the white paper to the Panel and in the presentation. Despite the Sponsor's emphasis on the BODE index, I don't intend to cover it for several reasons.

First of all, the index was an additional endpoint, not a primary or secondary one, and as Mr. Van Orden noted, there was no correction for multiplicity in determining its significance.

The index has only been in use since 2004 when it was reported as a predictor of death in COPD. It may be a promising composite, but there isn't a wealth of studies validating it as a predictor of response to treatment.

The index, as you heard, is a composite, and here you see the components, BMI, FEV_1, mMRC, and the six-minute walk. So you can see that three of the four components are part of the VENT assessments. FEV_1 and six-minute walk are the components of the VENT co-primary endpoint, and mMRC is a secondary effectiveness endpoint.

So for all of these reasons, I'm only going to focus on the metrics and analyses that I've already mentioned.

Now, turning to safety, the primary safety endpoint specified in the protocol is the composite of major complications, the MCC, and the components are listed here. It's important to note that the components are not rated hierarchically and that the study was not powered to assess safety.

The Sponsor proposed a 30 percent equivalence delta in the non-inferiority hypothesis for the MCC, but FDA did not agree to this because it was felt to be far too high. As a result, no primary safety endpoint hypothesis was agreed on, and FDA stated that the date would be evaluated in its totality in a risk benefit assessment.

Other safety endpoints included survival, another composite or progression to death, LVRS or lung transplantation, rehospitalization and adverse events, and these parameters were followed through one year.

That completes the study overview, and now I'd like to look at the results from a clinical perspective.

This slide shows a graphical view of the

data accountability that Mr. Van Orden already described. This wedge here and here represents the prespecified visit window in the protocol. It includes, as you can see, about 60 percent of the data at 6 months for both treatment and controls and about 70 percent at 12 months. So you can see that a sizable proportion of data are missing based on the prespecified window visits, about 40 percent at 6 months and 30 percent at 1 year.

Even using the post-hoc extended windows, over 20 percent of the data are missing, and Mr. Van Orden has already addressed the statistical issues that may arise from this, and you will need to consider those, too, in terms of the clinical evaluation.

Now, let's turn to the co-primary endpoint components, FEV_1 and six-minute walk. The blue bars here, here, and here show the mean delta or the difference in the percentage changes in the components between the treated and control groups. The yellow bars represent the confidence intervals, and the green bar, which doesn't show up so green here, the vertical green bar is provided for reference. It represents the clinically important difference of 15 percent identified by the Sponsor.

I want to point out that the protocol specifies achieving statistically significant differences between the treatment and controls. The bar is here for clinical relevance. The differences for both components are statistically significant. You will need to evaluate the clinical relevance of these statistical effects.

The six-month data were the basis of the primary effectiveness evaluation prespecified by the Sponsor, but in clinically evaluating the effectiveness of the treatment, one might like to consider the durability of the treatment as well.

To that end, the 12-month data in this graph show that the statistically significant difference in FEV_1 is maintained at approximately the same level while the change in the 6-minute walk test is smaller and does not maintain the statistical significance. Again, the green bar indicates the 15 percent clinically important level, and it's included for reference.

So the effect of the device is not maintained at 12 months in one of the two components of the primary effective endpoint.

Looking at the proportion of patients who met the prespecified clinically important difference

for FEV_1 of 15 percent in the responder analysis, the difference in responders met statistical significance, but a substantial proportion of patients, over 77 percent, in the treatment group, were non-responders.

For the second component of the primary effectiveness endpoint, the six-minute walk, the difference between the valve and control groups was small and didn't meet statistical significance. So again, only one of the two components of the primary endpoint was statistically significant in the responder analysis related to a clinically important level of change.

Another way of looking at the co-primary endpoint response is shown in this FDA analysis showing percentages of responders to both of the components of the co-primary effectiveness endpoint. The differences between the treatment and control group, the difference is not statistically significant, and it's striking that only about seven percent actually had both a physiologic and a functional improvement.

Turning to the secondary effectiveness endpoints, here's the St. George's Respiratory Questionnaire, and you can see the mean delta of

minus 3.4 at 6 months which was statistically significant and the mean at 12 months which was slightly lower, minus 3 and wasn't statistically significant.

Again, the green bar is shown for a reference, and it indicates the clinically important difference of minus 8.

For the mMRC scale, you again see the statistically significant but small delta of minus 0.3 at 6 months and the 0 delta at 12 months. You can see both here in relation to the clinically meaningful change of minus 1.

As I mentioned before, both the mMRC and the St. George's Respiratory Questionnaire are quality of life instruments, and you might be surprised at the magnitude of the deltas considering that you might expect a positive placebo effect in the treated group and a negative placebo effect in the control group. You will, of course, have to decide the clinical significance of these changes.

And for the last metric, cycle ergometry, you see the same pattern, statistical significance at 6 months but not at 12 months, and you can see the delta in relation to the clinically important change of 10 lots.

The Sponsor, as you've heard, has also provided a subset analysis of a group that they've identified as high responders with respect to the primary endpoint, and this group is called high heterogeneity based on the Core Laboratory software assessment.

FDA and the Sponsor disagree as to whether or not this subgroup was prespecified, but the subgroup was at the least not clearly defined in the VENT protocol.

During the PMA process, the Sponsor provided their definition, which is actually based on two different measures of heterogeneity, one at TLC and one at RV, and both measures are used in different analyses. So it doesn't appear to be uniquely defined.

Despite these controversies, it is clear that while the group may show better performance, there is a statistically significant association of the high heterogeneity subgroup with a composite of progression to LVRS of death. So it's also associated with increased risk.

Turning to the safety results, you've already heard that that MCC at six months was numerically worse in the treated group but that this

difference wasn't statistically significant.

Once again, FDA and the Sponsor never came to an agreement on the delta for this endpoint, and the study wasn't powered to detect differences in safety.

You can see here a graphical comparison of the components of the MCC at six months. All are numerically worse in the treated group except for empyema, which didn't occur in either group, but I'd like to focus for a minute on the deaths which were perhaps the most striking component.

In addition to the numbers of deaths, it's instructive to look at the causes of death as well. In the first six months, there were no deaths in the control group. In contrast, in the first 6 months, there were 6 deaths in the valve-treated group which as you saw represented 2.8 percent of patients. Two deaths occurred within the first three weeks after valve implantation.

The first patient developed ischemic colitis two days after the valve procedure. She underwent surgery and remained ventilator dependent with respiratory failure until her death.

The second patient experienced massive hemoptysis eight days after valve placement. He had

a cardiac arrest and was intubated and on mechanic ventilation. We heard from the Sponsor that this hemoptysis was felt not to be probably, but not definitely, valve-related. Reading the actual reports, I would perhaps come to a different conclusion because the bronchoscopy showed the bleeding was predominantly from the right upper lobe where the valves had been placed, and it certainly occurred close after valve placement.

The next four patients died between four and slight over five months post-procedure.

This third one on the slide died at four months from respiratory failure from a COPD exacerbation. At bronchoscopy, mucous plugs were found occluding a number of airways and the valves, but he had also been hospitalized at two months and again at three months for COPD exacerbations also requiring bronchoscopy from mucous plugging. So the hospitalization, the final one, was the third in four months.

The next patient died from a COPD exacerbation with respiratory failure, which began at three and a half months and required mechanical ventilation until his death. This episode was associated with pneumonia in an area that was not

served by the valve.

The fifth patient died of metastatic cancer, and this is the only completely COPD-unrelated death in the group.

