Cochlear®

Hear now. And always

CLINICAL INVESTIGATION OF NEW CI DELIVERY MODELS: STATISTICAL ANALYSIS PLAN

MAY 30, 2018

Cochlear Americas
13059 East Peakview Ave., Centennial CO, 80111 USA
Tel: 303 790 9010 Toll Free: 800 523 5798 Fax: 303 792 9025
www.Cochlear.com/US

Table of Contents

1 Purpose

The purpose of this Statistical Analysis Plan (SAP) is to document analyses that are planned for the Clinical Investigation of New CI Delivery Models before they are performed. The results of the analyses documented here are to be presented in the clinical study final report. Selected analyses may also be presented in interim reports.

2 Scope

This SAP should be read in conjunction with the Clinical Study Protocol and Case Report Forms. Changes to the versions of the Clinical Study Protocol (protocol) and Case Report Forms (CRFs) cited above may necessitate updates to the SAP.

3 Applicable Documents

Document Number	Document Title
CAM5417 Protocol	Clinical Investigation of New CI Delivery Models in an Nucleus CI Population Version 2.0

4 Software

Statistical analyses will be conducted in SAS version 9.3 or above (SAS Institute, Cary, N.C.), R version 3.2 or above (R Core Team, http://www.R-project.org) or another validated statistical software package.

5 Abbreviations

The following list provides acronyms used in this document and their meaning:

1. AE – adverse event
2. CI – Cochlear Implant
3. CRF – case report form
4. ITT – intent-to-treat
5. PP – per-protocol
6. SAE – serious adverse event
7. SOC – standard of care

6 Design and Objectives

This is a nonrandomized, multi-site, repeated-measures study design to 1) compare the performance in subjects using an expert clinician compared to an Artificial Intelligence MAP on a CNC Monosyllabic Word score and 2) to compare speech recognition. The primary objective is to evaluate the effectiveness of the AI MAP.

The study will contain a minimum of 80 subjects and a maximum of 190 subjects, 30-70 in Group 1 and 50-120 in Group 2.

7 Study Variables

7.1 Primary Variables

There are two primary effectiveness endpoints that will be evaluated within two different groups. The newly implanted group and the existing subjects.

The first primary effectiveness endpoint is the mean improvement of CNC word recognition in quiet at the 6 month time point in comparison to the preoperative baseline score with an appropriately fit hearing aid in the implanted ear. The mean will be collected from the group of newly implanted subjects.

The second primary effectiveness endpoint is the mean CNC word recognition in quiet at the 1 month time point with AI MAP as compared to EC MAP. The mean scores will be collected from the existing subject population.

7.2 Secondary Variables

Additional variables that will be collected and analyzed include clinician satisfaction, subject satisfaction, various psychoacoustic test metrics, AzBio sentences in noise, and CNC monosyllabic phoneme score at 60 dB. The clinician satisfaction, AzBio sentences, CNC monosyllabic words and psychoacoustic test metrics will be collected for both groups of subjects. Subject satisfaction will be collected on the existing subject group only.

8 Data Collection

Table 1 outlines the visits, visit windows, and testing requirements at each evaluation time point for the newly implanted group (Group 1).

Table 1. Follow-up Schedule and Testing Requirements for Newly implanted Subjects

Event	Pre-Operative	Visit A (Initial Activation	Visit B	Visit C	Visit D
Visit Window	**Before implant**	**2-4 weeks post-surgery**	**1 month post-Visit A**	**3 months post-Visit A**	**6 months post-Visit A**
Create AI automaps		√			
Clinical Counseling		√	√	√	√
CNC Words	√			√	√
AzBio Sentences in noise				√	√
Program with AI Map			√	√	√
Administer Psychoacoustic Tests (via direct connect):					
Audiogram			√		
Phoneme discrimination			√		
Speech audiometry				√	√
Loudness scaling				√	√
Pending outcome requests (if any)				√	√

Table 2 outlines the visits, visit windows, and testing requirements at each evaluation time point for the existing subjects group (Group 2).

Table 2. Follow-up Schedule and Testing Requirements for Existing Subjects

Event	**Visit A**	**Visit B**
Visit Window	**NA**	**1 month post-Visit A**
Create AI automaps	√	
CNC Words in sound	√	√
Clinical Counseling	√	√
AzBio Sentences in noise	√	√
Administer Psychoacoustic Tests (via direct connect):		
Audiometry	√	
Phoneme discrimination	√	
Speech audiometry	√	√
Loudness scaling	√	√
Pending outcome requests (if any)		√
Subject Questionnaire		√

9 Sample Size

Sample size for this trial is based upon separate power calculations for the two primary effectiveness endpoints. The study will have a maximum enrollment of 190 subjects with a maximum of 70 newly implanted subjects and a maximum of 120 existing, or previously implanted, subjects.

