PINNACLE

Statistical Analysis Plan

PR1087-002, Version B

15 October 2019

Approvals: Signatures on File

Statistician, NAMSA	Signature	Date
Reviewed by:		
	See e-signature on GrandAvenue	
Consultant Regulatory Affairs Libra Medical	Signature	Date
	See e-signature on GrandAvenue	
Clinical Manager	Signature	Date

Revision History		
Rev	Release Date	Description
A	19 Aug 2019	Initial version
B	15 Oct 2019	Updates based on FDA feedback and protocol revisions: • Revised definition of study success • Added details on the statistical test for the primary efficacy endpoint • Added new secondary endpoint for IPSS improvement • Added details for secondary endpoints • Changes to analyses for handling missing data

Table of Contents

1 List of Abbreviations and Phrases 5
2 Introduction 5
2.1 Objectives 6
3 Study Design Summary 6
3.1 Overall Study Design and Plan 6
3.2 Test and Control Arm Description 7
3.3 Cross-over Description 7
3.4 Pharmacokinetics Arm Description 8
3.5 Primary Endpoints 8
3.5.1 Primary Efficacy Endpoint 1 8
3.5.2 Primary Safety Endpoint 1 9
3.6 Secondary Endpoints 9
3.6.1 Secondary Endpoint 1: Average IPSS improvement in the Test arm at 12 months 10
3.6.2 Secondary Endpoint 2: Percentage of responders at 3 months 10
3.6.3 Secondary Endpoint 3: Durability – Percentage of responders 11
3.6.4 Secondary Endpoint 4: Qmax 11
3.6.5 Ancillary Endpoints and Other Analyses 11
3.7 Pharmacokinetics Arm 12
3.8 Randomization and Blinding 12
3.8.1 Randomization Plan 12
3.8.2 Blinding 12
3.9 Cross-over of Control Subjects 12
4 Study Sample Justification 13
4.1 Primary Efficacy Endpoint Sample Size Considerations 13
5 Analysis Populations Set 13
6 Poolability of Data 14
6.1 Sites 14
6.2 Subgroup Analysis 14
7 Missing data 15

7.1 Missing Data Handling Rules 15
7.1.1 For the Primary Efficacy Endpoint 15
7.1.2 For the Primary Safety Endpoint 15
7.2 Multiple Responses 16
7.3 Sensitivity Analyses for Missing Data 16

1 List of Abbreviations and Phrases

AE	Adverse Event
BPH	Benign Prostatic Hyperplasia
DCB	Drug Coated Balloon
DMC	Data Monitoring Committee
GI	Gastrointestinal
IPP	Intravesical Prostatic Protrusion
IPSS	International Prostate Symptom Score
IRB	Institutional Review Board
ITT	Intent-to-Treat
PP	Per Protocol
Qmax	Peak Flow Rate
QoL	Quality of Life
Optilume	Urotronic's Optilume BPH Catheter System
PK	Pharmacokinetics
SAE	Serious Adverse Event
UADE	Unanticipated Adverse Device Effect
UTI	Urinary Tract Infection

2 Introduction

The Optilume™ BPH Catheter System (Optilume) is designed to treat subjects with benign prostatic hyperplasia (BPH). It consists of two investigational catheters:

- Optilume™ BPH Prostatic Pre-dilation Catheter
- Optilume™ BPH Prostatic Dilation DCB Catheter

Treatment with the Optilume is intended for men with BPH who meet all the study inclusion criteria and none of the study exclusion criteria. All analyses will be conducted using SAS Version 9.4 or later, or similarly well accepted analysis software (e.g. R).

2.1 Objectives

The primary objectives of the trial are to evaluate the safety and efficacy of Optilume in the treatment of BPH. Primary efficacy will be assessed based on the reduction in the International Prostate Symptom Score (IPSS). The primary safety endpoint will be assessed based on major device-related serious complications. Safety will also be assessed from the reported serious adverse events (SAEs) and unanticipated adverse device effects (UADEs).

3 Study Design Summary

3.1 Overall Study Design and Plan

This is a prospective, multi-center, double blind, randomized controlled clinical trial in a 2:1 allocation of Test versus sham Control with a Pharmacokinetics (PK) arm of 15 non-randomized subjects. For the randomized portion of the study, 147 subjects will be randomized and treated with either the Optilume BPH Catheter System (Test device) or a sham device at up to 20 clinical sites. Five (5) of the 20 clinical sites will also be participating in the PK arm of the study. Fifteen (15) non-randomized subjects will be treated in the PK arm. Total enrollment for the randomized portion and the PK arm will be 162 subjects.

