STATISTICAL ANALYSIS PLAN

For

Clinical Evaluation of the Cochlear Nucleus® CI532 Cochlear Implant in Adults

Investigation Number: CLTD5685

August 4, 2019
Version 5.0

Study Sponsor:

Cochlear Limited
1 University Avenue
Macquarie University, NSW 2109
Australia

Table of Contents

1.0 Purpose

This Supplemental Analysis Plan (SAP) describes the statistical methods to be used during the reporting and analysis of data collected for the Clinical Evaluation of the Cochlear Nucleus® CI532 Cochlear Implant in Adults. This supplemental plan should be read in conjunction with the Study Protocol.

2.0 Clinical Study Design Overview

This clinical investigation will be conducted as a multicenter, prospective, nonrandomized, single-subject, repeated-measures design in which each subject serves as his/her own control. This approach accommodates the heterogeneity that characterizes hearing-impaired populations. Blinding procedures are not appropriate for this trial design, as it is not possible to conceal the presence, or absence, of a cochlear implant (CI) system from recipients and/or clinical Investigators.

Subjects in this study will be implanted with the commercially available Nucleus CI532 cochlear implant and fit with the CP1000/Nucleus 7 Sound Processor. All testing will be complete with CP1000/Nucleus 7 Sound Processor.

3.0 Subject Population

Study sites will enroll 100 adult cochlear implant candidates at up to 15 North American cochlear implant centers.

Cochlear Americas will adhere to the National Coverage Determination (NCD) for Cochlear Implantation for enrollment of Medicare beneficiaries into the Clinical Evaluation of the Cochlear Nucleus® CI532 Cochlear Implant in Adults. Specifically, the following NCD criteria will be met:

1. Diagnosis of bilateral moderate-to-profound sensorineural hearing impairment with limited benefit from appropriate hearing (or vibrotactile) aids, defined as test scores of less than or equal to 40% correct in the best-aided listening condition on tape-recorded tests of open-set sentence recognition;
2. Cognitive ability to use auditory clues and a willingness to undergo an extended program of rehabilitation;
3. Freedom from middle ear infection, an accessible cochlear lumen that is structurally suited to implantation, and freedom from lesions in the auditory nerve and acoustic areas of the central nervous system;
4. No contraindications to surgery.

Candidates must also meet the following inclusion and exclusion requirements as stated in the study protocol:

Inclusion Criteria:

1. Adults 18 years or older who have a bilateral postlinguistic sensorineural hearing loss.
2. Limited benefit from amplification as defined by test scores of 40% correct or less in the ear to be implanted and 50% or less in the contralateral ear on a recorded monosyllabic word test
    I. Consistent with the Minimum Speech Test Battery (2011), it is required that all subjects be evaluated at 60 dBA presentation level.
3. Bilateral moderate sloping to profound hearing loss

4. Minimum of 30 days experience with appropriately fit bilateral amplification, fit using the standardized NAL fitting method
5. Proficient in English
6. Ability to complete testing

Exclusion Criteria:

1. Previous cochlear implantation
2. Pre-linguistically deafened (onset of hearing loss at less than two years of age)
3. Ossification or any other cochlear anomaly that might prevent complete insertion of the electrode array
4. Duration of severe to profound hearing loss greater than 20 years
5. Diagnosis of retro-cochlear pathology
6. Diagnosis of auditory neuropathy
7. Unrealistic expectations on the part of the subject regarding the possible benefits, risks, and limitations that are inherent to the surgical procedure and use of the prosthetic device
8. Unwillingness or inability to comply with all investigational requirements
9. Additional cognitive, medical or social handicaps that would prevent completion of all study requirements

4.0 Data Analyses

4.1 Sample Size Estimate

To calculate the minimum sample size required to reject the null hypothesis (H0) that "Group mean CNC words scores in the best unilateral listening condition measured at 6 months post sound processor activation will not be superior to the group mean score in the preoperative, best unilateral condition", the following values were chosen:

- A minimum clinically meaningful difference between pre to post activation at 6 months of 15% for CNC words in quiet, based on clinical consensus.
- An expected standard deviation of difference scores of 22.4% for CNC words in quiet. This SD is based on the estimated standard deviation of difference score data collected in completed clinical studies
- A significance level $\alpha = 0.05$ (two-tailed).
- A desired power of 0.9 i.e. there is 90% chance of detecting a real change between the experimental programs.

