Pulmonx Corporation

Clinical Investigational Plan #630-0012_H

Statistical Analysis Plan – Revised 26 October 2017
(supersedes Appendix 19 from CIP #630-0012_H)

Change History:

SAP Section	Change	Rationale
Study Design	Update the enrolled study participants to 190.	To reflect the number of study participants enrolled in the study.
Secondary Effectiveness Endpoints	Analyses of secondary endpoints were updated to limit the number of statistical comparisons. Now limited to Absolute Change in FEV1, 6MWD, and SGRQ score. The use of Hochberg method was specified to address multiplicity.	To control the family-wise Type I error for the secondary endpoints, analyses of secondary endpoints were limited.
Safety Endpoint	Deleted “of the EBV treatment arm”.	To reflect that adverse events will be compared between treatment arms during the short-term and long-term periods.
Additional Measures	Additional measures were updated, specifically, Treatment Lobe Volume Reduction (TLVR) for the treatment arm only was moved from Secondary to Additional.	Moved to “Additional Measures” since this parameter indicates the mechanism of action and unable to compare to a Control group as no TLVR is expected without an intervention.
Statistical Methods – I	Confidence intervals were removed from adverse event reporting. In place of presenting confidence intervals, adverse event rates for those events occurring in at least 3% of subjects in either group will be compared using a Fisher’s Exact test.	Making adverse event reporting consistent with typical way of presenting and analyzing adverse events.
Statistical Methods II (C)	Pseudo-sites clarified to be created from sites within the same country.	Clarification.
Statistical Methods III	Definitions of populations clarified.	Clarification.
Statistical Methods IV	Update to “three” secondary effectiveness variables in the text and in Table 3.	To reflect the modified secondary effectiveness endpoints.
Statistical Methods IV (B)	Update to analyses of secondary endpoints, including use of Hochberg method.	To control the family-wise Type I error for the secondary endpoints,

		analyses of secondary endpoints were limited.
Statistical Methods IV (C)	Update to analyses of additional effectiveness endpoints.	As a result of modifying the secondary effectiveness endpoints, additional measures were also updated.
Statistical Methods V	Clarify that covariate analyses are to be performed on primary and secondary endpoints only. Removed variables that will not have data available for control subjects. Deleted "Follow-up" column as no "Follow-up" variables will be used in covariate analysis.	Clarification.
Safety Analyses VII	Updated to include Control arm in summary of adverse events and specify Fisher's Exact tests of events occurring in at least 3% of either treatment arm.	In place of presenting confidence intervals, adverse event rates for those events occurring in at least 3% of subjects will be compared using a Fisher's Exact test.
Statistical Methods IX	Included imputation strategy for subject deaths and clarified secondary endpoints would utilize same imputation method as primary endpoint.	Subject death prior to endpoint visit should be considered a failure. Secondary endpoints should account for missing data.
Statistical Methods X	Clarification to statistical tests.	Clarification.
Other	Tables and text formatted.	Ease of reading.

Analysis Summary

Study Title: Lung Function Improvement after Bronchoscopic Lung Volume Reduction with Pulmonx Endobronchial Valves used in Treatment of Emphysema (Clinical Protocol 630-0012).

Patient Population: Male and female patients with severe heterogeneous emphysema meeting the study eligibility criteria.

Study Design: This will be a multi-center, prospective, randomized, controlled study with EBV treatment statistically evaluated using Intent-to-Treat (ITT) analyses. A maximum of 183 ITT study participants, who meet study entry criteria, consisting of screening eligibility criteria, baseline eligibility criteria, and procedure eligibility criteria, were planned to be enrolled. A total of 190 study participants were enrolled. Safety and effectiveness of bronchoscopic lung volume reduction (BLVR) using the Pulmonx EBV will be evaluated at 1 year. An interim analysis designed to evaluate effectiveness for continuing crossover of control participants at 1 year to EBV treatment will be performed when 74 study participants have completed the 1-year follow-up. For study participants who have been treated with EBV, a secondary valve intervention such as valve removal, replacement, or adjustment may be considered during the study follow-up. Long-term data will be collected annually for EBV-treated study participants through 5 years. Per the regulatory plan agreed to with FDA, 1-year of follow-up is required pre-approval and the remaining 4 years of follow-up will be conducted post-approval.