Subjects will be randomized in a 2:1 ratio to Test:Control. Randomized subjects will be stratified by IPSS score (≤ 19 or > 19) prior to study enrollment as shown in Figure 1. All subjects will be followed up post-treatment at Foley removal, 14 days (telephone call), 30 days, 3 months, 6 months and 12 months. In addition, all subjects treated with Test device will be followed annually through 5 years.

Individual site enrollment in the randomized portion of the study may not exceed 30% of the total randomized study enrollment.

The study duration is anticipated to be approximately 8 years.

The schematic of the study treatment arms and cross-over are shown in **Figure 1**.

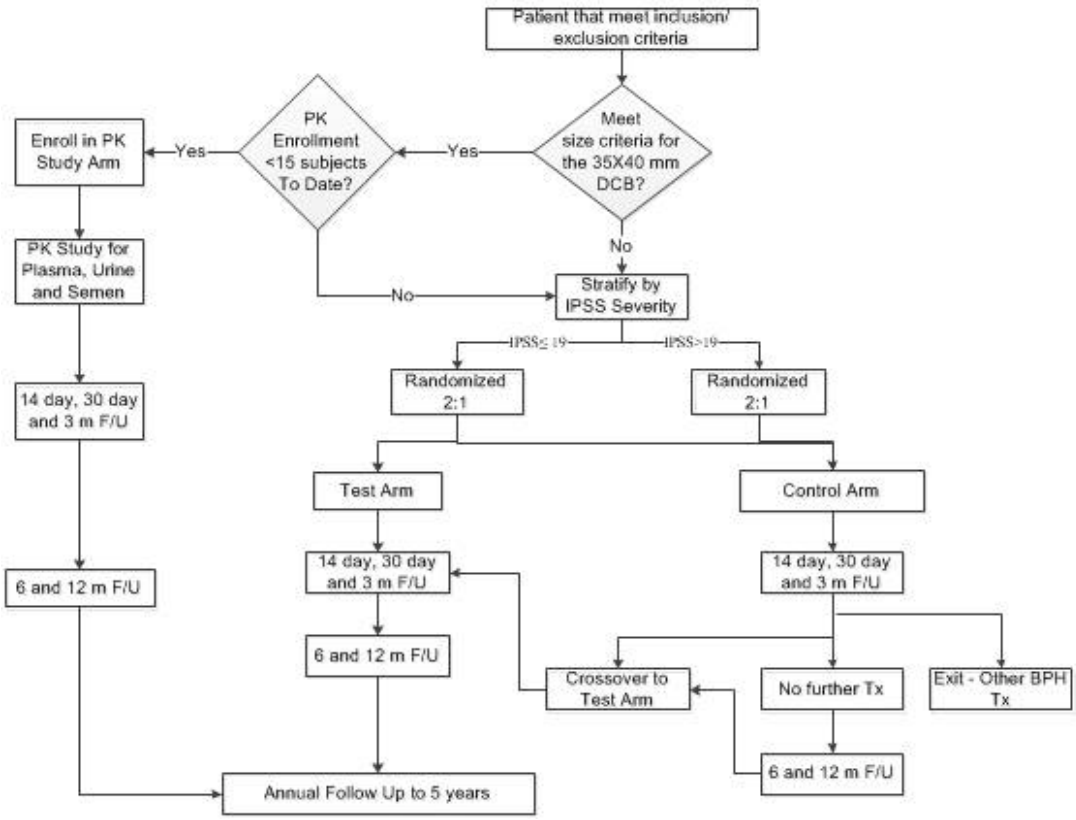

FIGURE 1: STUDY DESIGN FLOWCHART

At 3 and 12 months after randomization, the efficacy of the Test device will be demonstrated by comparison to the Control group. Additionally, all subject treated with the Test device will continue to complete questionnaires and assessments through 5 years post procedure. All subjects randomized to the Test arm, the PK Study arm, or who later receive the Test device, will be monitored for serious device- or procedure-related adverse events.

3.2 Test and Control Arm Description

Subjects randomized to the Test arm will be treated with the Urotronic Optilume BPH Catheter System (Test device).

