STATISTICAL ANALYSIS PLAN

A PROSPECTIVE, SINGLE-ARM, MULTICENTER STUDY TO EVALUATE EFFECTIVENESS AND SAFETY OF TYMPANOSTOMY TUBE PLACEMENT USING THE TULA IONTOPHORESIS AND TUBE DELIVERY SYSTEMS FOR CHILDREN IN AN OFFICE SETTING (OTTER; in-Office Tympanostomy Tube placemEnt in childRen)

Document Number: CPR007004, Revision E Effective Date: 11/29/2017

Protocol Number: CPR007001 Rev E

[redacted]

Signature: _____

Date: _____

[redacted]

Signature: _____

Date: _____

[redacted]

Signature: _____

Date: _____

TABLE OF CONTENTS

1.0 Scope

This is the Statistical Analysis Plan (SAP) for the final analysis of data collected under Protocol Number CPR007001.

2.0 Trial Objectives

The primary objective is to confirm safety and effectiveness (procedural success and tube placement tolerability) of tympanostomy tube placement using the TULA Iontophoresis and Tube Delivery Systems for children in an in-office procedure.

The trial will also evaluate anesthesia effectiveness, tube patency and tube retention.

3.0 Trial Endpoints

3.1 Co-Primary Endpoints

- **Procedural Success:** Proportion of subjects in the pivotal cohort with successful placement of Tusker Medical tympanostomy tubes in all indicated ears in an office procedure.

- **Tube Placement Tolerability:** Mean subject-reported pain score following TDS tube placement using the Faces Pain Scale-Revised (pivotal cohort children ages 5 and older)

3.2 Safety Endpoint

Occurrence of adverse events, by subject.

3.3 Secondary Efficacy Endpoints

Tube Patency: Tube Patency is the proportion of subjects in the pivotal cohort, in which a Tusker Medical tube was successfully placed, with functionally patent tube(s) in all successfully treated ears at the 3-week post-procedure follow-up visit.

Tube Retention: Tube Retention is the proportion of subjects in the pivotal cohort, in which a Tusker Medical tube(s) was successfully placed, with presence of a Tusker Medical tube across the TM in all successfully treated ears at the 3-week post-procedure follow-up visit.

Anesthesia Effectiveness: Proportion of subjects in the pivotal cohort, who completed iontophoresis in all indicated ears, with adequate anesthesia for TT placement in all treated ears as determined by physician's evaluation of TM anesthesia prior to tube placement.

[redacted]

[redacted]

[redacted]

[redacted]

4.0 Trial Design

This study is a prospective, single-arm, multicenter evaluation of safety and effectiveness of iontophoresis using the IPS and TDS for TT placement in-office for children indicated for tympanostomy tube placement.

The study consists of OR lead-in, in-office lead-in and pivotal cohort subjects. [redacted] Study enrollment and procedures will occur in two stages.

Stage 1 (OR Lead-In Cohort): Stage 1 consists of tube placement procedures in the OR under general anesthesia. [redacted] each investigator will enroll 2 lead-in subjects in the operating room (OR) undergoing tube placement using the Tube Delivery System to place the tympanostomy tube(s) under general anesthesia (ie, no lidocaine iontophoresis). Stage 1 is comprised of the OR Lead-In Cohort up to 100 subjects).

Stage 2 (In-Office Lead-In Cohort and Pivotal Cohort): Stage 2 consists of in-office procedures using both the lidocaine Iontophoresis System and the Tube Delivery System for the in-office tube placement procedure, including both the Office Lead-In Cohort and the Pivotal Cohort. [redacted] each investigator will enroll 2 lead-in subjects in-office undergoing tube placement using the TDS with local anesthesia facilitated by the Iontophoresis System (up to 100 subjects in the Lead-In Office Cohort). Investigators may enroll into the Pivotal Cohort following completion of their lead-in office procedures. The pivotal cohort will include 222 children (ages 6 months through 12 years, inclusive) with evaluable data indicated for tympanostomy tube placement enrolled at up to 25 investigational centers (minimum of 15 centers).[1] The pivotal phase will include 102 children ages 5 years and older, and 120 children with evaluable data less than 5 years of age.