And the sixth patient died of respiratory failure from a COPD exacerbation. He also had a complicated course prior to his death, with a hospitalization for a COPD exacerbation at six days after valve placement and another hospitalization at two months for a valve-related hemoptysis. That hospitalization was complicated by a lung infection and a myocardial infarction. So he had a rough time prior to the final episode.

Turning to the 6- to 12-month timeframe, there were three deaths in the control group, representing 3.4 percent of patients. The first was due to lung cancer. The second occurred at home and was due to gradually worsening COPD. The last was related to post-operative complications from a pulmonary wedge resection for a nodule in the lung.

In the treated group, there were two deaths in the 6- to 12-month timeframe representing slightly under one percent. The first was related to a COPD exacerbation associated with a non-valve-related pneumonia. The second patient died of metastatic

cancer but had experienced a COPD exacerbation requiring mechanical ventilation at two and a half months post-valve placement.

So, in summary, the 12-month death rates were comparable, but the pattern was interesting with only one of three control deaths being related to a COPD exacerbation and that occurred in the 6- to 12-month timeframe. Three-quarters of the valve group deaths occurred early, and three-quarters were related to COPD exacerbations. One was caused by valve-related hemoptysis, and three of the eight patients had more than one hospitalization for COPD exacerbations prior to death.

Now, let's turn to another aspect of safety. I'm not sure why this is shown up as black here on this slide, but the control box should be in red and comes to about here and should say 10 percent within there. Some mystery of the computer hardware here.

Here you see the valve group and serious adverse events of COPD exacerbations at 12 months, and the valve group was 23 percent and the control group, which you can't see, is 10 percent.

Once again, you can also see that the valve group also had significantly more adverse events of

COPD exacerbations at 12 months, 72 percent, and it was actually 57 percent in this group coming to about here in the control group. It looks good on the computer screen here, just not on your projector.

Because the adverse events as to post-serious adverse events may sometimes be more relevant to the clinician because they tend to be more numerous, this bar graph shows the 12-month rates of components of COPD and pulmonary-related adverse events. The rates of pulmonary infections, other than pneumonia distal to a valve, increased shortness of breath, non-cardiac chest pain, and hypoxemia are all statistically significant.

In terms of valve-specific events, you see here serious adverse events at 12 months. All valve-related events occurred in 16 percent of patients, and you can see the component rates for serious bronchial pathology, expectoration, or migration of valves and distal pneumonia.

Valve removal was not considered a serious adverse event by the Sponsor, but it's shown here for reference on this column because additional bronchoscopies are required for valve removal, and approximately 14 percent of the treated ITT population had one or more valves removed.

And here's the last safety measure that I want to highlight, and that's hospitalizations. This graph, once again, the control group should be in red, and this screen doesn't seem to like red, but it does show 12-month hospitalization rates. The rates are higher in the valve group than the control, almost 40 percent in the EBV group, and if you could see it, 25 percent in the control group. And that difference is statistically significant.

So far I've summarized the clinical aspects of event trial data. For reference, the European trial, which had a nearly identical protocol and patient demographics, had similar trends in effectiveness, although they were small and not statistically significant. The MCC was also significantly worse in the valve-treated group, 13.5 percent versus 3.3 percent, and this difference was statistically significant.

So this provides you some additional information in a similar patient population for you to consider with respect to both safety and effectiveness.

Clinically, we also need to consider whether the instructions for use provided by the Sponsor are adequate to obtain reasonably similar

safety and effectiveness, and there are three issues that I would like you to consider with respect to the instructions.

First, the method of patient and target lobe selection is not the same in the instructions for use as in the trial. As you saw, the VENT trial used a non-FDA approved software-based algorithm to assess heterogeneity and choose the target lobe. The instructions for use state only that the most involved lobe should be chosen on radiographic assessments; chest radiograph, chest CT, or hydro-CT are not specified. And I should be clear that the Sponsor cannot advise the use of the software system they used in the trial because it isn't FDA approved.

You will need to consider whether or not the instructions provided can be used with similar effectiveness and safety.

Secondly, the instructions for use don't specify how many lobes should be treated, and the VENT trial didn't treat more than a single lobe.

And last, training was provided in the use of the device, in the VENT trial as you've heard, but none is actually included in the instructions for use, and you should consider whether or not this training may influence the safety or effectiveness of

the device in general use.

So, in summary, these are some of the clinical issues that have been raised. There may be issues related to the interpretation of results based on the extent of missing data and post-hoc redefined visit windows. There may be a difference between clinical and statistical significance of the effectiveness endpoints. The Sponsor has mentioned an FDA drug center draft guidance regarding endpoints, but that same guidance repeatedly states that endpoints must be "clinically meaningful and the magnitude of the improvement should be clinically relevant."

With respect to the risk benefit ratio, you will have to make a qualitative judgment based on the totality of the data presented. You may want to consider the deltas at 6 and 12 months for both primary endpoint components and the secondary effectiveness endpoints. And you will need to decide the clinical significance of these changes.

With respect to risk, I've highlighted some issues related to deaths, COPD exacerbations, and hospitalizations for you to consider.

The Sponsor has stressed that the device is removable, but we should perhaps also point out that

we have no long-term studies about the consequences of the device in terms of subsequent lung volume reduction surgery or transplantation.

And finally, there are some issues related to the instructions for use that could potentially affect the safety and effectiveness of the device in general use by effects on patient and target lobe selection.

And now I'd like to introduce my colleague, Dr. Jiping Chen, who will discuss the postmarket issues.

MS. CHEN: Thanks, Dr. Shure. Good morning, distinguished members of the Panel and members of the audience.

My name is Jiping Chen, and I'm one of the epidemiologists in the Division of Postmarket Surveillance in the Office of Surveillance and Biometrics.

As the epidemiologist in the peer review team, I'm responsible for working with the Sponsor for the development of a post-approval study protocol.

The Sponsor has submitted post-approval study protocol for the extended follow-up of the premarket cohort and one PAS outline for the new

patients. In the event that the PMA is approved, we'll continue to work with the Sponsor to develop detailed PAS protocols that both Agency and Sponsor can agree on.

Here's the outline for my presentation today. First, I will discuss the general principles that are utilized when thinking about the need for and designing post-approval studies. Then I will comment on the rationale for postmarket questions that the premarket study was not designed to answer but that may be addressed in the post-approval study. Then I will present the FDA assessment of the PAS, and finally I will discuss PAS issues that we would like the Panel members to discuss on the design of the post-approval study if the PMA is approved.

Before we talk about post-approval studies, we need to clarify a few things. The discussion of a post-approval study prior to a formal recommendation on the approvability of this PMA should not be interpreted to mean FDA is suggesting the Panel find the device approvable.

The plan to conduct a PAS does not decrease the threshold of evidence required to find the device approvable.

The premarket data submitted to the Agency

and discussed today must stand on its own in demonstrating a reasonable assurance of safety and effectiveness in order for the device to be found approvable.

There are two general principles for post-approval studies. The main objective of conducting post-approval studies is to evaluate the device performance and the potential device-related problems in a broader population over an extended period of time after premarket establishment of reasonable evidence of device safety and effectiveness.

Post-approval studies should not be used to evaluate unresolved issues from the premarket phase that are important to the initial establishment of device safety and effectiveness.

