Statistical Analysis Plan

For the ReActiv8-B Clinical Trial

A Clinical Trial under an Investigational Device Exemption (IDE)

ReActiv8 Implantable Neurostimulation System for Chronic Low Back Pain

Sponsor
MML US Inc.
15070 23rd Avenue North
Plymouth, MN 55447

Document Number 950061, Rev. A
10 December 2015

NCT02577354

Revision History

Rev	DCR#	Change Description	Effective Date
01	DCR-00760	Initial Release	14-Jan-2015
02	DCR-00893	1. Change primary endpoint to NRS a. Changed ODI to be a secondary endpoint b. Changed statistical hypotheses accordingly 2. Changes to ensure consistency between CIP and SAP 3. Minor spelling, grammatical and typographical errors fixed.	24-Apr-2015
A	DCR-01273	1. Change primary endpoint to VAS 2. Change the procedure to address multiple testing for the secondary objectives 3. Supply details of the planned interim analysis 4. Clarify planned secondary and supplemental analyses 5. Minor spelling, grammatical and typographical errors fixed	10-Dec-2015

Contents

1. Overview

This Statistical Analysis Plan (SAP) describes the statistical methods to be used during the reporting and analysis of data collected under the ReActiv8-B Clinical Investigation Plan (CIP).

2. Scope

This SAP should be read in conjunction with the CIP and case report forms (CRFs). This version of the SAP has been developed with respect to the CIP version 950057, Rev A and updated per 950057, Rev B. Any further changes to the CIP or CRFs may necessitate updates to the SAP.

3. Primary Efficacy Objective

The Primary Efficacy Endpoint is a comparison of responder rates between Treatment and Control groups, where a "responder" is a Subject with ≥30% reduction from baseline in average low back pain VAS without any increase from baseline in pain medication and/or muscle relaxants prescribed and taken in the two weeks prior to the Primary Endpoint Assessment Visit.

Note: The primary outcome measure is intended to assess the efficacy of 120 days of stimulation. Thus the visit for assessment of the primary outcome must occur no sooner than 120 days post randomization.

4. Secondary Efficacy Objectives

There are six specified secondary efficacy objectives. The first five objectives have associated endpoints and hypotheses and will be statistically tested at the Primary Endpoint Assessment Visit. The sixth objective, consisting of multiple endpoints, will be evaluated after the Outcome Post Crossover Visit but does not have associated statistical hypotheses.

1. Comparison of change from baseline in Oswestry Disability Index (ODI) between Treatment and Control groups at the Primary Endpoint Assessment Visit
2. Comparison of change from baseline in EQ-5D between Treatment and Control groups at the Primary Endpoint Assessment Visit
3. Comparison of Percent Pain Relief between Treatment and Control groups reported by the Subject at the Primary Endpoint Assessment Visit
4. Comparison of Subject Global Impression of Change at the Primary Endpoint Assessment Visit
5. Comparison of number of Subjects with Resolution of Low Back Pain (remitters or cure, defined as a VAS score ≤ 2.5cm) at the Primary Endpoint Assessment Visit
6. Evaluation of changes in primary and secondary efficacy metrics in the Crossover group following the Outcome Post Crossover Visit.

5. Trial Design

The study is an international, multi-center, prospective, randomized, sham controlled, blinded trial with an adaptive statistical design. Subjects and investigators are blinded to treatment assignment, and the assessment of primary endpoint data is blinded. Subjects meeting all eligibility criteria who have been implanted with ReActiv8 will be randomized to either the Treatment Arm or the Control Arm. Randomization will be performed according to a random permuted block design stratified by clinical site with a 1:1 allocation ratio for Treatment vs. Control.

All Subjects are implanted with the ReActiv8, and then at the Randomization Visit are instructed to deliver stimulation for 60 minutes per day, in two 30-minute sessions (morning and evening). Subjects are randomized (1:1) to the Treatment or Control Arm at the Randomization Visit.

Subjects randomized to the Treatment Arm of the study will have ReActiv8 programmed to deliver stimulation at a Subject-appropriate level (The Treatment group). Determination of the Subject appropriate level is described in the Implant and Programming Manual. Subjects in the Treatment group may or may not feel muscle contraction during stimulation.