Primary Effectiveness Endpoint
The percentage of study participants in the Endobronchial Valve (EBV) treatment arm who meet the threshold of $\geq$15% improved forced expiratory volume in one second (FEV_1) as compared to the control arm at 1 year. Improved FEV_1 will be calculated by determining the percentage change for FEV_1 from baseline to 1-year post-procedure for individual study participants.

Secondary Effectiveness Endpoints

1) FEV_1: Difference between study arms in 'absolute change from baseline' for FEV1 at 1 year.

2) 6-Minute Walk Distance (6MWD): Difference between study arms in 'absolute change from baseline' for 6MWD at 1 year.

3) St. George's Respiratory Questionnaire: Difference between study arms in 'absolute change from baseline' for SGRQ score at 1 year.

Safety Endpoint
Evaluation of the short- and long-term adverse events profile during the treatment period, defined as the day of the study procedure until 45 days after the study procedure (short), and in the post-treatment period, defined as 46 days after the study procedure until the 1-year follow-up visit (long).

Additional Measures

- Treatment Lobe Volume Reduction (TLVR) for the treatment arm only
- Spirometry, including FEV_1, forced vital capacity (FVC) and the ratio of FEV_1/FVC
- Body plethysmography, including residual volume (RV), inspiratory capacity (IC), functional residual capacity (FRC), TLC, and the ratios of RV/TLC and IC/TLC
- SGRQ Total and domain (i.e. 'symptoms', 'activity' and 'impacts on daily life') scores
- Modified Medical Research Council (mMRC) Dyspnea Scale Score
- BODE Index
- Transitional Dyspnea Index (TDI) from Baseline Dyspnea Index (BDI)
- COPD Assessment Test (CAT)
- SF-36 Health Survey score
- EQ-5D Health Survey score
- Health Care Utilization Questionnaire
- 6MWD test
- Borg scale dyspnea scores before and after 6MWD test
- Change in use of 'maintenance' medications, including bronchodilators, corticosteroids, antibiotics, and anti-inflammatories
- Pulmonary rehabilitation compliance diary responses
- EXACT-PRO diary responses
- Health status change responses
- Carbon Monoxide Diffusing Capacity (DL_{CO})
- Lung radiographic features

Interim Analysis

An interim data analysis designed to evaluate effectiveness for continuing crossover of control arm study participants at 1-year to EBV treatment will be performed when 74 (50% of the total 147) study participants have completed the 1-year follow-up. The interim analysis will be reviewed by the DSMB and FDA. If crossover of control arm study participants is found to be justified by the interim analysis, then crossover of control arm study participants may be continued. If the DSMB recommends not continuing crossing control arm participants to EBV treatment, then those control arm patients who have not yet crossed over will exit from the study per protocol after the 1-year visit.

Long-Term Follow-up

Data will be collected annually through 5 years. Per the regulatory plan agreed to with FDA, 1 year of follow-up is required pre-approval and the remaining 4 years will be conducted post-approval. The 2-, 3-, 4-, and 5-year data will consist of FEV_1 and adverse events.

Definition of Study Success: The study will be a success if the difference between the EBV Treatment arm and Control arm for the percentage of study participants meeting the threshold of $\geq$15% improved in FEV_1 differs significantly (two-sided test at $p \leq 0.05$) in favor of the treatment group at 1 year. This evaluation will be conducted at 1-year post-randomization.

Number of Study Participants: 190

Statistical Methods

I. Descriptive Statistics
Means, standard deviations, medians, and confidence intervals will be reported for all continuous variables. Dichotomous variables will be reported as percentages and the numerator and denominator will be reported and defined.