The reasons for conducting post-approval studies are to gather postmarket information, including long-term performance of the device, data on how device performs in the real world in a broader patient population that is treated by community-based physicians as opposed to highly selected patients treated by investigators in clinical trials. Evaluation of effectiveness of training programs for use of device, the evaluation of device performance in subgroups of patients since clinical trial tend to

have limited numbers of patients or not patients at all in certain -- subgroups of the general patient population. In addition, post-approval studies are needed to monitor adverse events, especially rare adverse events that were not observed in clinical trials.

And finally, we conduct post-approval studies to address issues and concerns that the Panel members may raise based on their experience and observations.

Here are three questions that our review team considers important in assessing the long-term safety and effectiveness of the device and may be addressed in post-approval studies.

First, what will the real world performance of the device be in the more general population of patients and providers?

Second, what is the long-term safety and effectiveness of the device postmarket?

And, finally, is there need of a postmarket failure analysis for removed or expectorated valves?

The Sponsor has briefly described their PAS plans earlier this morning. We would like to bring to your attention a few issues regarding the Sponsor PAS outline for the new patients. Please be reminded

that the PAS is a prospective, single-arm, open-label, observational study to evaluate the training effectiveness and the long-term device safety and effectiveness in real world settings. All endpoints will be evaluated with descriptive statistics.

Here is the FDA assessment of this PAS outline. First, study design. The fact that the pivotal study met primary effectiveness endpoint at 6 months but not at 12 months raises a concern about the durability of the device effect.

Is a single-arm study with descriptive statistics the most appropriate design for a PAS?

FDA is concerned about the appropriateness of the study design to address device long-term safety and effectiveness without an appropriate comparison group. FDA has discussed with the Sponsor regarding the possibility of comparing device effectiveness between EBV subjects with those who received the lung volume reduction surgery and/or standard of care control from the National Emphysema Treatment Trial.

It is known that VENT trial and the NETT trial has similar inclusion and exclusion criteria. However, consideration have to be given to the fact that EBV was implanted unilaterally and LVRS was

performed bilaterally when designing the PAS.

We would like the Panel members to discuss if there's a need to compare EBV subject with LVRS subjects and standard of care controls to address device long-term safety and effective postmarket.

Second, effectiveness endpoint. Long-term effectiveness will be assessed by evaluating the post-bronchodilator spirometry at one, two and three years post-procedure.

We pose a question as to whether spirometry alone is sufficient to address device long-term effectiveness postmarket.

Given the information described in IFU, it is not appropriate to not consider six-minute walk test as a effectiveness endpoint.

We would like the Panel members to discuss if there's need for evaluation of six-minute walk test in addition to spirometry as effectiveness endpoint.

Third, safety endpoints. The Sponsor will estimate the serious adverse events rates at one, two and three years post-procedure. To be most meaningful and interpretable, all adverse events including death should be documented and those assessed to be procedure or device-related clearly

noted, and summary frequency is provided.

Assessing only SAEs will potentially underestimate the rate of adverse events and is not sufficient for evaluating device long-term safety profile.

As stated by FDA clinician Dr. Shure and statistician Mr. Van Orden earlier, device safety remains a concern in the premarket study. Therefore, FDA's uncertain whether it will be more appropriate to include all adverse events, not just serious adverse events, to adequately interpret the device long-term safety profile.

We would like the Panel members to discuss what safety endpoints should be addressed in the post-approval study.

Fourth, FDA is also concerned about the proposed duration of follow-up. Is a follow-up of three years appropriate?

We would like Panel members to discuss appropriate duration of follow-up to address device long-term safety and effectiveness.

Finally, sample size. The sample size calculation is based on the assumption that the upper one-sided 95 percent confidence limit of the rate of valve expectoration and migration is less than 10

percent, that is observed rate 6 percent plus 4 percent -- with 6 percent being the highest expectorate rate postmarket.

FDA is concerned because of the following: first, the calculation is not hypothesis driven, and the study power is not estimate. Sample size calculation based on a study hypothesis is needed to ensure that the study will have sufficient power to test a hypothesis.

Second, the migration or expectoration rate used in the sample size calculation is less than what was observed premarket. The validity of this assumption is not clear. Although Sponsor claims that postmarket training will play a role in reducing the rate, the effectiveness of training has not been assessed yet.

We would like the Panel members to discuss the appropriateness of the migration/expectoration rate of 6 percent for the postmarket period and to discuss what will be an appropriate safety hypothesis for the post-approval study?

Based on the Sponsor's proposed PAS and our initial assessment, we will be asking the Panel during your afternoon deliberations to discuss whether the proposed PAS plan for new patients is

appropriate to address device long-term safety and effectiveness and make recommendations.

The issues to be discussed include the appropriateness of the study design and control selection, the appropriateness of the assumption that the valve expectoration or migration rate is lower than that observed in the premarket in sample size calculation, and the appropriateness of effectiveness safety endpoints and duration of follow-up.

In addition, we would like the Panel to discuss any additional issues or questions that can be addressed in your post-approval study and make recommendations if the device gets approved.

This concludes my presentation as well as the FDA presentation this morning. We welcome any questions that you may have. Thank you.

DR. BIRNBACH: I'd like to thank the FDA speakers for their presentations. Does anyone on the Panel have any questions specifically for the FDA? You may also ask the FDA questions later today. So if I don't get to you now.

DR. VASSILIADES: This is not directed to anyone in particular, so whoever wants to answer.

First was the EU data. Was that data available prior to the start of the U.S. trial? And

what was the discussion about why that was not rolled into? Was it simply because the data was not statistically significant in terms of why that data was not rolled in and used as part of the trial?

DR. SWAIN: Julie Swain. We might ask the Sponsor to answer that question.

DR. VASSILIADES: Okay.

DR. SWAIN: That would be most appropriate.

DR. VASSILIADES: Okay. So we'll come back to that later.

DR. SWAIN: Sorry, guys.

DR. VASSILIADES: Another question. I'd like the FDA's opinion, since we seem to be making a big point about the window extension, and I understand from a theoretical standpoint why that's important, to sort of adhere to their agreed-upon windows at the beginning, but from a practical standpoint, with this study in particular, is this really a big deal as to whether someone comes in three weeks of the procedure versus two weeks?

DR. SHURE: That's always an interesting practical, clinical sort of a question. The issue here I think is that we don't know. The windows were prespecified. They were narrow windows, but they were chosen by the Sponsor, and I think the concern

is that the study was completed, the data were available. Whether the analysis had been done or not, we don't know, but if you notice, the six-month window was widened asymmetrically. It wasn't widened to the same window as NETT. We just don't know, and I think Mr. Van Orden pointed out that it's not verifiable that the hypothesis, that looking inside or outside the window would be the same. We just don't know.

DR. VASSILIADES: Next question is does anyone have an opinion or a professional assessment of why the high heterogeneous subgroup has a higher increase in safety issues? You mentioned that several times, but I didn't quite understand. Maybe there's a rationale for that or it's just a noticed outcome? The reason I'm asking, of course, is, you know, at least from my impression, this seems to be the target population. I understand that these subgroups weren't identified and they were done post-hoc, but it seems to me on a theoretical and physiologic basis that these are the kinds of patients that you're wanting to treat and identify, yet they also are the population that has the highest or the least favorable safety profile, and I just want to know more about that.

DR. BIRNBACH: For the transcriber, can you, although you've already introduced yourself, each time you come to the microphone, reintroduce yourself so they know who you are.

DR. SHURE: I'm going to jump in for a moment, and I'm Dr. Shure, in front of my statistical colleague because I think part of this is clinical.

First of all, as you understand, I can't tell you why but I do -- and I think that we have differences with the Sponsor in terms of both the definitions of this subgroup and the significance of it.