Based on these assumptions for words in quiet, a sample size of 26 subjects is required to achieve a power of 90%.

AzBio Sentences in Noise

- To calculate the minimum sample size required to reject the null hypothesis (H0) that "Group mean AzBio sentence in noise scores (SNR +10) in the best unilateral listening condition measured at 6 months post sound processor activation will not be superior to the group mean score in the preoperative, best unilateral condition", the following values were chosen:

- A minimum clinically meaningful difference between pre to post activation at 6 months of 15% for sentences in noise, based on previous clinical consensus.
- An expected standard deviation of difference scores of 21.3% for AzBio sentences in noise. This SD is based on previous clinical trial data with newly implanted subjects.
- A significance level of $\alpha = 0.05$ (two-tailed).
- A desired power of 0.9
- Based on these assumptions, a sample size of 24 subjects is required to achieve a power of 90%.

Health Utility Index (HUI)

To calculate the minimum sample size required to reject the null hypothesis (H0) that "Group mean utility score (HUI) at 6 months post sound processor activation will not be superior to the scores measured preoperatively the following values were chosen:

- A minimum clinically meaningful difference between pre to post activation at 6 months of 0.12 utility score
- An expected standard deviation of difference scores of 0.18, this is a conservative value based on the SD of the difference measured after hearing aid fitting
- A significance level of $\alpha = 0.05$ (two-tailed).
- A desired power of 0.9

Based on these assumptions, a sample size of 26 subjects is required to achieve a power of 90%.

A sample size of 100 adult subjects will be enrolled in the clinical investigation, this large sample will allow for additional analyses to be conducted specifically the additional patient reported outcome measures, where a large sample size is typically desired. Additionally the sample size provides greater justification for generalizability to the wider clinical population.

4.2 Study Population

All subjects who are consented and enrolled into the clinical study will constitute the intention-to-treat (ITT) population for the purposes of adverse event reporting. Only subjects implanted with the CI532 and fit with the CP1000/N7 Sound Processor and completed per the protocol will be considered as the completed cases (CC) population and per protocol (PP).

4.3 Determining Efficacy

Efficacy of the Cochlear Nucleus CI532 cochlear implant system will be determined by a comparison of preoperative vs. postoperative outcomes measures. The speech measures for this purpose are the CNC Word Test and the AzBio sentences in noise. The primary and secondary efficacy endpoints will be based on analyses of designated test measures at the 6-month postactivation.

4.3.1 *Primary Efficacy Objective*

The primary efficacy objective for this study is to understand if cochlear implantation with a Cochlear Nucleus CI532 cochlear implant and the use of the CP1000/N7 sound processor in adult patients results in improved speech understanding at 6 months postactivation, as measured by performance on an open-set word recognition test, in the best unilateral listening condition of the implanted ear.

- On the CNC word measure, the group mean score for best unilateral listening condition at 6 months will be better than the group mean score in the preoperative, unilateral aided condition.

The primary hypothesis to be tested in this analysis is the following: that the 6 month postoperative performance is significantly different from preoperative performance in the treated ear. The null and alternative hypotheses are given below.

H01: ìPost – ìPre ≤ 0 versus Ha1: ìPost – ìPre > 0

Where ìPre is the CNC word score in quiet obtained with a hearing aid preoperatively in the ear to be implanted, and ìPost is the CNC word score in quiet obtained at 6 months post sound processor activation in the treated ear. The test will be based on a one-sample t-test of the paired difference in pre and post results and performed at the one-sided 0.025 alpha level. If there is significant evidence that the assumptions of the t-test do not hold (i.e. $p<0.05$ from a Shapiro-Wilks test of normality), a Wilcoxon signed rank test will be used.

4.3.2 *Secondary Efficacy Objective*

The secondary efficacy objective will be determined by further comparison of group means for preoperative vs. 6 months postoperative outcome measures in best unilateral as measured by performance on sentence recognition in noise. The speech measure for this purpose is the AzBio sentences in noise at a +10 dB Signal-to-noise-ratio

- On the AzBio sentences-in-noise measure at a +10 SNR, the group mean score for best unilateral listening condition at 6 months will be better than the group mean score in the preoperative, unilateral aided condition.

The hypothesis test is as follows:

H01: ìPost – ìPre ≤ 0 versus Ha1: ìPost – ìPre > 0

Where ìPost is the post-implant value and ìPre is the pre-implant value. Each test will be based on a one a sample paired t-test of the difference in pre and post results. All tests will be performed at the one-sided 0.025 alpha level.