Subjects randomized to the Control arm will receive a sham device. A cystoscopy will be performed with a flexible cystoscope followed by insertion of the 21 Fr Optilume BPH, Prostatic Pre-dilation Catheter (Sham device) within the sheath. The catheter hub of the sham device has been modified to prevent inflation of the balloon.

3.3 Cross-over Description

If a subject's condition did not improve after the assigned index treatment, and the subject and physician are considering alternative treatments, the subject's treatment assignment may be unblinded after the 3-Month follow-up visit. Under this condition, the subjects in the Control

arm may be allowed to consider crossing-over to the Test arm of the study. Any subjects who cross-over or receive alternative treatments before completing the 12-Month follow-up visit will automatically be considered a treatment failure.

The safety and efficacy data collected post-cross-over will be summarized separately.

3.4 Pharmacokinetics Arm Description

In a subset of 15 non-randomized subjects at selected sites, small amounts of blood, urine and semen will be collected for pharmacokinetic testing and sperm quality testing.

All subjects who are enrolled in the Pharmacokinetics arm of the study will be treated per the study protocol and device IFU with the 35 x 40 mm drug coated balloon (DCB) Test device (largest diameter/length).

3.5 Primary Endpoints

Study success will be based on the success of the primary efficacy endpoint 1 and secondary endpoint 1.

3.5.1 Primary Efficacy Endpoint 1

The primary efficacy endpoint is the change in IPSS. The endpoint will be evaluated based on the change or reduction in IPSS from baseline. The reduction in IPSS at 3 months for subjects randomized to the Control arm will be compared to the reduction in IPSS at 12 months for subjects randomized to the Test arm.

For the this endpoint to be a success, the observed improvement in IPSS at 12 months post-treatment in the Test arm must be at least 25% greater than that of the Control arm at 3 months based on the following statistical hypotheses:

$$HH_{oo}: \pi\pi_{TTTTTTTT,12mmoo} - 1.25\pi\pi_{CCooCCTTCCooCC,3mmoo} \leq 0$$

$$HH_{AA}: \pi\pi_{TTTTTTTT,12mmoo} - 1.25\pi\pi_{CCooCCTTCCooCC,3mmoo} > 0$$

where:

$\pi\pi_{TTTTTTTT,12mm}$ is the mean reduction in IPSS from baseline for the Test arm at 12 months

$\pi\pi_{CCooCCTTCCooCC,3mm}$ is the mean reduction in IPSS from baseline for the Control arm at 3 months

The statistical hypothesis test for the primary efficacy endpoint will be based on a two-sample t-tests at the one-sided 0.025 alpha level (equivalent to a two-sided 0.05 alpha level). Incorporation of the criterion that the Test arm results must be at least 25% greater than that of the Control arm will be done via multiplication of the individual control group values by 1.25. This produces a mean scaled by an increase of 25% and can be directly applied to the standard

two-sample t-test for comparison of means. In pseudo SAS code, this can be performed as follows:

```
data pinnacle;
set pinnacle;
if group = 'Control' then primary_endpoint = 1.25*ipss_change;
else if group = 'Treatment' then primary_endpoint = ipss_change;
run;

proc ttest data = pinnacle sides=L alpha=0.025;
class group;
var primary_endpoint;
run;
```

In the event there is evidence the normality assumption of a t-test do not hold based on a Shapiro-Wilks test, an alternative non-parametric test, the Mann-Whitney U test (Wilcoxon rank-sum Z test with a continuity correction, SAS PROC NPAR1WAY) will be used.

3.5.2 Primary Safety Endpoint 1

The primary safety endpoint is a composite endpoint of major device-related serious complications, defined as any of the following events through 12 months:

- Device-related rectal fistula or gastrointestinal (GI) fistula
- Device-related formation of fistula between the rectum and urethra
- Device-related new onset severe urinary retention lasting > 14 consecutive days post-healing
- Device-related unresolved new onset stress urinary incontinence by 90 days
- Device-related bleeding requiring transfusion
- Device-related urethra or prostatic capsule rupture requiring surgical intervention

The primary safety endpoint will be analyzed with descriptive statistics and nominal 95% confidence interval; there is no formal statistical hypothesis test planned for this endpoint.

Analysis will be on events that occur prior to the window close date for the scheduled 12-month study visit.