However, physiologically, and this is certainly something for the Panel to discus in the afternoon, you might, this is just a hypothesis, to speculate, that a group with relatively more lung destruction and perhaps more heterogeneity, if that's the case, might be a sicker population or it might be more susceptible to adverse events. I think that's also something you can discuss during the afternoon with the Sponsor, but we don't have a clinical answer for you on that, but there is a definite association. Mr. Van Orden.

MR. VAN ORDEN: Yeah, from a statistical point of view, obviously I'm not a clinician, if you

look at the percents that are in those groups, there's not a wide gap. It's more because it's a continuous variable, and it may be the reason why the high heterogeneity subgroup shows such good results because it doesn't include those sick patients that died.

DR. SHURE: So you can see more of an effect in sicker patients, you're saying but the --

MR. VAN ORDEN: The sicker patients aren't included.

DR. SHURE: -- perhaps.

MR. VAN ORDEN: I wouldn't --

DR. SURE: I don't think we know.

MR. VAN ORDEN: I wouldn't read a lot into the higher death rate, but I also wouldn't read a lot into the higher --

DR. VASSILIADES: You know, if the Sponsor wants to comment on that during their time again later, then I'd be more than happy to listen to that. I have one last question. Were you trying to imply evidence, I mean I'm also concerned about the placebo effect and naturally any unblinded trial, there's going to be a placebo effect of some sort and particularly with some of the subjectivity of the primary endpoints, but you've mentioned we see in the

data that the 6-minute walk test deteriorated in terms of its results in superiority over the control from 6 months to 12 months. Is it your assessment that that is a result of the placebo effect potentially wearing off or the possible placebo effect in the treatment group wearing off from 6 months to 12 months?

DR. SHURE: That's certainly possible. That's certainly possible.

DR. VASSILIADES: Okay. Yeah.

DR. SHURE: It's certainly possible.

DR. VASSILIADES: All right.

DR. BIRNBACH: Dr. Wilcox.

DR. WILCOX: Thank you. I have a couple of questions perhaps better answered by the Sponsor. I'll just put them on the floor. The first is at some pointed we noted there were 70 major violations, and so that patients in whom those occurred were dropped out of the calculations. Is that correct?

MR. VAN ORDEN: No, they were dropped out of the Per Protocol population. They were included in the Completed Cases --

DR. WILCOX: But in the PP group, they were left -- they were dropped out?

MR. VAN ORDEN: They were included in the

primary analysis, these major protocol violators.

DR. WILCOX: I don't understand.

DR. SHURE: Not in the Per Protocol.

DR. WILCOX: Not in the Per Protocol, but my question is did that make any difference?

MR. VAN ORDEN: In the Per Protocol population, the six-minute walk test was not significant.

DR. WILCOX: Was not significant.

DR. SHURE: So when those patients were excluded --

DR. WILCOX: When they were excluded --

DR. SHURE: -- it was not significant.

DR. WILCOX: If they were included, they were significant -- it was significant?

DR. SHURE: Right.

DR. WILCOX: Okay. Thank you. And do you know if any sort of -- we've talked about the risk benefit analysis. Do you know if there was any sort of cost benefit analysis in terms of -- between the two groups --

DR. BIRNBACH: That's outside the area of this discussion. Dr. Wiswell first and then --

DR. SHURE: I'm sorry. Was that --

DR. WISWELL: I have a question about the

data for the high heterogeneous groups. Other than the -- and it's about the safety. Other than the lung reduction surgery and deaths being increased in that subgroup population, did you analyze other safety data, MCC, AEs, SAEs, rehospitalizations, in that subgroup population, that 40 percent that were highly heterogeneous?

MR. VAN ORDEN: This safety analysis wasn't done looking at the high heterogeneous. It was not a designed analysis to look at the subgroup. We were just looking at -- there was a prespecified analysis to look at factors that might lead to death, may lead to death or LVRS, and the high heterogeneity variable was significant in that analysis. We did not do a lot of analyses with -- just based on that subgroup.

DR. WISWELL: I put forth maybe to the Sponsors if they can respond this afternoon with other safety data, just in that roughly 40 percent overall, the patients to see what the other safety profile data would be, please.

DR. BIRNBACH: Dr. Loeb.

DR. LOEB: This is the second Panel that I've been on, and so I want to ask for your opinion, maybe having more background, but it seems like the whole FDA presentation is that there were a lot of

problems identified in this study. I think when looking at any study, especially a study of magnitude, it's easy -- well, yeah, it's easy after the fact to look and say where there were a lot of problems. Every journal club, that's what you spend your time doing, saying, well, this study could have been done better, and you can find that in every study.

Can you give me your gut feeling or from your experience, do you think that this study had a lot more problems than you would anticipate given the size and complexity of the patient population and what you're looking at? Does it strike you that there are more problems than there should have been for this type of a study?

DR. SWAIN: This is Julie Swain, and having been on the Panel for a decade and chaired it for several years, this is exactly the question you all discuss. It's not a question, although we, the FDA folks, deal with studies every single day. It's not one, a question we should answer, and it's a question because you're on the Panel, you have experience in clinical trials, you need to answer that.

DR. BIRNBACH: You took the words right out of my mouth. Dr. Halabi, and then we're going to go

to Dr. Ries.

DR. HALABI: So the first question for the statistician is have you done an analysis on the primary endpoint, if you use the prespecified window, even though you have 58 percent and 61 percent of the data that are not missing? And was that statistically significant?

MR. VAN ORDEN: I believe the Sponsor presented one at just the prespecified window earlier today, and the six-minute walk test with p-value was .025, which is the cutoff for significance or not.

DR. HALABI: Okay. And the next question had to do again with the protocol violations. So when you exclude the patients, I believe there were about 62 patients who had -- excuse me -- 49 patients and 20 patients in the control arm had important clinical violation. When you excluded that in your Per Protocol analysis, what window did you use? Did you use the post-hoc window or the --

MR. VAN ORDEN: That includes the extended window, yes.

DR. HALABI: The extended window. And this was statistically not significant for the six-minute walk, but again the question is have you done any analysis that used the prespecified window?

MR. VAN ORDEN: For the Per Protocol?

DR. HALABI: Actually for all of them, whether -- yeah, Per Protocol.

MR. VAN ORDEN: No analysis has been done on just the prespecified window for the Per Protocol or ITT populations.

DR. HALABI: And then the final question had to do with the missing data. Did you compare the baseline data for those patients which were missing versus those were not because I know the missing at random assumption is not verifiable, but at least you do have a sense of who are those patients that are missing at baseline.

MR. VAN ORDEN: I mean, as I said in the presentation, that there were more that withdrew or never came in, but as far as other baseline characteristics, I don't believe there were any significant differences.

DR. HALABI: Okay. Thank you.

DR. BIRNBACH: Dr. Li.

DR. LI: Just to introduce myself to the Panel. I'm not a physician. I'm a Ph.D., and my expertise is in materials of design. So if I ask questions that are obvious to everybody else, I apologize.

One question I had, and this may be a completely stupid question, but I'm a little confused over how many valves were placed in each patient. Was it just one or -- because there was one diagram that had three, and if there were a different number of valves placed in each patient, was there any breakdown of the results versus the number of valves placed in a patient?