Subjects randomized to the Control Arm will have ReActiv8 programmed to deliver minimal stimulation (The Control group). Subjects in the Control group may or may not feel muscle contraction during stimulation.

All Subjects will have automatic electrical impedance measurement of electrodes performed immediately prior to each stimulation session.

IRB (or Ethics Committee) and FDA approval will be sought to enroll up to 800 Subjects, in up to 40 implanting sites to achieve an initial target of 116 randomized Subjects in the Intent to Treat (ITT) Cohort. An adaptive statistical design using one interim look when at least 50% of Subjects have completed the Primary Endpoint Assessment Visit, will be used for a potential sample size re-estimation.

A Subject is “enrolled” in the study at the time of informed consent. Subjects who are enrolled, meet the required inclusion and exclusion criteria, and continue consent will be scheduled for implant of the ReActiv8.

5.1. Randomization

Subjects will be randomized at the Randomization Visit 14 ± 3 days post implant. Subsequent visits are scheduled relative to the date of randomization with the primary efficacy endpoint assessed at the Primary Endpoint Assessment Visit. Randomization will be performed according to a random permuted block design stratified by clinical site with a 1:1 allocation ratio for Treatment vs. Control.

5.2. Blinding (Masking)

Subjects and investigators are blinded to treatment assignment, and endpoint assessments are also blinded. An assessment of the effectiveness of the blinding procedures will be performed at the Primary Endpoint Assessment Visit.

6. Sample Size Considerations

6.1. Sample Size Estimate

The sample size for the study is determined under the following assumptions for the primary efficacy endpoint:

- Minimum power of 80%
- Type I error of 5%
- Assumed primary efficacy success in Treatment group: 50%
- Assumed primary efficacy success in Control group: 25%

Under the above assumptions, a minimum of 116 evaluable Subjects is required in order to demonstrate superiority of the ReActiv8 treatment. The total sample size is increased by 10% to 128 to allow for attrition. The total sample size may be increased, but not decreased, at the time of the planned interim analysis to adjust for updates to the planning assumptions.

6.2. Level of Power and Significance

All tests of significance will be performed at the two-sided 0.05 significance level. The hypotheses associated with the first five secondary efficacy objectives will be tested using the Hochberg multiple comparison adjustment in the event the primary efficacy objective is met.

7. Data Structure and Handling

7.1. Data Handling and Transfer

Programming of analysis datasets, tables, figures and listings will be conducted during the data management phase of the study. Tables, figures, and listings may be reviewed prior to final data lock for data review. Any data values requiring investigation or correction will be identified, and protocol deviations will be reviewed. The final run of outputs will take place after the data are deemed final.

7.2. Missing Data and Sensitivity Analyses

The primary efficacy endpoint will be assessed in the ITT Cohort, consisting of all randomized Subjects not in the Surgical Roll In Cohort, with missing endpoint assessments examined in sensitivity analyses. Secondary efficacy endpoints will be assessed in the population of Subjects with complete data (baseline and at the Primary Endpoint Assessment Visit). Sensitivity analyses of the impact of missing data will include a tipping point analysis for the primary efficacy endpoint where the endpoint is computed with possible combinations of imputed success and failure for Subjects with missing primary efficacy assessment. Additionally, for the primary and secondary efficacy endpoints, a last observation carried forward analysis will be conducted.

7.3. Visit Windows

For analysis purposes, data will be analyzed according to the visit assigned within the database.

7.4. Pooling of Data Across Trial Sites

Poolability of data will be assessed using a Mantel-Haenszel (MH) test procedure to examine the null hypothesis of a common odds ratio of responders across sites. The data will be considered poolable if the p-value for the test is ≥ 0.15. If the p-value is < 0.15 additional analyses will be conducted to determine whether the site differences are due to imbalances in baseline factors that have evidence ($p < 0.20$) of an association with the primary outcome. These analyses may include an analysis of covariance on the percentage of pain relief, used to define responders, with site included as a random effect and adjustments for baseline covariates found to be associated with pain relief. If poolability across sites is not demonstrated, then results for the primary endpoint may be stratified and presented by site groupings.