3.6 Secondary Endpoints

The following secondary endpoints will be tested against the corresponding null hypotheses using a sequential gatekeeping strategy to control the type I error rate. Significance of the secondary endpoints will be based on sequential testing of the endpoints, proceeding in the order

listed, until a non-significant result is reached at which time no further claims of significance will be made. Additionally, significance for secondary endpoints will only be claimed if success for the hypothesis test for the primary endpoint is achieved. Each test will be performed at the one-sided 0.025 alpha level. Any other p-values or confidence intervals will be nominal and not adjusted for multiple testing; it is recognized this may limit the ability to use such p-values in product labeling.

For those endpoints without a specified hypothesis test, descriptive statistics will be used in reporting outcomes for other efficacy endpoints. Continuous variables will be summarized with means or medians, standard deviations. Adverse events, protocol deviations and device malfunctions will be summarized with descriptive statistics.

3.6.1 Secondary Endpoint 1: Average IPSS improvement in the Test arm at 12 months

The average IPSS improvement in the Test arm from baseline to 12 months will be assessed. The mean percent reduction in IPSS at 12 months for subjects randomized to the Test arm will be compared to a performance goal of 30%. For this endpoint to be a success, the mean percent reduction in IPSS from baseline to 12 months post-treatment in the Test arm must be statistically greater than 30% based on the following statistical hypotheses:

$$HH_{oo}: \pi\pi_{TTTTTTTT,12mmoo} \leq 30\%$$

$$HH_{AA}: \pi\pi_{TTTTTTTT,12mmoo} > 30\%$$

where:

$\pi\pi_{TTTTTTTT,12mm}$ is the mean percent reduction in IPSS from baseline for the Test arm at 12 months.

3.6.2 Secondary Endpoint 2: Percentage of responders at 3 months

A responder is defined as a subject who has an IPSS improvement ≥ 30% post-treatment as compared to baseline.

The responder rate at 3 months of subjects randomized to the Test arm will be compared to the responder rate of subjects in the Control arm.

$$H_{o3}: \mu_{1test,\ 3\ mo} \leq \mu_{1control,\ 3\ mo}$$

$$H_{a3}: \mu_{1test,\ 3\ mo} > \mu_{1control,\ 3\ mo}$$

where:

$\mu_{1test,\ 3\ mo}$ is the percentage of responders for the Test arm at 3 months

$\mu_{1control,\ 3\ mo}$ is the percentage of responders for the Control arm at 3 months

The test will be based on a two-sample exact test of binomial proportions.

3.6.3 Secondary Endpoint 3: Durability – Percentage of responders

A responder is defined as a subject who has an IPSS improvement of ≥ 30% post-treatment as compared to baseline.

The responder rate at 12 months of subjects randomized to the Test arm will be compared to the responder rate at 3 months of subjects in the Control arm.

$$H_{o2}: \mu_{2test,\ 12\ mo} \leq \mu_{2control,\ 3\ mo}$$

$$H_{a2}: \mu_{2test,\ 12\ mo} > \mu_{2control,\ 3\ mo}$$

where:

$\mu_{2test,\ 12\ mo}$ is the percentage of responders for the Test arm at 12 months

$\mu_{2control,\ 3\ mo}$ is the percentage of responders for the Control arm at 3 months

The test will be based on a two-sample exact test of binomial proportions.

3.6.4 Secondary Endpoint 4: Qmax

The maximum urinary flow rate (Qmax) provides an objective, non-invasive measurement of the subject's BPH symptoms. The change or increase in Qmax at 12 months for all treated subjects randomized to the Test arm will be compared to the change or increase in Qmax at 3 months for subjects randomized to the Control arm.

$$H_{o1}: \mu_{3test,12m} \leq \mu_{3control,3m}$$

$$H_{A1}: \mu_{3test,12m} > \mu_{3control,3m}$$

where:

$\mu_{3test,\ 12\ mo}$ is the change in Qmax for the Test arm at 12 months

$\mu_{3control,\ 3\ mo}$ is the change in Qmax for the Control arm at 3 months

The test will be based on a two-sample t-test of means.

3.6.5 Ancillary Endpoints and Other Analyses

The following Ancillary Endpoints are to provide additional characterization of the safety and efficacy of the Optilume BPH Catheter System in the treatment of LUTS/BPH.