DR. SHURE: The Sponsor can perhaps address this more in the afternoon session, but there was no limit on the number of valves that could be placed. The goal was to place valves either in lobar, segmental, or sub-segmental bronchi. Those are the major branches to a lobe of the lung until the bronchoscopist developed that the lobe had been entirely blocked off. That's called lobar exclusion in your material there. So there was no limit on that. I believe that the average was four valves, and the range was up to nine valves, but I don't have that information in front of me. And I believe that the breakdown was on, in analysis looking at lobar exclusion, not based on the number of valves but whether lobar exclusion was achieved. That's how some of this subset analyses were provided.

And the image-based analysis, which I guess

neither the Sponsor nor FDA actually covered here but is covered I think in your material, was based on lobar exclusion by image or by what the bronchoscopist had observed.

DR. LI: So is the answer then --

DR. SHURE: There is not a breakdown as far as I know by the number of valves, and I'm not really sure that would be relevant since the point is putting in as many as you need to, to totally block off the lobe.

DR. LI: I can understand it from that standpoint, but from the device side, it's a little hard to judge to see if the device is effective where you simply didn't put in enough valves or you put them in the wrong place. So it's not so much measuring the clinical outcome but more addressed of why you get the results that you get, but if the answer is no, that's fine. That answers my question.

The other question I had was one that seems to be glossed over, but in my reading, 45 percent of the 214 patients had a valve either removed or replaced for a variety of reasons. So that's basically almost half the patients had one or more valves removed, and based on your first answer, I don't know how many valves were removed in each

patient. So my question is did you look, did you break down the data again looking for why the data came out the way it did? Those patients that had valves that were removed or replaced versus any of the clinical outcomes.

DR. VAN ORDEN: We did look at those analyses and could not determine based on those analyses if there was a relationship between having the valve removed and the effectiveness of the device.

DR. LI: How about for a single --

DR. SHURE: Did you say 45 percent?

DR. LI: Yeah, and I just double-checked that. It was in your executive summary.

DR. SHURE: I think in part you need to separate out the valves that were removed and repositioned or replaced during the procedure itself versus those that were removed after the procedure.

DR. LI: Well, I think that's one breakdown. But, for instance, for bleeding for instance, it's not too outrageous to think that if you have to reposition a valve a couple of times, that you may have caused bleeding.

DR. SHURE: Right. There are data on that.

DR. LI: Okay. So my question is did

anybody look to see for any effects of replacement or movement of valves that had migrated or --

DR. SHURE: I don't think we received analyses like that, did we, Alvin?

MR. VAN ORDEN: The analyses we received didn't show any difference in the effectiveness based on removal of the device or not.

DR. LI: Okay. There were two things that I read in the Panel pack that I couldn't find the results for. One was the, I guess in the preplanning they were going to look at the effectiveness of essentially a surgeon experience. In other words, you know, where the people that have complications, were these the first valves they put in or was there any association with the experience or the number of valves placed by that physician versus the clinical outcome?

MR. VAN ORDEN: To my recollection, nothing was found with regards to --

DR. SHURE: I think the Sponsor will have to address that. I do not recall seeing that in the material that we received.

DR. LI: Okay. And then one last question if I can find it.

DR. BIRNBACH: While you're looking --

DR. LI: I'll save the rest for the Sponsor, thank you.

DR. BIRNBACH: Dr. Ries.

DR. RIES: I have several questions related to the FDA issues that you've raised for us, but they are also similar to some of the clarifications I want to address with the Sponsor. I'll hold those for later, but I have two I want to ask you now.

One is since an important part of the analysis is the issue of missing data and protocol violations, and I'm a little confused about what a protocol violation is, so I'd like to ask the FDA's interpretation, and then we could get information from the Sponsor.

As I understand the protocol, patients were screened when they came into the study and they met certain criteria. Then they went off for six or eight weeks of rehabilitation, and then they were prepared for -- after rehabilitation was really a baseline prior to the randomization and actual entry into the randomized portion of the study. There were 62 patients who were violators, but 23 according to the Sponsor were violations at initial screening. I assume that means before rehabilitation, but they qualified at the point of enrollment. Why do you

consider those violators?

DR. SHURE: The Sponsor identified those.

MR. VAN ORDEN: Yeah. We should clarify that there were lots of protocol violations, but the only ones we're concerned about are the ones that the Sponsor identified as clinically important and excluded from the protocol.

DR. RIES: What is the time point of which you define a violator? At the initial screening or at the time they're randomized, which is several months later --

DR. SHURE: Well, I think you have to have --

DR. RIES: -- after rehab.

DR. SHURE: Now, this was not our definition.

DR. RIES: I understand.

DR. SHURE: There were almost 3,000 protocol violations. We're not talking about the small ones, and the Sponsor identified the ones that were considered major. I understand your point. Well, what if they still met the inclusion criteria at the time after rehab? But the point I think for those would be that you still don't know if it's exactly the same patient population. If they didn't

meet it going in, you know, what does that mean? We don't know what it means, but they weren't handled the same way. And some, you know, were errors in, you know, choosing the value or whatever, but we just don't know how much of a difference that makes.

DR. RIES: I'll ask the Sponsor for more clarification this afternoon.

The second question maybe for now because I guess the one secondary analysis that sort of dropped off the screen was the supplemental oxygen use, and I wasn't sure exactly what -- but what was your perspective of --

DR. SHURE: I'm sorry. What was that again?

DR. RIES: You decided to exclude in your comments the supplemental oxygen use because --

DR. SHURE: The BODE?

DR. RIES: No, the supplemental oxygen.

DR. SHURE: Why did I exclude that?

DR. RIES: Yeah.

DR. SHURE: It was not -- there were recommendations in the protocol for oxygen use, but it was not subject to protocol. So in an unblinded study, that could be biased. So, in addition, what's the minimally important clinical difference for

change in supplemental oxygen use? I've never seen anything on that. I don't know how to interpret it. So that's why I chose not to look at that, and then you saw Mr. Van Orden's analysis of the possible statistical significance question related to it. But, you know, it's not subject to protocol and it could be a biased assessment.

DR. BIRNBACH: Dr. Halabi.

DR. HALABI: I have another question for the statistician just for clarification. So in the analysis you have 75 patients in the control group, and then when you excluded 20 violations, major protocol violations, did your sample size go down to 55 patients in the control group? How many patients you had also in the experimental arm?

MR. VAN ORDEN: I believe the Per Protocol numbers are on my slide in front of me but the second study design slide.

DR. HALABI: So 57 controls and 141 EBV, correct?

MR. VAN ORDEN: That's correct, yeah.

DR. SHURE: What slide number?

DR. HALABI: This is slide number 14.

DR. SHURE: Study design.

DR. HALABI: Yes.

DR. VAN ORDEN: There is some variability in some, how the Completed Cases and Per Protocol were -- I mean, depending on the endpoint and so -- yeah, all of the numbers I have presented are as reported by the Sponsor, but we could get down to it does depend on the case by case, and the numbers in the study design on slide 14 are max.

DR. HALABI: The maximum.

DR. VAN ORDEN: Yeah.

DR. HALABI: Okay. Thank you.

DR. BIRNBACH: Any other --

DR. VASSILIADES: One more question. Dr. Shure, I have a question for you. Since you're an expert in this area, in your opinion, are pulmonary function tests felt to be primarily an objective test, or is there some element of subjectivity in the interpretation as well as patient effort?