4.1 Design of the Pivotal Phase

The pivotal phase of the study aims to confirm safety, and determine whether the in-office tube placement procedure achieves an acceptable procedural success rate and acceptable tube placement tolerability as established by performance goals. The pivotal cohort will consist of a sample of 222 children including 102 children age 5 through 12 years old, and 120 children less than 5 years of age.

[1] Subjects who are consented for the study will be considered enrolled. Throughout the protocol and SAP, enrollment numbers will reflect consented enrolled subjects with evaluable data per SAP.

The study includes two primary endpoints. The first primary endpoint, Procedural Success, will be the proportion of pivotal cohort subjects with Tusker Medical tympanostomy tubes placed in all indicated ears in an office procedure. The statistical test for the Procedural Success primary endpoint is a hypothesis test against a procedural success performance goal. The second primary endpoint, Tube Placement Tolerability, will be the mean subject-reported pain score following TDS tube placement using the Faces Pain Scale-Revised (FPS-R) (pivotal cohort children ages 5 and older only). The statistical test for the Tube Placement Tolerability primary endpoint is a hypothesis test against a Tube Placement Tolerability performance goal.

[redacted] The study design thus employs a Bayesian hierarchical gatekeeper strategy whereby a successful first test will permit the trial to continue to the second statistical test. The test against the Procedural Success performance goal occurs in two steps:

1) The initial test against the performance goal is conducted on the older children in the pivotal cohort, once the study has accrued evaluable data for 102 children ages 5 through 12 years of age. [redacted] The older child group will be tested against the procedural success primary performance goal using a Bayesian hierarchical algorithm designed to borrow information from the enrolled younger children as appropriate given the performance results at the time of the first test. [redacted] If the older group achieves the performance goal, the second test is applied as in step 2.

2) Once the younger group achieves 120 evaluable subjects, they may be tested against the Procedural Success performance goal. The 102 older (5-12 years old) children are included in the Bayesian hierarchical model used to evaluate the younger subjects against the performance goal.

The second primary endpoint, Tube Placement Tolerability, will be tested for the older group only at step 1, above, at time of the initial test for Procedural Success for the 102 older children.

[redacted]

5.0 Study Cohorts

5.1 Lead-In OR Cohort

The Lead-In OR Cohort will consist of all children who underwent a lead-in procedure in the OR. The Lead-In OR Cohort safety and effectiveness data will be analyzed and presented separately.

5.2 Lead-in Office Cohort

The Lead-In Office Cohort will consist of all children who underwent a lead-in procedure in-office. The Lead-In Office Cohort safety and effectiveness data will be analyzed and presented separately.

5.3 Pivotal Cohort

The Pivotal Cohort will consist of children who underwent an in-office procedure using the TULA System (not including Lead-In subjects). The Pivotal Cohort safety and effectiveness data will be analyzed and presented separately.

6.0 Analysis Sets

The following analysis sets are defined for the Pivotal Cohort.

6.1 Full Analysis Set (FAS)

All subjects in whom lidocaine is introduced into the ear canal and for whom iontophoresis was initiated will be included in the FAS.

6.2 Per Protocol

The Per Protocol Set is a subset of the FAS and includes all FAS subjects without major protocol deviations. Major protocol deviations include:

• Eligibility violations
• Procedural deviations with the potential to affect anesthesia efficacy (eg, abbreviation of iontophoresis or use of incorrect local anesthesia drug).

6.3 Safety Analysis Set

The Safety Analysis Set will include all pivotal cohort subjects in which the lidocaine solution is introduced into the ear canal.

7.0 Subject Disposition

Subject disposition will be summarized with counts and percentages. Categories summarized will include the number of subjects who signed informed consent, number of screen failures (with reasons), and the number enrolled, completed, and discontinued, as well as reasons for discontinuation. The summary of disposition will also include the number and percentage of subjects included in each analysis set described in Section 6, and for each age strata and cohort.