8. Statistical Analyses

8.1. General Considerations

All statistical analyses will be conducted using SAS version 9.3 or later (SAS Institute Inc., Cary, NC) or R version 2.14 or later (R Development Core Team, 2012), and other validated statistical software as required. All output will be in Microsoft Word or RTF format, converted to PDF. Continuous variables will be summarized with means and standard deviations or as medians and interquartile ranges. Categorical variables will be summarized with the number and proportion of Subjects in each category. Binary outcomes will be presented as proportions with corresponding 95% confidence limits.

The primary efficacy endpoint and associated secondary efficacy endpoints are assessed at the Primary Endpoint Assessment Visit.

8.2. Study Populations

Several Cohorts are defined.

Intent to Treat (ITT) Cohort consists of all randomized Subjects who are not in the Surgical Roll-In group. This is the primary analysis population for the primary efficacy endpoint. The primary efficacy endpoint will be assessed in the ITT cohort. Additionally, the following analysis populations are defined. The primary efficacy endpoint will also be assessed in the per-protocol and completer cohorts. The secondary efficacy endpoints will also be assessed in the ITT cohort.

Surgical Roll In Cohort consists of all Subjects implanted in the Surgical Roll In group.

Per-protocol: Subjects within the ITT cohort receiving the correct randomized treatment, meeting the enrollment criteria, compliant with delivery of stimulation at least 80% of the maximum possible sessions delivered, and who have no other major protocol deviations affecting the primary endpoint evaluation will be considered to be part of the Per Protocol population.

Completers: All randomized Subjects who complete the Primary Endpoint Assessment Visit and associated endpoint assessments, regardless of compliance with the delivery of stimulation or the requirement to maintain constant medications.

Crossover Cohort: All Subjects in the Crossover Group who met the inclusion criterion 2 for low back pain (Low Back Pain VAS of ≥6.0cm and ≤9.0cm (on a 10cm scale)) and inclusion criterion 3 for ODI (Oswestry Disability Index score ≥21% and ≤60%) at the time of crossover.

The primary efficacy endpoint and secondary efficacy endpoints will also be assessed in the Per-protocol and Completer cohorts.

8.3. Subject Disposition

Subject disposition for randomized Subjects will be presented by:

- Summary of Subjects per visit
- Summary of early withdrawal and reason for early withdrawal.

In addition, Subjects enrolled and not implanted, or withdrawing prior to randomization, will be summarized.

8.4. Demographics and Baseline Characteristics

Demographics and baseline characteristics of the data set will be summarized by randomization group. These parameters will include (but not be limited to):

- Age
- Gender
- Ethnicity
- BMI
- Medical history
- Low Back pain descriptive characteristics
- Duration of back pain
- Work status (including duration of “off work” and “part time” status for those Subjects adversely affected by back pain)
- Low Back pain medications and concomitant medications

8.5. Primary Efficacy Endpoint

8.5.1. Primary Analysis

The Primary Efficacy Endpoint is a comparison of responder rates between Treatment and Control groups, where a “responder” is a Subject with ≥30% reduction from baseline in average low back pain VAS, without any increase from baseline in pain medication and/or muscle relaxants prescribed and taken in the two weeks prior to the Primary Endpoint Assessment Visit.

Note: The primary outcome measure is intended to assess the efficacy of 120 days of stimulation. Thus the visit for assessment of the primary outcome must occur no sooner than 120 days post randomization.

The primary efficacy endpoint will be assessed by the following hypotheses:

H0: $\Pi VAS_T = \Pi VAS_C$

HA: $\Pi VAS_T \neq \Pi VAS_C$,

where ΠVAS_T is the proportion of Subjects meeting primary success criteria in the Treatment group and ΠVAS_C is the proportion of Subjects meeting success criteria in the Control group. The hypotheses will be tested using a two-sided binomial test for a difference in proportions. The objective will be met if a statistically significant difference favoring the Treatment group is found.

The study will be considered a success if the primary efficacy endpoint is met.