II. Analyses of the Patient Populations
These analyses are intended to determine the similarity of two treatment groups and similarity of patients from different study sites with respect to important demographic or other variables, either known or suspected to have an influence on the outcome variables. The absence of similarity for any variable will identify that variable as a potential covariate in subsequent analyses.

A. Comparability of Treatment Groups – To assess the success of the randomization process, the demographic and prognostic variables measured at study entry will be compared between the Treatment and Control arms. Continuous variables will be compared with the two-sample t-test or the Wilcoxon two-sample rank test, and categorical variables will be compared with the Fisher's exact test or a Chi-square test. Comparability analyses will be done with two-sided tests with significance level of 0.05.

B. Study Site Comparability –The appropriateness of pooling the data across study sites will be determined on a clinical basis, i.e., to ascertain if the sites used a common protocol, the sponsor adequately monitored the study to assure protocol compliance, and the data gathering and validation mechanisms were the same across all study sites (Meinert, 1986).

An analysis will be done to determine if the magnitude of the clinical effect of the primary outcome is maintained if sites and/or site by treatment interactions are included in an analysis of covariance model. We expect the majority of sites will show the treatment to be beneficial but statistically one would not expect all sites to show the treatment to be beneficial. For those sites with contrary results, an analysis will be attempted to determine what factors at those sites led to the result (See expert statistical testimony from Dispute Resolution Panel transcript September 6, 2001).

C. In study sites with small numbers of patients, it will not be possible to evaluate site or site by treatment interaction. The reason is that what may appear to be a site by treatment interaction may be a small numbers phenomenon. For example, if there were only three patients from a given study site with one in the control group and two in the treated group, one success in each arm will appear to be a site by treatment interaction (100% success in the controls but 50% success in the treated arm). Hence, study sites, within the same country, with fewer than three patients in either treatment arm will be combined into one or more pseudo-sites to allow the comparison to be done. The size of any pseudo-site created in this way will not exceed the size of the study site with the largest enrollment.

III. Analyses Populations

A. Geographic Cohort: This study will be conducted at clinical sites inside and outside the U.S.

B. Endpoint Analyses: Primary and secondary study endpoints will be analyzed utilizing an intent-to-treat (ITT) population defined as all randomized patients analyzed by the groups to which they were randomly assigned (EBV treatment or control).

C. Secondary Analyses: Secondary analyses will also be performed on the study primary and secondary endpoints on a Completed Cases (CC) basis and Per-Protocol (PP) basis.

i. The CC population is defined as all randomized and eligible patients who received study-directed treatment and attended the 1-Year follow-up visit.

ii. The PP population is defined as all randomized patients who meet study eligibility criteria, who were treated as randomly assigned, and had follow-up within window for the primary endpoint (i.e., 1-Year FEV_1). Any visits where protocol violations occurred will be reviewed for possible exclusion from impacted analyses.

D. Safety Analysis Population: Both the ITT analysis population and 'As Treated' (AT) analysis population will be used to assess the safety data. For the AT analysis, study participants will be analyzed based on the treatment they actually received.

E. Secondary Valve Procedures: If applicable, statistical analyses will also be conducted after sorting patients by the actual treatment received, specifically valve removal, valve replacement, and valve adjustment.

IV. Effectiveness Analyses

There is one primary effectiveness variable, three secondary effectiveness variables, and several additional effectiveness variables.

A. Analysis of the Primary Effectiveness Endpoint

Both an interim and an end of study analysis for the primary effectiveness endpoint are planned. The primary effectiveness endpoint is the difference between the EBV treatment arm and control arm in percentage of study participants who reach a threshold of $\geq$15% improved FEV_1, collected post-bronchodilator, at 1 year (see Table 1). The FEV_1 value will be calculated by determining the percentage change for FEV_1 from baseline to 1-year post-procedure using: ((FEV_1 at 1-year follow-up subtracted from FEV_1 at baseline) / (FEV_1 at baseline)) for individual study participants. The two arms will be compared using the standard normal Z-statistic.