An additional secondary objective will be evaluated by comparing of the group mean for preoperative vs. 6 months postoperative health utility outcome as measured on the HUI3.

The hypothesis test is as follows:

H01: ìPost – ìPre ≤ 0 versus Ha1: ìPost – ìPre > 0

Where ìPost is the post-implant value and ìPre is the pre-implant value. Each test will be based on a one sample paired t-test of the difference in pre and post results. All tests will be performed at the one-sided 0.025 alpha level.

Additional supportive efficacy analyses will include:

- An analysis of individual data for speech perception measures (CNC word scores and AzBio Sentences in noise scores) to establish the proportions of those subjects showing a significant

improvement, no change, and a significant decrement in performance. Individual scores obtained at 6 months will be compared with those obtained, on the same measures preoperatively, based on the binomial model (see Thornton and Raffin, 1978). A similar analysis will be completed for data obtained at 3 and 12 months.

Additionally, for the HUI3, the proportion of individual subjects showing significant improvement (i.e. a change pre to 6 months on the multi score of ≥0.03) will be reported. For the individual attributes a change pre to 6 months of ≥0.05 is considered significant, (Horsman, et al.,2003) The proportion of subjects demonstrating a significant change will be reported.

Both group mean scores and proportion of individual subjects showing significant benefit, no change in performance and significant decrement in performance will be documented and reported in a detailed annual progress report and in subsequent publications.

For all speech perception measures individual data analysis will be performed and reported in a peer reviewed journal and in the Annual Progress Report or presentation.

4.4 Type 1 Error Control

The hypotheses for the primary and secondary Endpoints are to be tested formally according to a fixed sequence testing procedure. This testing procedure is based on the principle of a closed testing procedure and is used to control the type I error. In the fixed sequence testing, all the following superiority tests will be conducted using a two-tailed 95% CI (alpha=0.025 one-sided).

The Primary Endpoint will be tested first. Only if the superiority test for the Primary Endpoint is successful, then the secondary endpoints will be tested. If the superiority test for the Primary Endpoint fails, the testing procedure stops and no further testing will be performed.

4.5 Additional Statistical Analyses

4.5.1 Analysis of Demographic Characteristics

The demographic characteristics of the study group will be presented descriptively. Quantitative variables such as age, duration hearing loss and duration of severe to profound hearing loss will be presented with mean, standard deviation, median, minimum and maximum. Other demographic factors such as etiology and medical comorbidities will be tabulated for each subject.

4.5.2 Subgroup Analyses

The consistency of the primary endpoint will be examined across subgroups of subjects defined by the following baseline characteristics: age, gender, and duration of hearing loss, baseline CNC word scores, and existing medical comorbidities. Any significant difference between subgroups on endpoints will be explored with additional analyses. Additionally, for the analysis of age subjects will be stratified into 2 groups according to age at time of surgery: 1). <65 years and 2). ≥65 years. The effect of age may further be investigated by examining outcomes by decade.

Additional analyses will be completed to evaluate speech perception in quiet and noise at 6 months in the best bilateral listening condition – for most subjects this will be the bimodal condition (CI + contralateral hearing aid), and for others this will be the combined condition CI + Acoustic hearing (same ear) + contralateral hearing aid). Additionally, a subgroup analysis will be done for: Subjects using bimodal hearing and subjects using combined hearing. Differences between and within the subgroup will be evaluated.

Subgroup analysis will be done for: Subjects using hybrid hearing (acoustic + electric in the implanted and subjects using electric alone. Differences between and within the subgroup will be evaluated for primary and secondary endpoints.

The number of subjects falling into each group following surgery cannot be predicted consequently, sample size is not adjusted based on these subgroup analyses.

In addition, the consistency of the primary endpoint will be examined across investigational sites by testing for an effect of site in ANOVA model. Any potential variation between sites in the primary endpoint will be explored by assessing whether or not there are differences in baseline characteristics between subjects at sites that might explain the results.

Exploratory analyses may be conducted to determine factors that predict performance on a CNC word score at 6 months.