8.0 Demographic Characteristics and Medical History

Demographic characteristics including age, sex and race/ethnicity will be descriptively summarized for all enrolled subjects in each of the defined study cohorts. Significant medical history will also be summarized for all enrolled subjects in each of the defined study cohorts. Demographic characteristics and medical history will be presented separately for each of the Analysis Sets defined for the Pivotal Cohort, and for each age strata.

10.0 Primary Endpoint Analyses

10.1 Procedural Success

There are two primary endpoints for the study. The first primary efficacy endpoint is Procedural Success defined as the proportion of subjects in the pivotal cohort FAS achieving procedural success evaluated against a performance goal of 68%.

If both ears of a subject are indicated for TT placement, the procedure must be successful in both ears for the procedure to be counted as a success on a subject-level for the Procedural Success primary efficacy endpoint. If either ear does not receive a tube, the procedure is not considered a success. If a subject requires only a unilateral tube, the procedure is considered a success when the indicated ear has received the tube. The procedural success definition is aligned with the desired clinical outcome; the child has Tusker tubes placed in all indicated ears while avoiding general anesthesia and an OR visit. The statistical test for the pivotal cohort procedural success primary endpoint is a Bayesian hypothesis test against a procedural success performance goal.

1) The initial test against the performance goal is conducted on the older children in the pivotal cohort, once the study has accrued evaluable data for 102 children ages 5 through 12 years of age.

The primary hypothesis is

Test 1:
H_0: $p_{5+} \leq 68\%$ vs. H_a: $p_{5+} > 68\%$,

where p_{5+} is the proportion of subjects 5 through 12 years old achieving procedural success.

This hypothesis will be tested using a Bayesian hierarchical model evaluating whether the posterior probability of H_a is at least 0.975, which is analogous to a classical hypothesis test at the 0.025 significance level in the FAS comprising only subjects 5 through 12 years old. This Bayesian analysis applies the test to the older group, while borrowing data from all enrolled evaluable subjects less than 5 years old, as appropriate given the performance results at the time of the first test. [redacted] If the older group achieves the performance goal, the second test is applied as described in step 2. Procedural success will also be summarized in the Per Protocol set comprising subjects 5 through 12 years of age.

2) Once the younger group achieves 120 evaluable subjects, they may be tested against the Procedural Success performance goal. The second test evaluates the younger subjects against the performance goal while borrowing data from the 102 older subjects using a Bayesian hierarchical model as appropriate given the performance of the older group.

Test 2:
H_0: $p_{<5} \leq 68\%$ vs. H_a: $p_{<5} > 68\%$,

where $p_{<5}$ is the proportion of subjects in the younger group (6 months through <5 years old) achieving procedural success.

Test 2 is also performed using a Bayesian hierarchical model, requiring a 97.5% posterior probability of H_a. [redacted]

[redacted]

10.2 Tube Placement Tolerability

The second primary endpoint is Tube Placement Tolerability defined as mean subject-reported pain score following TDS tube placement using the Faces Pain Scale-Revised (FPS-R) (pivotal cohort children ages 5 and older only).

[redacted]

If both ears of a subject are indicated for tympanostomy tube placement, the tube placement pain score is collected after the second ear tube placement is completed (one tube placement pain score per subject). The FPS-R score will be collected after any manipulations required to place the tube so that the pain score fully reflects the activities required to place the tube. [redacted]

The mean pain score for subjects in the pivotal cohort FAS, for whom tube placement was successful, will be evaluated against a performance goal of 4.2 pain score.

The test against the performance goal is conducted on only the older children in the pivotal cohort, once the study has accrued evaluable data for 102 children ages 5 through 12 years of age.

The statistical test for the Tube Placement Tolerability primary endpoint is a hypothesis test against a Tube Placement Tolerability performance goal. [redacted] Only completers, ie, subjects with FPS-R scores after successful tube insertion in all indicated ears, will be included in the primary analysis.