DR. SHURE: Okay. Pulmonary function tests, well, let's talk about spirometry, which is what -- we'll talk about both spirometry and body box measurements. That's lung volumes. You would think that they are not subject to placebo effect, but they are subject to variation, both day-to-day, week-to-week, and month-to-month, in an individual and that's

well recognized. One of the references that I have there from the American Thoracic Society gives guidelines for what the minimum change has to be to overcome that variability. It's not variability in the test itself. It's variability in the patient. So that's why over months you need to see a 15 percent change in FEV_1 for it to be clinically meaningful. And that's where the 12 percent difference comes for a bronchodilator response. That's an acute response right away to a drug. To treatments that are longer, you need different metrics, and those are well described.

The FEV_1, you could breathe out, it's how fast you can breathe out, you know, basically, and you could breathe out slower, but there are well defined standards, and I am sure that they adhere to them in this study, you know, they're done to ATS guidelines which means you look for specific things to be sure that the test is interpretable, and you exclude tests that are not interpretable, and these are well standardized. There are error codes, for example, if between three measurements there's too much variation in the FEV_1, that's not reliable, and that will be excluded.

So there are a lot of controls set in to

deal with the noise, if you will, in performing these tests. So there's variability over time per patient, but there shouldn't be, one would think, a placebo effect. That is I'm trying hard enough because if you're not trying hard enough, it should show up in the shape of the curve and in the error codes because there's going to be variability. It's possible not to show up, but it's less likely.

DR. VASSILIADES: And what is your assessment with respect to six-minute walk test?

DR. SHURE: The six-minute walk test, and I've provided you with references on that, is well standardized. As many functional tests are, it has a psychological component to it as well because, you know, if you're not motivated, you're not going to walk as far, perhaps. So it's partly how you feel as well as, you know, your absolutely muscle strength. It's a composite measure, if you will, of function, and that's recognized. But it's conducted in a very standardized way so that it should be reasonably reproducible, but again that's where, you know, the variation comes in and the 15 percent or greater standard is set. But conducted according to the standards, it's really a well recognized test. It correlates, as you can see, added to FEV_1 and in the

BODE index, it becomes even by itself, FEV_1 and six-minute walk by themselves are reasonable predictors. Combined, they're even perhaps better predictors of outcomes.

But they are -- that may be more subject to a placebo effect than the FEV_1 conceivably because there is this psychological component to your performance, but it is conducted in a very standardized way that should get the best performance.

DR. WILCOX: Does the standard include measuring PO_2?

DR. SHURE: Does it include what?

DR. WILCOX: Measuring PO_2 in the patient?

DR. SHURE: Measuring PO_2?

DR. WILCOX: Yes, to determine the effort.

DR. SHURE: You mean de-saturation.

DR. WILCOX: Yes.

DR. SHURE: No, there's an assessment of fatigue and an assessment of dyspnea, the BODE index that goes along with that. Many people will monitor oxygen saturation, but it is not stopped for de-saturation unless it's dangerous. It's stopped because the patient cannot either walk six minutes or cannot -- or stops.

DR. WILCOX: A number of us believe that you need to have some sort of assessment of effort, and that being the better one in that test, de-saturation. If they walk themselves to de-saturation then --

DR. SHURE: Well, based on the way it's performed, it should be the patient's best effort. They're encouraged to walk as far as they can, and it's done in a very standardized way, and those guidelines are clearly provided and are part of the protocol.

DR. BIRNBACH: Before we break, just before the coffee break, there was a request by the Sponsor for clarification of something that Dr. Dominik had asked. Are you okay with that? This is your last chance for homework before the afternoon.

(No response.)

DR. BIRNBACH: Okay. We will now break for lunch. It is 2 minutes to 12:00. We will reconvene again in this room in one hour at 2 minutes to 1:00. Please take any personal belongings you may want with you at this time. The ballroom will be secured by FDA staff during the lunch break. You will not be allowed back into the room until we reconvene.

I'd like to again remind the Panel members

that there should be no discussion of the PMA during break amongst yourselves, with the Sponsor, the FDA, or with the public.

(Whereupon, at 11:58 a.m., a luncheon recess was taken.)

A F T E R N O O N S E S S I O N

(1:00 p.m.)

DR. BIRNBACH: Welcome back. Before we proceed with the Panel discussion, I would like to ask to the Sponsor to come forward to address any detailed issue raised in the morning session that the Sponsor had been asked to address after the lunch break, and that includes not only the statistical issues brought up by Dr. Dominik, but I think that Dr. Li and Dr. Vassiliades also had some unanswered questions that need to be addressed.

MR. McCUTCHEON: Thank you, Mr. Chairman. John McCutcheon with Emphasys Medical. We'll attempt to answer all those questions. If I missed anything, please just readdress them.

First of all, there was a question on why we did not pool the European data, and there was a change in our FDA review team when we first submitted the PMA. So initially we were in the Plastics and Reconstructive division, and during discussions with them, they had asked that we not pool the data. So we agreed to do that and amended the statistical analysis plan at that time. So that may be aware to this current review team.

Another question was on the extended

windows, why did we do it, why is it asymmetrical, and what are the impact on outcomes? I think in Dr. Sciurba's presentation, we can show it again, he actually showed that there was no impact on outcomes whether you use the narrow windows or the extended windows. So we showed that sensitivity analysis earlier this morning.

DR. BIRNBACH: If I might break in, why did you have different windows?

MR. McCUTCHEON: The asymmetry?

DR. BIRNBACH: Yes.

MR. McCUTCHEON: We did that because we thought that it would be more conservative because it's a progressive disease, and we thought the patients, if you're going too early, we'd be picking up healthier patients perhaps. So we wanted to make it asymmetrical, and it was at the suggestion of our statistician that we do that, so that we're not biasing it in our favor.

So, in other words, anybody that we picked up that was outside the window going forward were more likely to be further progressed in their disease. So we felt that that was a -- yeah, and there weren't that many on the low side. Most the patients missed it on the high side, came in later

rather than earlier, but it was really an arbitrary decision trying to apply that logic that we thought is a more conservative way to do it.

The reasons why and the timing, it was before the data was unblinded, and it was done by our statistician per his recommendation, and I think we were naive when we set the windows for this patient population, and plus or minus 14 days turned out not to be practical, and I think you'll find by talking to our clinicians that we ran a very tight trial. We didn't make exemptions and allow any, although there were some protocol deviations, none were condoned by the Sponsor, and in this case, we just set too tight of a window.

A lot of these patients have COPD exacerbations and aren't available in that tight of a window and then the rescheduling takes some time. So we were just naive the way we set it. When we benchmarked it to other studies, they're much longer in this patient population and so we're trying to walk that balance.

It was strictly a numbers issue though because again when we do the sensitivity analysis, it makes no difference. The outcome's the same. We just lose a lot of power using the narrow window.

DR. DOMINIK: Can I ask a follow-up to that? You mentioned before the data were unblinded. So this was a unblinded study. So you're suggesting that the dataset did not have treatment group identify on it at the time.

DR. CHIACCHIERINI: The dataset did have treatment group identify and, in fact, we had to look at that to make sure it was balanced between the treated and control group. The sole reason why we extended the period was to make sure that there was a significant population followed in the total six-month window. And we actually did that without looking at the result. We didn't look at the FEV_1 or the six-minute walk. We looked strictly at participation in the actual observation of that visit.

DR. BIRNBACH: And again, for the transcriber, please reintroduce yourself every time you go to the mic.

DR. CHIACCHIERINI: I'm Dr. Chiacchierini.

DR. BIRNBACH: Dr. Dominik, does that answer your question?

DR. DOMINIK: That's helpful.

MR. McCUTCHEON: I'm going to ask Dr. Ernst to come up. There were several questions on

competence intervals for the individual adverse events and the baseline comparisons of the available data versus the imputed ITT group.