For the primary analysis in the ITT population, the impact of Subjects with missing endpoint assessment will be examined in sensitivity analyses.

Records of pain medications will be collected along with all other medications used for treatment of low back pain, which are also being collected for analysis of secondary and cost-effectiveness endpoints. The Subject will report medications taken at each scheduled follow-up visit. Rescue medications taken on an exceptional basis for acute pain conditions other than back pain will also be documented and their possible effect examined as part of sensitivity analyses.

The number and percent of randomized Subjects meeting the responder criteria will be presented by randomization group with associated 95% asymptotic confidence limits. The difference in proportions and associated 95% confidence limits will be presented along with the p-value from a binomial test for the test of the hypothesis associated with the primary objective.

The individual components of the primary efficacy endpoint (VAS and medications) will be analyzed and presented separately. Cumulative proportion of responder analysis will also be used as a way of presenting this data.

8.5.2. Supporting Analyses

The primary efficacy endpoint will also be analyzed in the Per-protocol and Completer cohorts.

8.6. Secondary Efficacy Endpoints (SEE)

The hypotheses will be tested utilizing an overall Type I error of 5%. The overall Type I error will be maintained by utilizing Hochberg's method to test the collection of the first five secondary objectives. These will be tested only if the primary efficacy endpoint is met. Secondary efficacy endpoints will be analyzed in the Per Protocol population.

8.6.1. SEE 1: ODI

The ODI efficacy objective will be tested under the following hypotheses:

H0: $\mu ODI_T = \mu ODI_C$

HA: $\mu ODI_T \neq \mu ODI_C$

Where μODI_T is the mean change in ODI in the Treatment group and μODI_C is the mean change in ODI in the Control group at the Primary Endpoint Assessment Visit. The hypotheses will be tested using a two-sided two-sample t-test for a difference in mean changes.

8.6.2. SEE 2: EQ-5D

The EQ-5D efficacy objective will be assessed under the following hypotheses:

H0: $\mu EQ5D_T = \mu EQ5D_C$

HA: $\mu EQ5D_T \neq \mu EQ5D_C$,

where $\mu EQ5D_T$ is the mean change in EQ-5D (index score) in the Treatment group and $\mu EQ5D_C$ is the mean change in EQ-5D in the Control group from baseline to the Primary Endpoint Assessment Visit. The hypotheses will be tested using a two-sided two-sample t-test for a difference in mean changes.

8.6.3. SEE 3: Percent Pain Relief

The secondary Percent Pain Relief efficacy objective will be tested under the following hypotheses:

H_0: $\mu PPR_T = \mu PPR_C$

H_A: $\mu PPR_T \neq \mu PPR_C$

Where μPPR_T is the mean percent pain relief in the Treatment group and μPPR_C is the mean percent pain relief in the Control group at the Primary Endpoint Assessment Visit. The hypotheses will be tested using a two-sided two-sample t-test for a difference in mean changes.

8.6.4. SEE 4: Subject Global Impression of Change

The secondary Subject Global Impression of Change efficacy objective will be tested under the following hypotheses:

H_0: $_{DISTRIBUTION}SGIC_T = {}_{DISTRIBUTION}SGIC_C$

H_A: $_{DISTRIBUTION}SGIC_T \neq {}_{DISTRIBUTION}SGIC_C$

Where $_{DISTRIBUTION}SGIC_T$ is the distribution of the Global Impression of Change in the Treatment group and $_{DISTRIBUTION}SGIC_C$ is the distribution of the Global Impression of Change in the Control group at the Primary Endpoint Assessment Visit. The hypotheses will be tested using a two-sided Mann-Whitney test.

8.6.5. SEE 5: Resolution of Low Back Pain

The secondary Resolution efficacy objective will be tested under the following hypotheses:

H0: $\Pi RLBP_T = \Pi RLBP_C$

HA: $\Pi RLBP_T \neq \Pi RLBP_C$

Where $\Pi RLBP_T$ is the proportion of Subjects with a VAS back pain assessment for the previous 7 days of ≤2.5 in the Treatment group and $\Pi RLBP_C$ is the proportion of Subjects with a VAS back pain assessment for the previous 7 days of ≤2.5 in the Control group at the Primary Endpoint Assessment Visit. The hypotheses will be tested using a two-sided binomial test for a difference in proportions.