Factors that will be considered are:

- Demographic variables such as age, degree of hearing loss, duration of hearing loss
- Other preimplant variables such as: CNC word score, AzBio +10 score, AzBio +5 score, PTA, LFPTA, SSQ score, HUI score and MoCA score (pass/fail)
- Surgical factors, such as surgical approach, speed of insertion of sheath and electrode and number of white markers visible
- Electrode factors such as, electrode location, WF, basal angle of insertion and apical angle of insertion
- Post implant variables: hours of device use, CNC word score at 3 months, AzBio +10 at 3 months, HUI and SSQ scores at 6 months

4.5.3 Change in Audiometric Thresholds

This is not a hearing preservation study and as such loss of low frequency hearing will not be tracked as an adverse event, changes in low frequency will be documented and reported. Specifically:

- The change in the low frequency pure tone average audiometric threshold in dB (125 Hz-750 Hz LFPTA) at each interval (IA, 1, 3, 6 and 12 months) will be calculated for each subject.
- The change in the audiometric threshold in dB at 500 Hz at each interval (IA, 1, 3, 6 and 12 months) will be calculated
- Additionally the effect of age on the change in LFPTA at each interval will be calculated. The effect of age will further be analyzed by stratifying the subjects into 2 groups based on their age at time of enrolment: <65 years of age and ≥ 65 years of age. Differences between and within the groups will be evaluated.

Additionally, the audiometric results will be reported according to the AAO Consensus on minimal reporting (Adunka, 2018)

These results will be documented in an Annual Progress Report and in publication(s).

4.5.4 Imaging

Post-operative CT Scans will be analyzed through 3D reconstruction to report on:

- Placement of the electrode in the cochlea: Scala Tympani versus Scala Vestibuli
- Measurement of modiolar proximity also known as the wrapping factor, scores will range between 0 -1, with 0 being the closest to the modious and 1 being furthest on the lateral wall. Wrapping factor will be reported as a percentage.

Possible exploratory analysis may be completed to determine if there is a correlation between the electrode placement and wrapping factor any surgical technique factors as reported on the surgical questionnaire with speech perception (CNC word score and AzBio Sentences in noise).

Electrode position will be also correlated with T& C levels and impedance values after at least 60 days of device use.

4.5.5 Patient Reported Outcomes

Hearing Aid Experience Questionnaire and the Nucleus 7 Questionnaire

These two device use questionnaires were developed by the Sponsor and are used to collect information regarding device usability, subjective preferences, and satisfaction with regards to device use in various listening conditions. The baseline is the Hearing Aid Experience Questionnaire and the follow-up is the N7 questionnaire completed at the 6 month interval. The differences between the N7 questionnaire and baseline (i.e. change in baseline) will be analyzed using a repeated measures, mixed effects model, with the device use difference as the response and variables such as speech perception, age and gender as the explan*atory variables.* Additionally, the proportion of subjects reporting being satisfied a score of >50 will also be reported.

***Speech*, Spatial, and Qualities of Speech Questionnaire (SSQ)**

The SSQ is a close-ended questionnaire composed of several subjective questions related to quality of speech. The baseline and the follow-up SSQ at the 6 month interval will be recorded for each subject. The differences between the 6-month SSQ and baseline SSQ will be analyzed using a one-sample t-test of the paired difference in pre and post results and performed at the one-sided 0.025 alpha level. If there is significant evidence that the assumptions of the t-test do not hold (i.e. $p<0.05$ from a Shapiro-Wilks test of normality), a Wilcoxon signed rank test will be used. Additionally, the differences between the 6-month SSQ and baseline SSQ (i.e. change in baseline) may be analyzed using a repeated measures, mixed effects model, with the SSQ difference as the response and variables such as speech perception, age and gender as the explanatory variables. To determine individual postimplant benefit for the categorization system developed by Noble et al. (2009) will be used (≤1 = no benefit; >1–2 = benefit; >2–4 = high benefit; >4 = very high benefit) with the proportion of subjects falling into each category will be reported.

Mini-Tinnitus Questionnaire (Mini TQ)

The Mini TQ is a 12-item questionnaire that defines a general dimension of distress as a result of tinnitus. The Mini TQ will be recorded for each subject at baseline, at six month and 12 months follow-up. The differences between each follow-up Mini TQ and baseline Mini TQ will be analyzed by reporting the proportion of subjects who reported for each of the statements “TRUE”, PARTLY TRUE”, NOT TRUE. Additionally, the differences between each follow-up Mini TQ and baseline Mini TQ (i.e. change in baseline) may be analyzed using a repeated measure, mixed effects model, with the Mini TQ difference as the response and variables such as speech perception, age and gender as the explanatory variables.