Table 1. Primary Effectiveness Endpoint

Variable	Unit	Values Statistically Evaluated
Forced Expiratory Volume (FEV_1)	$\geq$15% improved	percentage of participants

An interim data analysis designed to evaluate effectiveness for continuing crossover of control arm study participants at the 1-year follow-up to EBV treatment will be performed when 74 (50% of the total 147) study participants have completed the 1-year follow-up. The interim analysis will be reviewed by the DSMB.

To account for the interim analysis, the power spending function, defined as:

$$\alpha(t) = \alpha t^{\phi}$$

will be used to preserve an overall type I error rate of 0.05 for the study. Using the nTerim program (Statistical Solutions) to calculate the power spending function with:

α =.05, ϕ=2.3 and *t*=.5, $\alpha(t)$ =.01,

the value of the Z-statistic must exceed 2.571 (nominal alpha <0.01) for the null hypothesis to be rejected at this interim look (see Table 2).

Table 2. Power Spending Function for Interim Analysis

Looks	**1**	**2**
Time	0.5	1
Lower Bound	-2.57057	-2.00360
Upper Bound	2.57057	2.00360
Nominal Alpha	0.01015	0.04511
Incremental Alpha	0.01015	0.03985
Cumulative Alpha	0.01015	0.0500
Exit Probability	38.75	50.52
Cumulative Exit Probability	38.75	89.28

If Z≥2.571 then continuing crossover of control arm study participants will be strongly justified since the p-value will be ≤0.01. This observation may provide evidence to stop the trial early.

The study hypothesis will be tested again at the end of the study. The Z-statistic will be calculated again, and by taking into account the interim analysis, will have a final critical boundary value of 2.004, per the nTerim program. If the trial is not stopped as a result of the interim analysis, then the final Z-statistic must be greater than or equal to 2.004 in order to reject the null hypothesis at the final analysis (at the overall 2-sided 5% significance level). Test statistic is "Z" as defined as:

$$Z = \frac{p_T - p_C}{\sqrt{p(1-p)\left(\frac{1}{n_T} + \frac{1}{n_C}\right)}}$$

Where,

p_T and p_C are the proportions of success for the primary endpoint in the treatment and control groups, respectively, *p* is the pooled estimate of the success rate, and

n_T and n_C are the sample sizes obtained in the treatment and control groups, respectively. Two sets of (p, p_T, p_C, n_T, n_C) will be obtained: one at the interim analysis and one at end of the study.

B. Analysis of the Secondary Effectiveness Endpoints

The following describes the analysis techniques to be used for each of the secondary endpoints. To control the family-wise type I error rate at 5%, the Hochberg step-up procedure will be utilized (Hochberg, 1988).

a. FEV_1

Difference between study arms in 'absolute change from baseline' for FEV1 score at 1 year. Descriptive statistics will include means, standard deviations and 95% confidence intervals. An analysis of covariance (ANCOVA) with factor of treatment and baseline FEV1 as a covariate will be used to test the difference between treatment arms. P-value will be adjusted for multiple imputation.

b. 6-Minute Walk Distance (6MWD)

Difference between study arms in 'absolute change from baseline' for 6MWD at 1 year. Descriptive statistics will include means, standard deviations and 95% confidence intervals. An analysis of covariance (ANCOVA) with factor of treatment and baseline 6MWD as a covariate will be used to test the difference between treatment arms. P-value will be adjusted for multiple imputation.

c. St. George's Respiratory Questionnaire (SGRQ)

Difference between study arms in 'absolute change from baseline' for SGRQ score at 1 year. Descriptive statistics will include means, standard deviations and 95% confidence intervals. An analysis of covariance (ANCOVA) with factor of treatment and baseline SGRQ as a covariate will be used to test the difference between treatment arms. P-value will be adjusted for multiple imputation.

d. The secondary effectiveness endpoints are summarized in Table 3.