The primary hypothesis is

$$H_o: \ p_{5+} \geq 4.2$$
$$\text{vs.}$$
$$H_a: p_{5+} < 4.2$$

Where p_{5+} is the mean Tube Placement FPS-R for subjects 5 through 12 years old.

This hypothesis will be tested using a standard classical t-test at the 0.025 significance level in the FAS comprising only subjects 5 through 12 years old. [redacted]

[redacted]

11.0 Secondary Endpoint Analyses

[redacted] The secondary endpoints (Tube Patency, Tube Retention and Anesthesia Effectiveness) are functions of the tube and iontophoresis device performance and are not anticipated to vary by subject age. Therefore, the second test on the full (pooled) pivotal cohort is appropriate.

11.1 Tube Patency

Tube Patency is the proportion of subjects in the pivotal cohort FAS, in which a Tusker Medical tube was successfully placed, with functionally patent tubes in all successfully treated ears at the 3-week post-procedure follow-up visit.

The statistical test for the pivotal cohort Tube Patency secondary endpoint is a hypothesis test for the proportion of subjects with functional tube patency at the 3-week follow-up visit against a Tube Patency performance goal. The mid-P[2,3,4] method for a single proportion will be used to test

$$H_0: p \leq 80\% \text{ vs. } H_a: p > 80\%$$

This endpoint will be considered met if the p-value from the test is less than 0.025.

[2] Berry G, Armitage P. Mid-p confidence intervals, a brief review. The Statistician, 1995; 44:417-423.
[3] Lynderen S, Laake P. Power comparison of two-sided exact tests for association in 2x2 contingency tables using standard, mid p, and randomized test version. Statistics in Medicine, 2003; 22:3859-3871.
[4] Lancaster HO. Significance tests in discrete distributions. JASA, 1961; 56:223-234.

11.2 Tube Retention

Tube Retention is the proportion of subjects in the pivotal cohort FAS, in which a Tusker Medical tube was successfully placed, with presence of a Tusker Medical tube across the tympanic membrane in all successfully treated ears at the 3-week post-procedure follow-up visit.

The statistical test for the pivotal cohort Tube Retention secondary endpoint is a hypothesis test for the proportion of subjects with a Tusker tube at the 3-week follow-up visit against a Tube Retention performance goal. A mid-P method for a single proportion will be used to test

$$H_0: p \leq 88\% \text{ vs. } H_a: p > 88\%$$

This endpoint will be considered met if the p-value from the test is less than 0.025.

11.3 Anesthesia Effectiveness

Anesthesia Effectiveness is the proportion of subjects in the pivotal cohort FAS, who completed iontophoresis for all indicated ears, with adequate anesthesia for TT placement in all treated ears as determined by physician's evaluation of TM anesthesia prior to TT placement.

The statistical test for the Anesthesia Effectiveness secondary endpoint is a hypothesis test for the proportion of subjects with adequate anesthesia for TT placement in all ears completing iontophoresis against an anesthesia effectiveness performance goal. The mid-P method for a single proportion will be used to test

$$H_0: p \leq 85\% \text{ vs. } H_a: p > 85\%$$

This endpoint will be considered met if the p-value from the test is less than 0.025.

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

14.0 Safety Endpoint Analyses

14.1 Adverse Events

Adverse events (AEs) reported during the study will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The number and percentage of subjects reporting adverse events will be summarized by MedDRA system organ class and preferred term by cohort and in total.

The study will present the totality of safety data, by subject including:

- Serious device, procedure, and drug-related events
- Other serious events
- Non-serious device, procedure, drug-related events
- Other non-serious events

For all AE tables, a subject reporting the same AE more than once will be counted once when calculating the number and percentage of subjects with that particular event. If a subject reports the same AE more than once or has the same AE on multiple occasions, the maximum severity grade recorded for the event will be presented.