- A1: Additional responder analyses with a responder defined as IPSS improvement of 35%, 40% and 50%
- A2: Change in PVR
- A3: Change in sexual function
- A4: Change in BPH-II

- A5: Change in quality of life
- A6: Change in pain score
- A7: Procedure parameters
- A8: Change in Qmax

3.7 Pharmacokinetics Arm

Descriptive statistics will be used to analyze all data from the PK arm. Nominal confidence intervals and p-values may be used without adjustment, but these will not be generated for the purposes of supporting labeling claims.

3.8 Randomization and Blinding

3.8.1 Randomization Plan

Subjects will be randomized in a 2:1 allocation of Test vs Control. Randomization will be stratified by IPSS severity at baseline, with separate permuted blocks for study sites.

Those subjects who do not meet inclusion/exclusion criteria after baseline evaluation will be counted as screening failures and will not be enrolled in the study. Each subject will be randomized prior to initiation of the test/control procedure. Only randomized subjects will be considered enrolled and evaluable.

3.8.2 Blinding

The treating physicians in the Randomization arm of the study will be blinded. A blinded person at the study site will be conducting all study follow-up visits and administering all subject questionnaires through the 12-Month follow-up visit to maintain blinding.

Subjects in the Randomization arm of the study will be blinded to the treatment received through their 12-Month follow-up visit, at which point the subjects will be unblinded. Blinding will be broken only to protect the subject's health. If a non-urgent clinical need requires that the subject be unblinded prior to applicable follow-up and if time allows, the physician will notify the Sponsor prior to unblinding the subject.

3.9 Cross-over of Control Subjects

If a subject’s condition did not improve after the assigned index treatment, and the subject and physician are considering alternative treatments, the subject’s treatment assignment may be unblinded after the 3-Month follow-up visit. After the 3-Month follow-up visit, the subjects in the Control arm who are symptomatic, and continue to meet the study entrance criteria (all inclusion criteria and no exclusion criteria) will be given the option to receive treatment with the Test device and be followed according to the full Test arm follow-up schedule. A cross-over treatment must be complete before the end of the subject’s 12-Month follow-up visit window.

The safety and efficacy data collected post-cross-over will be summarized separately. Cross-over subjects will be followed out to 5 years, starting from the cross-over procedure, under the same subject ID.

4 Study Sample Justification

4.1 Primary Efficacy Endpoint Sample Size Considerations

Sample size for the primary efficacy endpoint was based on the following assumptions:

- Statistical power of 90%
- One-sided 0.025 alpha
- 2:1 randomization allocation
- A common standard deviation of 7
- Mean improvement in the Control arm of 6 points and mean improvement in the Test arm of 12.5 points
- 10% loss of follow-up rate

Based on these assumptions, a sample size of 132 evaluable subjects (88 Test and 44 Control) will provide greater than 90% power. Assuming a 10% loss of follow-up rate, the randomization sample size is 147 (approximately 98 Test and 49 Control).

With the 15 subjects in the PK arm, the total number of subjects required is 162 (147+15).

5 Analysis Populations Set

All subjects enrolled in the study (including those withdrawn from the investigation or lost to follow-up) will be accounted for and documented.

The primary endpoint analyses and the secondary endpoints will be performed on the intent-to-treat (ITT) population, under which all randomized subjects will be included for the analysis, regardless of whether or not the subjects received the treatment to which they were randomized during the appropriate time period.

In addition to the ITT analysis, the primary endpoint analyses will also be performed on the as treated population (i.e. subjects analyzed based on the treatment actually received) and per-protocol (PP) population, (i.e., subjects treated and followed per the protocol) where appropriate.

The per protocol analysis set will exclude subjects with the following significant protocol violations:

- Subjects with significant violations of the inclusion and exclusion criteria including the following:
    - Subject's rights are violated or did not give consent and subject data was requested to be excluded by IRB

- Subject's obstructive symptoms are primarily due to stricture or bladder neck contracture in addition to or rather than BPH
- Subject had UTI at the time of treatment
- Subject had radiation of the pelvic region
- Subject had a previous surgery for BPH
- Subject did not complete wash out of drugs
- Subject had bladder or sphincter dysfunctions or anything else that would confound the results
- Subject had a psychiatric or cognitive disorder that prevents him from adequately answering the study questionnaires
- Subjects who are not available or are unblinded before the Primary Endpoint evaluation as required by the cohort unless their exit is due to treatment failure. Subjects who are considered treatment failures (and received non-study treatment or prematurely crossed over) will be assigned the worst possible values for the missing data in the analyses.