Table 3. Secondary Effectiveness Endpoints

Variable	Unit	Values Statistically Assessed
FEV_1	Liters	mean absolute change
6-Minute Walk Distance	Meters	mean absolute change
SGRQ Total Score	Points	mean absolute change

C. Analysis of Additional Effectiveness Endpoints

a. Supporting Evidence for Effectiveness

The following additional effectiveness endpoints will be measured for both study arms. These are expected to provide supporting evidence of the effectiveness of EBV treatment. Results will be described with summary statistics. These endpoints will be described for each study arm separately and comparatively between arms by calculating mean change or difference in proportions, whichever is appropriate for the variable being analyzed. The definitions for these additional effectiveness endpoints are shown in Table 4.

Table 4. Additional Effectiveness Endpoints

Spirometry Measures		
FEV_1	percent	mean percent change
FEV_1	Liters	mean absolute change at 6 months
FEV_1	≥12% improved	Percentage of participants
Forced Vital Capacity (FVC)	liters	mean absolute and percent change
FEV_1/FVC	ratio	mean absolute change
DL_{CO}	percentage	mean absolute and percent change
Body Plethysmography Measures		
Residual Volume (RV)	liters	mean absolute and percent change
Inspiratory Capacity (IC)	liters	mean absolute and percent change
Functional Residual Capacity (FRC)		mean absolute and percent change
Total Lung Capacity (TLC)	liters	mean absolute and percent change
RV/TLC	ratio	mean absolute change
IC/TLC	ratio	mean absolute change
Patient-Reported Health Status Measures		
SGRQ Total Score	percent	mean percent change
SGRQ Total Score	Points	mean absolute change at 6 months
SGRQ Total Score	≥4 points improved	percentage of participants
	≥8 points improved	percentage of participants
SGRQ Domain Scores		
‘Symptoms’	points	mean absolute change
‘Activity’	points	mean absolute change
‘Impacts Daily Life’	points	mean absolute change
mMRC Score	semi-quantitative scale score	mean absolute change
COPD Assessment Test (CAT)	points	mean absolute change
Transitional Dyspnea Index (TDI)	points	mean absolute change
	≤1 point improved	percent of participants
Short Form (SF)-36 Health Survey	points	mean absolute change
EQ-5D Health Survey	points	mean absolute change
Other Measures		
Target Lobe Volume Reduction (TLVR)	milliliters	mean absolute change
TLVR	percent	mean percent change
BODE Index	semi-quantitative scale score	mean absolute change
	≤1, 0, ≥1 change	percentage of participants
6MWD	percent	mean percent change
6MWD	Meters	mean absolute change at 6 months
6MWD	≥25 meters improved	percentage of participants
6MWD	≥54 meters improved	percentage of participants
6MWD – subjects stratified by distance walked at baseline	meters	mean absolute change
Borg Scale Dyspnea Score: Before 6MWD test	semi-quantitative scale score	mean absolute change

Borg Scale Dyspnea Score: After 6MWD test	semi-quantitative scale score	mean absolute change

b. Informational Purposes:
The following additional effectiveness endpoints will be measured in both treatment groups for informational purposes. Results will be described with summary statistics (see Table 5). These endpoints will be described for each treatment group separately and comparatively between groups by calculating mean change or difference in proportions, whichever is appropriate for the variable being analyzed.