Overall summaries of serious and non-serious AEs will be presented. The overall summaries will include the number and percentage of subjects experiencing AEs classified by treatment, relationship to device/procedure/study drug, MedDRA SOC (System Organ Class) and preferred term. Summaries of AEs will also be presented by severity for each treatment group including the number and percentage of subjects experiencing AEs classified by relationship to device/procedure/study drug, SOC and preferred term.

[redacted]

[redacted]

Safety data will be presented separately as follows:
• Lead-In OR Cohort
• Lead-In Office Cohort
• Pivotal Cohort

[redacted]

Complete subject listings of all AEs will be provided. For each AE the following will be specified: start and stop dates, severity grade, MedDRA SOC, relationship to study device, relationship to procedure, relationship to study drug, action taken, outcome of the adverse event and seriousness. Any unanticipated adverse events will be noted.

14.2 Erythema Assessment at Return Electrode Location

Erythema observed for the skin under the return electrode will be categorized by grade level (0 through 4) prior to and immediately after iontophoresis. Erythema findings will be summarized by number and proportion of subjects in the Safety Analysis Set for each erythema grade level for each Cohort. Number and proportion of subjects reporting discomfort at the return electrode location will be summarized.

14.3 Audiometric Assessments

Change in audiometric measurements (shifts) from baseline to the 3-week follow-up visit for subjects in the Safety Analysis Set will be presented using descriptive statistics (mean, SD, median, minimum and maximum). Shifts in air conduction (AC) and bone conduction (BC) hearing thresholds will be reported for each frequency (500, 1000, 2000 and 4000 Hz, and air conduction only at 8000Hz) and for each Cohort on a by-ear basis (treatment ears only). Shifts from baseline to each follow-up visit in air-bone gap (air conduction minus bone conduction) will also be reported for each frequency (500, 1000, 2000 and 4000 Hz) for each Cohort on a by-ear basis. Shifts in air conduction pure tone average (AC PTA), bone conduction PTA (BCA) and air-bone gap averages will be presented for each follow-up time point for each Cohort. For all PTA measurements, the PTA will be calculated by adding the threshold levels in decibels (dB) obtained at each frequency tested divided by number of frequencies tested.

Shift tables for air conduction pure tone averages greater than 15 dB (>15dB) by ear will be calculated, and presented for each Cohort for each follow-up time point.

Degree of hearing loss severity will be defined according to the protocol (slight through profound) and presented by subject for each Cohort and by relationship to study device, procedure or study drug.

Due to the differences in audiometric requirements for children dependent on age, all audiometry results will be presented by subjects less than 4 years of age, greater than 4 years of age, and overall.

14.4 Tympanometry

Tympanometry results are used for assessment of tube patency and for changes in middle ear condition. Tympanometry will be performed for indicated ears at screening and at each follow-up visit. The number and percentage of ears with tympanogram types A, B or C will be presented for all subjects in the Safety Analysis Set at screening and at the follow-up visits.

14.5 Otoscopic Assessments

Otoscopic examination will be performed for all ears at screening and at each follow-up visit. The number and percentage of indicated ears (at screening) and treated ears (at procedure and follow-up visits) in each Cohort with abnormal findings for the external acoustic meatus or tympanic membrane will be presented for all subjects in the Safety Analysis set at screening, procedure and at the follow-up time visits.

14.6 Cranial Nerve Function

Cranial nerve function physical examination will be performed for all in-office procedure subjects at screening, post-procedure and at the 3-week follow-up visit. The number and percentage of subjects with abnormal findings in the Safety Analysis set for each In-Office Cohort will be presented for the screening, procedure and the 3-week follow-up visit.

15.0 Prior and Concomitant Medications

Prior and concomitant medications including prescription and over-the-counter drugs will be recorded. Medications will be coded to identify the drug class and preferred drug name.

Concomitant medications will include all medications that started on or after the procedure through the final follow-up visit. Prior medications will include all medications that started within 28 days prior and stopped prior to the day of procedure.