8.6.6. SEE 6: Crossover Group Analysis

Changes in all primary and secondary efficacy metrics will be evaluated in the Crossover group following the Outcome Post Crossover Visit. Outcome measures will be assessed for all Crossover Subjects, and a separate analysis will be performed for all Subjects who met the inclusion criteria for low back pain at the Primary Endpoint Assessment Visit (i.e.: prior to Crossover) (prior week average Low Back Pain VAS of ≥6.0cm and ≤9.0cm (on a 10cm scale)) and inclusion criterion for ODI (Oswestry Disability Index score ≥21% and ≤60%). The analysis will also explore correlation of outcomes to amount of stimulation delivered.

8.7. Interim Analysis

A single interim analyses of the primary efficacy endpoint will occur when at least 50% of the original planned study enrollment has completed the Primary Endpoint Assessment Visit. This analysis will be limited to results from this cohort and will be used for purposes of a potential adaptive sample size increase. The interim analysis plan has the following characteristics:

- Only a single interim analysis is planned.
- The interim analysis can lead only to an increase in the study size, not a decrease in sample size or early study termination for benefit.
- If the conditional power at the time of the interim analysis is less than 30% or greater than 90%, then there will be no change in the study size. If between 30% and 90%, the study size may be increased up to a maximum of twice its originally designed size (that is, 2 x 116) to regain the original 80% power objective.
- The interim analysis will be performed by a third-party independent statistician under the direction of the Data Monitoring Committee (DMC), and detailed interim results, other than a DMC recommendation regarding the findings, will remain blinded to the sponsor and other study participants to prevent introduction of operational bias.

- On the basis of the interim analysis, the DMC may recommend one of three actions: continue the study as is, increase the sample size, or as is always an option, potentially stopping the study for reasons of safety or futility.

At the time of an interim analysis, based on the estimated conditional power, a sample size re-estimation may be performed using the method of Cui et al.[a] The results will not be used to decrease the original sample size estimate or provide a basis for early study stopping for positive efficacy results. However, if it is determined that an excessive number of additional Subjects are required, a decision to recommend stopping the study early for futility may be made.

Since there is no potential for early termination benefit, no alpha adjustment to the final p-value is required other than that associated with the use of the Cui et al procedure.

8.8. Safety Analysis

8.8.1. Primary Safety Assessment

The primary safety assessment is of serious device and/or procedure related adverse events in all Subjects in the Intent to Treat Cohort at the Primary Endpoint Assessment Visit. All reported adverse events will be documented and reported with summary statistics presented for observed rates. There are no formal, statistical hypotheses being tested in the safety assessment.

There will be a separate safety analysis of Subjects in the Surgical Roll In Cohort. If numbers permit, a Poolability analysis of the ITT and the Surgical Roll In Cohort will be conducted.

8.8.2. Supporting Safety Analysis

The PMA Submission will include 12 month safety data on all Subjects implanted with ReActiv8 including all Subjects implanted in the ReActiv8-B Trial for whom 12 month safety data are available, and other persons implanted with ReActiv8 in other clinical trials (including the ReActiv8-A PMCF study) and all post-marketing registries that are registered on www.clinicaltrials.gov. This data will be provided as a separate analysis of the safety on all implanted persons as an appendix to the ReActiv8-B trial results.

A separate safety assessment will be performed on the Surgical Roll In Cohort, including a poolability analysis.

8.9. Supporting Analyses

The supporting efficacy analyses will include:

- Once data are complete, all primary and secondary outcome measures will be assessed at 12 months and compared to baseline within the Treatment group. The analysis will also explore correlation of outcomes to amount of stimulation delivered.

a Cui, L., Hung, H. M., & Wang, S. J. (1999). Modification of sample size in group sequential clinical trials. Biometrics, 55, 853–857. doi:10.1111/j.0006-341X.1999.00853.x

Note: For all supporting analyses, amount of stimulation delivered will be defined as the number of minutes stimulation was delivered divided by the number of minutes that stimulation could have been delivered.