DR. ERNST: Thank you, John. My name is Armin Ernst again for the record.

These are some of the slides that have been requested. I'll just put that up here. It has the confidence intervals for the MCCs as well as the p-values. It doesn't really change anything statistically or clinically, and you'll see it remains the same that the driver is really the post-obstructive pneumonia and the MCC rate.

I'd like to take the opportunity while you look at it just to remind everybody in the Panel, though, when you look at the numbers of deaths, three in the control and eight in the treatment group, that this was a two to one randomization. So this was not equally distributed. So if it would have been equal, it would have been reasonable to expect six deaths in the control group.

Also, the timing of deaths, that there was nothing in the first six months, remember when you look at the actual time of deaths in the control group, it was at six, six and a half, and seven. So if one of those patients, just a single one would

have died two weeks earlier, it would have been a completely overlapping graph. So this is probably a random event.

But anyway, does that --

DR. DOMINIK: The additional safety events that were compared between treatment groups, the one I think, your slide 60, COPD, and there's six or seven events of interest there where you have treatment and control proportions of events --

DR. ERNST: The non --

DR. DOMINIK: Yes. Do you have confidence intervals about the difference --

DR. ERNST: No, we do not.

DR. DOMINIK: -- in proportions for those? Okay.

DR. ERNST: So this is for the MCC group, the prespecified endpoint.

I would also like to show you these slides because that had come up. If there is a different patient population that withdrew that would be identifiable in terms of its characteristics from the patients that remained in, and so here on the left side, in the left column you see the means for the patients who made it through the trial and the right side are the ones that withdrew, and you can see that

these are really comparable populations and those statistics, there are no significant differences. So nobody seemed to have been favored in the exit of this trial. This also applies to all the pulmonary function tests that I have shown before. Again, there is no significant difference. They're well matched, the patients who completed and the patients that withdrew, and that also applies to the rest of the variables.

There is -- let me just go back where I see this. The only thing where there is a small difference, even though it did not achieve statistical significance, I just want to point this out, is the diabetes, but that I doubt has any clinical significance either. Does that answer the questions of the Panel regarding --

DR. BIRNBACH: I believe so. Dr. Halabi.

DR. HALABI: Could you define those dates? What do you mean by withdrew? Are these based on the window, or this is based on the patients who were counted as missing and was that missing using the Per Protocol or --

MR. McCUTCHEON: These were based on the patients that remained on the extended windows, and it does not include the patients who died. So these

were patients who dropped out versus patients who remained in the extended window.

DR. CHIACCHIERINI: The term withdrew meant --

DR. BIRNBACH: Could you introduce yourself?

DR. CHIACCHIERINI: I'm sorry. I'm Dr. Chiacchierini again. The term withdrew means that they withdrew prior to the end of the six-month window without a visit. So if they withdrew prior to that time, if they withdrew in the window and did not come in for a visit, they were considered to have withdrawn, and we actually would have had to impute a value for them as well.

DR. HALABI: Thank you.

MR. McCUTCHEON: John McCutcheon again. Are there other questions that we can answer for the Panel?

DR. BIRNBACH: Is there anyone else on the Panel who would like to ask any questions? Dr. Li.

DR. LI: If I may, if I could ask some questions about the device again. Was there any effect of experience, there were several different designs of devices used and the number of valves for a patient. Did you find any correlations of any of

those features with any of the performance criteria?

MR. McCUTCHEON: No, we did not. We did a lot of the multivariate analysis using valve version, and that was never significant. We didn't identify any learning effect or learning curve in operator experience, and there was also no correlation with number of valves with either safety or outcome measures that we could find. And the average, you asked earlier, I believe that it was 3.8 valves per subject on the average were used.

DR. LI: Okay. If you could bear with me, there was this one number that pops up several places that 45 percent of the subjects had one or more of the valves removed or replaced. How was it determined that that valve needed to be removed or replaced? And was that immediately recognized or did, you know, minutes pass or some time pass before they recognized that the device had to be removed or replaced?

DR. STRANGE: Charlie Strange. It's interesting that 44 percent of people coming into the procedure had a valve repositioned during that initial bronchoscopy, and the way we do that is by visual assessment. I think you saw the picture of the little knee of that valve sitting outside of an

orifice, but this is not a hard thing to do. You grab the edge of the valve, you pull it out, you take it, you repackage it inside the loading catheter again, and have another loading catheter ready to put the second valve in. So it adds probably three or four minutes to the procedure if you fire a valve and it doesn't seat perfectly. And it's done very frequently. In the analysis plan, there was a comparison between the people that had repositioned valves versus those that had correct fires the very first time, and there was no difference in any safety or efficacy outcome between that 44 percent and the others that did not have repositioning at that initial visit.

DR. LI: And again, just so I understand the procedure, if the valve needs to be repositioned, it's completely removed, and a new valve is put back in.

DR. STRANGE: Occasionally, if a valve is seated too deeply, then you can grab the edge and wiggle it a little bit and move it forward into a correct position. The majority of the 44 percent had a valve placed too proximally. That means it is at risk for expectoration or migration, and those were always removed. You can't really push a valve

further and deeper bronchoscopically.

DR. LI: Okay. And then, I'm not sure how to ask you this question. How do you know the valve is actually working? If you put, you know, for instance up to nine valves in the patient and there's no change of lung volume for instance, how do you know actually the valve is working, or if you put in nine valves, how do you know all nine are all actually working if it's just a visual assessment of a physician?

DR. STRANGE: Well, if you remember the photographs of the valves with the four pictures across there, you can actually see these valves with respirations indent, having a concavity to them, and then you'll actually see them open with air exhalation on expiration.

DR. LI: So that's confirmed in every seated valve then, that that valve opens and closes?

DR. STRANGE: It's visually assessed. I think you'll also find in the PMA packet that on six-month CT scan, there was a finding that some valves were not completely occlusive, and I think this is a way of going forward. We can further target just to make sure that new investigators, new physicians can really focus on this proper placement, get good

seating, and I think the company has, in addition, added that marker valve that properly place the valve. There's some training issues about getting these valves done correctly the first time.

DR. LI: And then my last question, there was one section there where there was an evaluation of using imaging to determine technical success or procedural success. And in the summary that I saw that at 6 months, only 56 percent of the devices evaluated were considered at imaging a technical success, and 85 percent of those 56 percent had a least one valve that was not fully occlusive. So that kind of goes back to my question about how do you know that these things are actually -- and I guess what I'm struggling for is your hypothesis of relieving the pressure seems reasonable, but the clinical results really just aren't as impressive I would expect them to be if that was simply the solution. So either that isn't the actual solution or the valves really aren't working as well as they should. So that's my question.

MR. McCUTCHEON: Given the nature of the study, and this truly was a landmark study, I'm not aware of any other device in this area that's had 100 percent HRCT follow-up, pre and post. We also had

bronchoscopic video. So when Jonathan Goldin's, Dr. Goldin's Core Lab identified any valves that had a leak, you can detect it on CT scan, we were then able to go back and pull the bronchoscopic video and match those up, and as a company, there was a collective learning curve, if you will, not an individual physician, but we've modified our techniques as the clinicians have over time. We were able to eventually train our eye to say this now bronchoscopically, we can tell it's at an angle, it may be in a distal bifurcation and not fully occlusive, something that we didn't appreciate until we had all this great CT follow-up. So there was some learning there, too, and we believe that we're getting better and better at placing the valves and knowing when they are placed.