Prior and concomitant medication data will be presented in a data listing for safety subjects.

16.0 Other Endpoint Analyses

For continuous and ordinal variables, descriptive statistics will include, at a minimum, the number of observations, mean, standard deviation, median, minimum, and

maximum. For categorical variables, the number and percentage of subjects will be presented.

17.0 Sensitivity Analyses and Handling of Missing Data

The number and proportion of subjects eligible for and compliant with all follow-up examinations will be presented. Subjects who withdraw from the study will be tabulated with the reasons for the withdrawal. Subjects who initiate the procedure, regardless of tolerability outcome or anesthesia effectiveness, will be encouraged to return for all follow-up assessments.

For the Procedural Success primary endpoint, all subjects who initiate iontophoresis are included in the primary analysis of this primary endpoint. Therefore, no missing data analyses for the Procedural Success primary endpoint are planned.

For the Tube Placement Tolerability primary endpoint, all subjects in the pivotal cohort FAS for whom tube placement was successful are included in the primary analysis of this primary endpoint. [redacted]

Tube Retention secondary endpoint: Tube Retention is the proportion of subjects in the pivotal cohort FAS, in which a Tusker Medical tube was successfully placed, with presence of a Tusker Medical tube across the tympanic membrane in all successfully treated ears at the 3-week post-procedure follow-up visit. Subjects missing the 3-week follow-up visit will not be included in the Tube Retention secondary analysis.

[redacted]

Tube Patency secondary endpoint: Tube Patency is the proportion of subjects in the pivotal cohort, in which a Tusker Medical tube was successfully placed, with a functionally patent tube in all successfully treated ears at the 3-week post-procedure follow-up visit. Subjects missing the 3-week follow-up visit will not be included in the Tube Patency secondary analysis. [redacted]

[redacted]

Anesthesia Effectiveness secondary endpoint: Anesthesia Effectiveness is the proportion of subjects in the pivotal cohort FAS, who completed iontophoresis for all indicated ears, with adequate anesthesia for tympanostomy tube placement in all treated ears as determined by physician's evaluation of TM anesthesia prior to tube placement. Subjects who complete iontophoresis but do not complete TM tap assessment for all indicated ears will be counted as failures. [redacted]

[redacted]

21.0 Data Listings

Subject data will be summarized using listings and tables. All electronic case report form (eCRF) data will be listed per subject for all enrolled subjects. All listings will include the subject number and the identification of the study cohort.

22.0 Definition of Variables

22.1 Baseline

Baseline is defined as the last observation recorded prior to the treatment procedure.

Age will be calculated in years and fractions of year based on the date of informed consent relative to date of birth.

22.2 Study Day Calculation

Study Day 0 is the date of study procedure. Study Day is calculated relative to Study Day 0 and will appear in the listings where applicable.

Study Day will be calculated as:

Study Day = Date of event– Date of study procedure

No rules will be applied (eg, to assign missing day as 15th of month). Missing date data (day, month or year) will be queried.

23.0 Other General Considerations

Raw data will be presented with the exact precision that it was collected on the eCRF or other external data sources.

The number of decimal places to display for calculated data will be determined by the scale of measurement. No decimal places will be displayed if the smallest calculated value is 100; 1 decimal place will be displayed when all calculated values are within the interval (10, 100), with 10 being inclusive; 2 decimal places are displayed when all calculated values are within (1, 10), with 1 being inclusive; 3 decimal places are displayed for calculated values within (0.1, 1), with 0.1 being inclusive; and so on for even smaller scales of measurement.

Percentages will be reported with one decimal place. For summary statistics, means, medians, and confidence intervals will be displayed to one more decimal place than was determined above, dispersion statistics will have 2 more decimal places, and the minimum and maximum will be displayed to the same number of decimal places as determined by the rules above. Blank fields on an eCRF will be displayed as blank fields in corresponding listings.

24.0 References

1. ICH harmonised tripartite guideline - Statistical principles for clinical trials (E9) – Step 4, 05 Feb 1998.