Table 5. Informational Variables

Variable	**Unit**	**Values Statistically Assessed**
Maintenance Medications		
Bronchodilators	change in medication regimen	percentage of participants
Corticosteroids	change in medication regimen	percentage of participants
Antibiotics	change in medication regimen	percentage of participants
Anti-Inflammatories	change in medication regimen	percentage of participants
Pulmonary Rehabilitation Compliance Diary	reported adherence to program	percentage of participants
Adherence to program	number of sessions / week	mean absolute number
Intensity of program	length of reported sessions	mean absolute number
EXACT-PRO Diary Entries	points	mean absolute number
Health Utilization Measures	frequency of use	percentage of participants
Lung Radiographic Features		
TLVR	volumetric change	mean percentage change
valve occlusion	qualitative assessment	percentage of participants
inter-lobar fissures	percent complete	percentage of participants

V. Adjustment for Potential Confounders

The covariates shown in Table 6 may be potentially influential with respect to the effectiveness endpoints. If any of these covariates are found to be out of balance between the groups (at the two-sided 0.10 level), they will be included in a multivariate analysis, along with an indicator of treatment group, to evaluate their potential effect on the study conclusions. A multivariate logistic regression model will be used for primary and secondary endpoints.

Table 6. Covariates

	Baseline	Procedure
Gender	X	
Age	X	
Clinical Site	X	
6MWD	X	
Patient-reported health status measures (e.g. SGRQ)	X	
Body Mass Index	X	
BODE Index	X	
Medication Use	X	
Residual Volume (RV)	X	
RV % Predicted	X	
Total Lung Capacity (TLC)	X	
TLC % Predicted	X	
RV/TLC	X	
IC/TLC	X	
Vital Capacity	X	
Forced Vital Capacity (FVC)	X	
FVC % Predicted	X	
FEV_1	X	
FEV_1 % Predicted	X	
FEV_1 / FVC	X	
Lobar Volume	X	
Lobar Destruction Scores	X	
Ipsilateral DS Heterogeneity	X	
Severity of emphysema (GOLD classification)	X	
Comorbidities	X	
Target Treatment Lobe		X
Treatment Group		X

VI. Secondary Valve Procedures

Secondary valve procedures will be evaluated by examining the proportion of patients who undergo valve removal, valve replacement, or valve adjustment, and describing the reasons they occurred. Data for these study participants will be contrasted to data for study participants who did not receive a secondary valve procedure.

VII. Safety Analyses

Evaluation of the short- and long-term adverse events profile during the treatment period, defined as the day of the study procedure until 45 days after the study procedure (short), and in the post-treatment period, defined as 46 days after the study procedure until the 1-year follow-up visit (long). For the Control arm, the date of study procedure will be considered the date of the bronchoscopy assessment. Safety analyses will be performed by examining the percentage of patients that experiences adverse events. The rates and 95% exact confidence intervals will be presented on an event and on a per-patient basis. The percentage of subjects reporting adverse events occurring in at least 3% of subjects in either treatment arm will be compared using a Fisher's Exact test.

Adverse events will be categorized into clinically relevant groups (*e.g.:* stable pneumothorax with no intervention, pneumothorax resolved with chest tube insertion in less than 7 days,

prolonged air leak, etc.). Adverse events occurring in the treatment arm will be further categorized by severity and as device-related, procedure-related, or neither, and by those occurring during procedure hospitalization and those occurring post-discharge.

Rehospitalization rates will be reported by study arm on a Per-Patient (PPt) basis and on a Per-Event (PE) basis. In study participants who receive EBV treatment, adverse events will be evaluated by stratifying participants by use of conscious sedation and general anesthesia during the bronchoscopy procedure as well as evaluating number of valves received and sites EBV were placed.

VIII. Mortality

A Kaplan-Meier survival curve for each study arm will graphically display the time to death from the time of randomization and a log rank test will be used to compare Kaplan-Meier curves between study arms. A Cox proportional hazards model will be used to control for study site along with other significant co-variates as applicable.

IX. Patient Accountability and Missing Data

Every effort will be made to collect all data points in the study. The sponsor plans to minimize the amount of missing data by appropriate management of the prospective clinical trial, proper screening of study subjects, and training of participating investigators, monitors and study coordinators.