There's also an empirical test. If you get volume reduction -- and so if there's a diaphragmatic shift and some clear volume reduction, empirically you know you've got a great seal. It's just pure mechanics or physics. If you don't get that, there's only two reasons. Either there's a leak around the valve or you've got collateral flow, and again we're getting better and better at detecting both of those. But when you get the full seal, and you have a

complete fissure, you have volume reduction, and you get the clinical benefit associated with that.

DR. BIRNBACH: We seem to have seamlessly moved from your response to specific questions to our general Panel discussion. So, for the record, we are now proceeding with the official Panel discussion. I open the floor to the Panel members for questions either to the Sponsor or to the FDA, and we'll start with Dr. Willsie.

DR. WILLSIE: Okay. A couple of questions for the Sponsor again. I was actually going to ask you if you had gone back and reviewed the videotapes because I think that's important.

With the learning curves of the operator, was there improved efficacy of appropriate placement with increased experience by the operator?

MR. McCUTCHEON: John McCutcheon. We weren't able to quantitatively show a learning curve at all, but we do believe, and I again, I don't have data to support this, we do believe that over time placements got better and better and that the evidence was there were fewer expectorations. Our experience today in Europe is that that's a very rare event. We're seeing greater rates of atelectasis or volume reduction, and so we believe that we can now

more often, more frequently visually detect whether it's seated properly or not and getting much, much better results as a result of that.

DR. WILLSIE: And along those lines then, with the individuals who expectorated one of the devices, did someone go back independently and look to see whether the valve had been replaced correctly?

MR. McCUTCHEON: Absolutely. For every one that happened, where we -- and we didn't have 100 percent bronchoscopic video, but when we had them, we tried and occasionally someone didn't start the VCR, and so it wasn't 100 percent, but for most of them, and we could always identify a root cause. You saw the picture. That was a classic example of one being too proximal. We did an analysis by placement, and in the superior segment of the lower lobe in the B6, that's a very complex, sometimes trifurcates, we had a higher rate of expectorations from that particular segment. So we learned some anatomical things as well as delivery techniques.

DR. WILLSIE: Okay. One final question if I could. There's something in the instructions about encouraging your patient not to cough, and I guess I'd like to know how effective that was, and did you give them antitussives or do you recommend that

because with COPD patients, they cough, cough, cough, most of them. So -- or many of them.

MR. McCUTCHEON: I think that's a legacy that we put in five or six or seven years ago, and with no empirical reason, it just seemed like the thing to do to not have them coughing too hard once they had an implant.

DR. ERNST: You know, I completely agree with you. I think that's unrealistic, and it is probably not reasonable to expect that this has any correlation with valve expectoration. So even though this is in the instructions, it's probably a legacy event and really not a necessary component.

DR. WILLSIE: Thank you.

DR. BIRNBACH: Dr. Marcus.

DR. MARCUS: Just a quick question. Is there any benefit of re-looking at valves a week later just to see how they're doing? Has that been looked at, at all?

DR. STRANGE: Charlie Strange. It hasn't been looked at. When we came into the six-month CT data though and saw that a fair number were misplaced, I think all of us have wondered if a one-month CT scan or a one-month re-look might be something that could be done in a post-approval

study.

DR. BIRNBACH: Dr. Ries.

DR. RIES: I have a couple of questions just to clarify some protocol and just to make sure I understand some of these issues properly. Back to the issue of protocol violations and the missing data. So my understanding is that you determine protocol violations at the point at which they were screened. Then they went through rehab and were enrolled in the study, but there were 23 patients who actually met entry criteria. Going in, was it the assumption that these violators were defined -- did you mean to define them at the point of screening, or was there an intention that you would then have a secondary screening after rehab?

DR. STRANGE: Just so everybody can be clear, we had screen one which you came in and signed the informed consent, a screen two where you went through the whole day of pulmonary function laboratory. The laboratory sent off for Codamine for instance that didn't come back for a while, and a fair number, and then your lung function assessments. The lung function assessments in the protocol were per ATS criteria. And a number of laboratories around American and in the trial don't use NHANES for

their spirometry, and so in the re-look by the Sponsor, when they came back around for monitoring, many of these, 23 ended up having, instead of a 45 percent FEV_1, now a 46 percent FEV_1 and violating inclusion criteria. The majority of those were spirometric. There were three individuals that didn't get their flu shot, one individual that came into the study on aspirin which was a protocol violation, and those 23, when they went out for pulmonary rehab and came back, had no other exclusions at the time of the baseline evaluation which was post-rehab and still got the full spirometry and full assessment at their baseline assessment. There were 39 individuals that at that baseline assessment had something wrong with their inclusion or exclusion criteria.

DR. RIES: Yeah, I was just a little confused of, you know, why you would set the inclusion criteria at the baseline, not at post-rehab, because I know for instance in NETT, because patients went through rehab and things changed in rehab, and so that the inclusion criteria were really set post-rehab. It seems like you didn't --

DR. STRANGE: I think we have to realize here that there were a fair number of people that

came in with their COPD that didn't have hyperinflation, that the screen failure rate was about two or three to one for the patients that eventually made it into the trial just because they weren't adequately hyperinflated to achieve benefit with our intervention.

DR. RIES: Then a question about missing data. I presume that all of the visits were required to be in-person visits. Is that correct?

DR. STRANGE: They were.

DR. RIES: Because, you know, one of the issues you have is a number of the -- the questionnaires could be administered at a distance, and was there any attempt for patients who didn't come in for an in-person visit to obtain any data at a distance?

DR. STRANGE: There was not. All these were in-person visits.

DR. RIES: And the questionnaires that you administered, the quality of life, the St. George's, and the quality of well-being, were those all self-administered versions or did you require an interview for those?

DR. STRANGE: I don't remember.

DR. SCIURBA: They were self-administered.

It was allowable to be self-administered.

DR. RIES: Okay. And then a specific question about the quality of well-being scale because, you know, one of the issues you have, I understand you removed deaths from the analysis obviously for most of the things, but one of the real advantages of the QWB instrument is that it does include death on the scale. So you don't have the problem of a survivor bias. I notice that in the QWB analysis, did you exclude the deaths from that because that would have a dramatic impact on that scale, if you're including the deaths or not including the deaths.

DR. SCIURBA: The deaths were included on the QWB for the exact reason you described.

DR. RIES: Okay. And then would you get back, I'm having a problem understanding what has happened to this supplemental oxygen use secondary endpoint? What was that measure you were trying to derive and what happened to it?

DR. SCIURBA: My understanding, and John can correct me, he did the interaction with the FDA on this, is that there was a strong encouragement to include a supplemental oxygen criteria in the secondary outcome parameters. And so it was included

but, of course, it was based on patient questionnaires, liters per minute extrapolated through the day, and a calculation was based on that. It had no precedent validity. I think the FDA, the current group, acknowledges there's limitations in its value.

DR. RIES: So it was supposed to be a flow rate times whatever --

DR. SCIURBA: Correct.

DR. RIES: -- self-reported time --

DR. SCIURBA: Extrapolated flow rate based on historical reporting.

DR. RIES: And then if I can ask one, this is the big question I'm having trouble with, back to the issue of the volume reduction because obviously you didn't see volume reduction, so this may be sort of a conceptual question or a philosophical question about what this device is about. Is the expectation that mechanistically this has something to do with lung volume reduction or is it just lung volume redistribution now? We shouldn't really look at volume reduction. We're talking about benefits that are independent of overall volume reduction.

DR. SCIURBA: I mean I'm fairly confident that regardless of what we call it, that many of the