Patient Based Resource and Expenditure Questionnaire (PBRE)

The PBRE questionnaire (assessed at baseline, and 6 Months) will be used to collect self-reported costs of hearing loss pre- and post-cochlear implantation. The differences between the 6 month follow-up PBRE and baseline PBRE (i.e. change in baseline) will be analyzed by reporting the proportion of subjects responding on selected questions.

Montreal Cognitive Assessment -MoCA

The MoCA is a rapid screening instrument for mild cognitive dysfunction. It assesses different cognitive domains: attention and concentration, executive functions, memory, language, visuoconstructional skills, conceptual thinking, calculations, and orientation. The baseline and six month follow-up MoCA score will be recorded for each subject. The proportion of subjects who pass (i.e. score 26 or higher) and those who fail a score of below 26 will be reported. Subjects who demonstrate ≥2or ≤2 change in their score from baseline to 6-month will be reported as significant if it moves them into or out of a category (i.e. pass versus fail) (Personal communication Vincent Lin, MD.)

5.0 Adverse Event Reporting

An Adverse Event is the development of an untoward medical occurrence or the deterioration of a pre-existing medical condition following or during exposure to an investigational product, whether or not considered causally related to the product or the surgical procedure to implant it. An untoward medical condition can be symptoms (e.g., nausea), signs (e.g., tachycardia, fever) or clinically significant abnormal results of an investigation (e.g., laboratory findings, chest x-ray).

Adverse events that occur during this study may be associated with the implant procedure, including those from general anesthesia, or specifically associated with the use of the device. An adverse event will be considered to be device-related when, in the judgment of the Primary Investigator, there is a logical connection between the use of the device and the occurrence of the event, above and beyond the study procedure itself. Adverse events will be counted regardless of severity, seriousness, onset, duration, or relation to study treatment. A serious adverse event (SAE) is any untoward medical occurrence which:

- Results in death;
- Is life-threatening;
- Requires in-patient hospitalization for > 24 hours or prolongation of hospitalization which is not specifically required by the protocol;
- Results in permanent impairment of a body function or permanent damage to a body structure; or
- Requires medical or surgical intervention to preclude permanent impairment of a body function or permanent damage to a body structure.

For any SAE, if the Primary Investigator judges that there is a logical connection (caused or contributed to) between the use of the device and the occurrence, the SAE will be notes as device-related.

All adverse event rates will be reported as the number and frequency of events with corresponding 95% exact binomial confidence limits, as well as the number of events per patient-time (e.g., events per 10 patient years). These values will be qualitatively compared to the same rates observed in previous cochlear implant studies; no formal statistical comparisons will be conducted.

6.0 Justification of Pooling Across Study Sites

Pooling data from study sites will be done based on the following: all sites will have the same protocol, the sponsor will monitor the sites to assure protocol compliance, and the data gathering mechanism (case report forms and data acquisition) will be the same across all study sites (Meinert, 1986).

Consistency of the primary efficacy endpoints between sites will be assessed by testing for a difference between sites in the change in CNC word score from preoperative to 6 months postoperative via an analysis of variance model, with the change in CNC word score as the outcome and site as the factor. A p-value for the site factor of less than 0.10 will be considered evidence of differences between sites for the primary efficacy outcome. If there is evidence of a difference, additional analyses will be performed to explore the possible role of baseline characteristics to explain the results. Results for the primary efficacy endpoint will also be presented separately by site, irrespective of the test of differences between sites to help understand both qualitative and non-significant differences between sites

7.0 Missing Data

All efforts will be put forth to ensure near complete follow-up, with particular focus on the assessment of the primary outcome and occurrence of adverse events. A reminder of subject follow-up due date will be provided to participating centers to facilitate scheduling of the follow-up visit.

In the event a subject is withdrawn prior to the 6-month assessment, the subject will be noted as withdrawn and data collected to the time of withdrawal will be reported. In publication and presentation information and data on the withdrawn subjects will be discussed but not included in the final cohort that reached 6 months.

8.0 Reporting of Results

Results of the analyses outlined in this statistical analysis plan will be reported their entirety in planned publication(s) following the completion of endpoint data collection of all 100 subjects.

9.0 Amendment to Protocol

The original CIP was amended to include two additional study visits.