- In the Crossover Group, outcome measures will be assessed at the Outcome Post Crossover Visit for all Subjects who met the inclusion criteria for low back pain at the Primary Endpoint Assessment Visit (i.e.: prior to Crossover) (Prior week average Low Back Pain, VAS of ≥6.0cm and ≤9.0cm (on a 10cm scale)) and inclusion criterion for ODI (Oswestry Disability Index score ≥21% and ≤60%). The analysis will also explore correlation of outcomes to amount of stimulation delivered.
- Comparison of daily average VAS averaged over the prior 7 days of Journal entry with the single point VAS of recollection of prior 7 days of average pain.
- Comparison of EQ-5D VAS between Treatment and Control groups at each visit at which EQ-5D is recorded through 12 months.
- For the both the Treatment and Control groups, a cumulative proportion of responder analysis, so that the entire distribution of treatment response is depicted in a graph of the proportion of responders for all percentages of low back pain VAS reduction from 0% to 100%. In this manner, the number of Subjects with ≥50% low back pain VAS improvement can be determined, for example. The analysis will also be performed for the responder analysis of ODI.
- Subject Global Impression of Change at 12 months: The analysis will explore correlation of outcomes to amount of stimulation delivered.
- Treatment satisfaction questionnaire at 12 months: The analysis will explore correlation of outcomes to amount of stimulation delivered.
- Clinical Global Impression – Global Improvement (clinician assessment) at 12 months: The analysis will also explore correlation of outcomes to amount of stimulation delivered.
- Change in opioids used for treatment of low back pain from baseline to 12 months. The analysis will also explore correlation of outcomes to amount of stimulation delivered.

Additional, ad hoc exploratory analyses may also be conducted.

Note that the supporting analyses may be incomplete at the time of PMA submission since not all Subjects may have reached 12 months post-randomization.

8.10. Health Economics

Cost-effectiveness of ReActiv8 therapy, extrapolated to the lifetime of the device, will be measured. Key economic outcome measures to be recorded will be:

- Change in work status at all visits.
- Change in consumption of medications taken for low back pain at all visits compared to Baseline. Note that an analysis may be performed of different categories of medications including for example opioids, muscle relaxants and anti-depressants.
- Assessment of health care utilization (office visits, hospital visits, emergency room visits, and other therapies such as physical therapy) at all visits.

Note that both the Treatment and Control group Subjects are required to keep medications constant from randomization until the Primary Endpoint Assessment Visit, and other health care utilization will be as per the protocol (e.g.: no additional therapies).

Following the Primary Endpoint Assessment Visit, all Subjects will be permitted to adjust medications and to use other health care resources. For the Crossover Group, this may yield some information on how Subjects will adjust medications during the treatment phase.

Note that completion of the Health Economics assessments is likely to be incomplete at the time of PMA submission for ReActiv8, since not all Subjects will have reached the 12 month point at that time.

8.11. Subgroup Analyses and Predictors of Response

The primary efficacy endpoint will be analyzed in the subgroups defined below to demonstrate consistency of treatment effect. A logistic regression model will be used to test the interaction of randomization assignment and subgroup/covariate. Subgroups to be analyzed will include, at a minimum:

- Gender
- Age (as both a continuous variable and grouped in ranges)
- Ethnicity
- Baseline VAS
- BMI (as both a continuous variable and grouped in ranges)
- Duration of work status adversely affected by CLBP

Multivariable analysis to explore predictors of response will be conducted in the Treatment group Subjects. Univariable regression models with covariates as defined below will be analyzed to determine predictors of response. All covariates with an association with response ($p<0.20$) will be included in a multivariable logistic regression model.

- Duration of back pain (as both a continuous variable and grouped in ranges)
- Baseline ODI (as both a continuous variable and grouped in ranges)
- Baseline EQ5D (as both a continuous variable and grouped in ranges)
- Baseline VAS (as both a continuous variable and grouped in ranges)
- Baseline opioid use
- Age
- Gender
- BMI
- Compliance with stimulation (the number of minutes that stimulation was actually delivered divided by the number of minutes that could have been delivered)