- The analysis for the primary endpoint will be performed by imputing missing data. Subject death prior to the 1-year visit date will be imputed as failure.
- For study participant FEV_1 data that is 'intermittent', missing outcomes will be imputed by linear interpolation using the FEV_1 value from the latest non-missing data point before the missed data point and the earliest non-missed data point after the missed data point.
- For study participants with truncated data (e.g. participants who drop out or are lost to follow-up), a multiple imputation strategy will be performed using the propensity score method. In brief, for a particular outcome, the propensities for study participants to have missing data (for each treatment group separately), modeled by logistic regression, are grouped into strata based on percentiles of the logistic propensity score model. Within a stratum, a study participant with a missing observation has an imputed value assigned by randomly choosing a value from among the study participants in the same stratum with non-missing observations. This procedure will be repeated 20 times on the entire dataset, resulting in 20 different 'complete' datasets allowing for estimation of the effect on the outcome of interest, accounting for missing data.

An additional analysis for the primary endpoint will be performed based on not imputing data. All partial data that is available on subjects who drop out during the course of the study will be included. In addition, sensitivity analyses such as worse-case or best-case imputation will be performed although it is recognized that they are biased.

Missing data for secondary endpoints will be imputed using the same methodology described above. Subject death prior to the 1-year visit date will be imputed baseline carried forward (0-point change).

X. Comparison of Crossed-Over Control versus Study Treatment Outcomes

a. Effectiveness Outcomes

Table 7. Effectiveness Endpoints for Control Arm Crossovers (comparison based on paired Control versus EBV Treatment differences)

Variable	Unit	Values Statistically Evaluated
Forced Expiratory Volume (FEV_1)	≥15% improved	percentage of participants
FEV_1	Liters	mean absolute change
FEV_1	percent	mean percent change
SGRQ Total Score	Points	mean absolute change
mMRC Score	semi-quantitative scale score	mean absolute change
6-Minute Walk Distance	Meters	mean absolute change
6MWD – subjects stratified by distance walked at baseline	meters	mean absolute change
6MWD	≥25 meters improved	percentage of participants
DL_{CO}	percentage	mean absolute and percent change
Residual Volume (RV)	liters	mean absolute and percent change
Inspiratory Capacity (IC)	liters	mean absolute and percent change
Total Lung Capacity (TLC)	liters	mean absolute and percent change
RV/TLC	ratio	mean absolute change
IC/TLC	ratio	mean absolute change

The time-course of outcomes will be described. A paired t-test will be used to test the difference between treatment modalities for the quantitative outcomes. For binary outcomes, a McNemar test will be used to test the difference between the two treatment modalities.

b. Safety Outcomes

Evaluation of the short- and long-term adverse events profile of crossed-over participants during the second year of follow-up, when they have the EBV treatment, will be assessed as in Section VII above. Additionally, the time-course of occurrence of Serious Adverse Events (SAEs) will be contrasted and compared for the first year (Control Treatment) versus second year (EBV Treatment) periods. For those SAEs that could occur with either Control or EBV Treatment, a McNemar test-statistic will be used to quantify the magnitude of the difference between the two treatment modalities.

XI. Randomization Assignment Method

Study participants who are determined to meet screening, baseline, and procedure eligibility criteria will be randomly assigned to Study Treatment (EBV or Control). Random assignment will be performed using a stratified permuted block design, generated separately for each clinical site, with assignment stratified by anatomical site of the planned treatment (e.g. right lung or left lung). Mixed block sizes will be used.

XIII. Device Malfunction Analyses

Results of any device malfunctions and their sequelae are to be presented descriptively. The rate and exact 95% confidence intervals will be computed.

XIV. Statistical Software

The parametric and non-parametric analysis of variance and other primary analyses will be done using SAS, Version 9.2 or later or StatXact.

References

1. Meinert, C. (1986). *Clinical Trials: Design, Conduct, and Analysis*. Oxford University Press, New York.

2. Medical Devices Dispute Resolution Panel Meeting of September 6, 2001 (Panel Transcript)

3. Hochberg (1988) A sharper Bonferroni procedure for multiple tests of significance. Biometrika 75(4):800-802