Other additional criteria may be added to the list to accommodate for unforeseen events that occurred during the conduct of the trial that resulted in noteworthy study protocol violations. Significant protocol violations will be summarized by category and by site.

All other ancillary endpoints will be analyzed for those subjects treated by the Test or Control device irrespective of the original randomization assignment (As Treated); subjects will be grouped based on the treatment received.

6 Poolability of Data

6.1 Sites

This study is designed and conducted as a multicenter randomized-control clinical trial. Data from all the sites will be pooled. All subjects will be treated and evaluated following the same protocol to ensure generalizability of the study results.

Heterogeneity in the treatment effect for the primary efficacy endpoint will be assessed via linear regression models with covariates for site, treatment group, and the interaction of site and treatment group. A p-value for the interaction term <0.15 will suggest evidence of variation in the treatment effect and will trigger additional exploratory analyses to attempt to further quantify and explain the variation. Sites with fewer than 4 subjects will be combined into a single super-site for these analyses.

6.2 Subgroup Analysis

Subgroup analyses will be performed for the 3-month primary efficacy endpoint to understand potential variation in the treatment effect. Statistical methods will follow the same as those outlined for poolability of sites (Section 7.1). Subgroups will be defined based on the following baseline factors:

- Subject age at the time of randomization (≤ median age vs. > median age)
- Subject race (White vs. Non-White)
- Baseline IPSS Score (≤ 19 vs. > 19, i.e. moderate vs. severe)
- Baseline prostate volume (≤ 40 g vs. > 40 g)
- Baseline presence of Intravesical Prostatic Protrusion (IPP)

7 Missing data

Every effort will be made to reduce the incidence of missing data. All available data on subjects who drop out during the study will be included.

7.1 Missing Data Handling Rules

7.1.1 For the Primary Efficacy Endpoint

Subjects who withdraw due to adverse events, lack of effectiveness, or who start other treatment to control BPH symptoms will be counted as having no improvement in IPSS score. For other subjects are missing data for the primary efficacy endpoint, data will be imputed via multiple imputation, separately by randomized treatment group. The following covariates that will be included in the multiple imputation model in the order listed:

- Sex
- Age,
- Race/ethnicity
- Baseline IPSS
- Baseline prostate volume
- Baseline presence of Intravesical Prostatic Protrusion

Missing covariates will be imputed via full conditional specification imputation via a logistic regression for categorical variables or linear regression for continuous variables in order to produce a data set with a monotone missing pattern. If there are fitting issues with the covariate imputation process, individual covariates may be omitted to still allow for imputation of the endpoint for the full randomized cohort. Alternatively, if there are fitting issues due to the separation or quasi-separation of data points, an augmented likelihood approach will be used. Following imputation of missing covariates if needed, imputation of the endpoint will be based on linear regression with 100 data sets. Inferences from the data sets will be combined to produce a single estimated treatment effect and p-value for the hypothesis test.

7.1.2 For the Primary Safety Endpoint

Unless there is evidence of occurrence of a primary safety endpoint event, subjects with missing data for the primary safety endpoint are presumed to not have experienced a primary safety endpoint event.

All subject data that are available on subjects who drop out during the course of the study will be included.

7.2 Multiple Responses

If a subject has multiple baseline evaluations, the last baseline value prior to the procedure will be used in the analysis.

If a subject has multiple post-baseline evaluations within the same time visit window period, the latest results will be used for the subject in all analyses unless there is scientific valid reason(s) to exclude one or more of the evaluations.

7.3 Sensitivity Analyses for Missing Data

For the primary efficacy endpoint, sensitivity analyses will be performed to evaluate the impact of missing data on study conclusion. This will include analysis based on only observed data, as well as analysis based on a tipping point analysis approach. For the tipping point analysis, every combination of best/worst case values will be imputed to examine under what circumstances the results change. Best/worst case values used for this will be the largest improvement/worst improvement observed in the trial.

For the primary safety endpoint, tipping point analyses will be performed to evaluate the impact of missing data on study conclusion. Every combination of having/not having a primary safety endpoint will be imputed and examined to determine under what circumstances the results change.