9.1 Primary Purpose

The primary purpose of these visits is to evaluate whether a change in cochlear implant programming, specifically an increase in the number of channels (maxima) will improve subject speech perception in quiet and noise.

9.2 Primary Effectiveness Endpoint Analysis

Phase 2 CNC Endpoint

Although the stated objective implies a statistical hypothesis test that the paired difference in CNC scores for 16 Maxima MAP compared to 8 Maxima MAP is different from 0, a test of non-inferiority will be stated.

The endpoint will compare CNC scores following 16 maxima MAP compared to the original programming and will be evaluated with a paired t-test. Formally the hypothesis is:

$$H_0: \mu_d \leq -M$$
$$H_a: \mu_d > -M$$

where μ_d is the mean of the paired difference in 16 maxima MAP (new device programming) CNC from 8 maxima MAP (original device programming) CNC and M represents the non-inferiority margin of 10%.

Non-inferiority Design:
All sample size assumptions performed in PASS2020.

Key Sample Size Assumptions:

- Paired t-Test for Non-inferiority
- Mean of paired differences = 0 *(i.e. assuming no difference in average CNC score for both programming)*
- No adjustment for attrition

In Table 1 below, sample sizes are presented under varying assumptions. The first column, "*Non-Inferiority Margin*" represents the magnitude of the margin of the non-inferiority for the mean of paired differences. For this study, higher mean values indicate a "better" response. The value "-10" is the distance below zero for which the new (16 maxima MAP) will still be considered non-inferior to the original MAP. In the FOX study Group 2, MAPs were compared in a group of existing CI users for which the SD was 10%. The estimated standard deviation for the 6 month primary endpoint (Change in CNC at 6 months) sample size calculations was 22.4%. The sample size estimations are presented for both possibilities to provide a range of requirements should the SD be higher than that observed in the FOX study.

Sample size requirements are provided for both 80% and 90% power. Additionally, sample estimates are presented for both a one-sided alpha=0.05 and one-sided alpha = 0.025.

The minimum sample size to maintain 90% power with a standard deviation of 10% is 13 subjects. Other options are provided below in Table 1.

Table 1.
Sample size calculations for Pair T-test of Non-Inferiority

Power	Assumptions		Total Sample Size Requirements	
	Non-Inferiority Margin	Standard Deviation	Minimum Sample Size One-sided Alpha = 0.025	Minimum Sample Size One-sided Alpha = 0.050
80%	-10.0	10.0%	10	8
	-10.0	22.4%	42	33
90%	-10.0	10.0%	**13**	11
	-10.0	22.4%	55	45

9.2 Additional Analyses

- Evaluation of the effect of performance in Quiet and Noise with Wrapping Factor (measure of modiolar proximity)
- Evaluation of the score obtained on the ACE-27 with speech perception outcomes (both original MAP and 16 maxima MAP)
- Observation of TIM measurements and data obtained from a 3D reconstruction of a CT scan.
- Observation of sound quality ratings for the original MAP and 16 maxima in both quiet and noise

10.0 References

Horsman, J., Furlong, W., Feeny, D, & Torrence, G. The Health Utilities Index (HUI) concepts measurement properties and applications. *Health & Quality of Life Outcomes*, 2003 1-54

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

Noble, W., Tyler, R. S., Dunn, C. C., et al. (2009). Younger- and older-age adults with unilateral and bilateral cochlear implants: Speech and spatial hearing self-ratings and performance. *Otol Neurotol*, 30, 921–929.

Thornton, A. R., & Raffin, M. J. (1978). Speech-discrimination scores modeled as a binomial variable. *J Speech Hear Res, 21*(3), 507-518

11.0 Revision History

Version	Change	Date
1	Introduction of document as Supplemental Analysis Plan for CMS approval	April 7, 2017
2	• Inclusion of analysis of individual data per request of CMS • Age group was changed from <70 and > 70 years to <65 and > 65 years to be consistent with Medicare coverage in the United States. • Reference to the IDE was removed as the IDE was	May 5, 2017

	closed in June of 2017 and the study was continued under IRB as a non-significant risk study	
3	• Clarification of intended treat population and the handling of missing data • Exploratory analyses added	October 2, 2018
4	• References added for SSQ and HUI individual analyses	April 8, 2019
5	• Addition of Section 9, to reflect the amendment to the CIP	August 4, 2019
6	• Change to the Hypothesis testing for the protocol amendment and sample size justification	March 5, 2019