For the newly implanted subjects (Group 1), analysis will be based on a paired t-test. For a one-sided α=0.05 with a standard deviation of 24 and a mean difference between preoperative and 6 month performance as small as 12.7%, at least 80% power will be provided for an evaluable sample size of 30 subjects.

For the existing subjects (Group 2), analysis will be based on a paired t-test for non-inferiority. For a one-sided α=0.05 with a standard deviation of 24 and non-inferiority margin of 10% with no assumed mean difference between AI MAP and EC MAP, at least 90% power will be provided for and evaluable sample size of 50 subjects.

10 Analysis Populations

The Intention-to-Treat (ITT) population will be comprised of all the subjects who are enrolled in the study. All primary analysis will be performed on the ITT population per randomization regardless of the treatment actually received.

11 Statistical Analyses

11.1 Primary Effectiveness Endpoint Analysis Group 1

The statistical analysis of the primary endpoint for Group 1 will be based upon a paired t-test. The hypothesis will compare the improvement of the mean CNC score from the preoperative visit to the 6 month visit by paired t-test. The statistical hypothesis is:

$$H_0: p_t = p_c$$
$$H_a: p_t \neq p_c$$

where p_t is the proportion of subjects with success of the primary endpoint in the treatment arm and p_c is the proportion of subjects with success of the primary endpoint in the control arm. The null hypothesis will be rejected at a 0.05 level.

Table X: Primary Effectiveness Endpoint (Group 1)

	Statistics	Screening	6-Month	Difference
CNC Word Recognition Test	Mean ± SD (N) Median (Min, Max)	13.9 ± 15.2 (31) 6.0 (0.0, 58.0)	60.2 ± 17.2 (31) 64.0 (18.0, 85.0)	46.3 ± 22.0 (31) 49.0 (-15.0, 84.0)
	P-value*	N/A	N/A	< 0.001
*P-value for paired difference based on a two-sided, paired t-test with a type I error rate of 5%.				

11.2 Primary Effectiveness Endpoint Analysis Group 2

The statistical analysis of the primary endpoint for Group 2 will be based upon a paired t-test for non-inferiority. The hypothesis will compare the mean CNC score from the baseline visit to the 1 month visit by paired t-test for non-inferiority with a margin on 10%. The statistical hypothesis is:

$$H_0: \mu_V \leq \mu_B - M$$
$$H_a: \mu_V > \mu_B - M$$

where μ_V is the mean CNC score at the 1-month visit and μ_B is the mean CNC score at the baseline visit and M is the non-inferiority margin of 10%. The null hypothesis will be rejected at a 0.05 level.

Table X: Primary Effectiveness Endpoint (Group 2)

	Statistics	Baseline	1-Month	Difference
CNC Word Recognition Test	Mean ± SD (N) Median (Min, Max)	60.1 ± 22.8 (56) 65.0 (3.0, 90.0)	59.1 ± 22.3 (56) 62.5 (6.0, 95.0)	-1.0 ± 10.6 (56) -3.5 (-21.0, 49.0)
	P-value*	N/A	N/A	< 0.001
*P-value for paired difference based on a one-sided, non-inferiority, paired t-test with a margin of 10% and a type I error rate of 5%.				

11.3 Additional Analysis

11.3.1 Analysis of Baseline Characteristics

Table X: Demographic Summary (Group 1)

	Mean ± SD, Median (Min, Max)
Age at enrollment	63.52 ± 16.77, 66 (24, 90)
Duration of hearing loss	28.29 ± 15.92, 26 (3, 56)
	Number and % of subjects
Etiology	
Aging	1 (3.23%)
Genetic	2 (6.45%)
Hereditary	2 (6.45%)
Large Vestibular Aqueduct Syndrome (LVAS)	2 (6.45%)
Meniere's Disease	1 (3.23%)
Meningitis	1 (3.23%)
Noise Exposure	4 (12.90%)
Ototoxic Drugs	1 (3.23%)
Unknown	16 (51.61%)
Ushers 2A	1 (3.23%)
Gender	
Female	15 (48.39%)
Male	16 (51.61%)
Ear Tested	
Left	15 (48.39%)
Right	16 (51.61%)

Table X: Demographic Summary (Group 2)

	Mean ± SD, Median (Min, Max)
Age at enrollment	56.98 ± 22.61, 68 (13, 89)
Duration of hearing loss	28.39 ± 17.92, 25 (1, 74)
Duration of device use (months)	79.30 ± 75.66, 48 (1, 305)
	Number and % of subjects
Etiology	
Cochlear Malformation	1 (1.79%)
Cytomegalovirus (CMV)	1 (1.79%)
Genetic	8 (14.29%)
Hereditary Presbycusis	1 (1.79%)
High Fever	1 (1.79%)
Large Vestibular Aqueduct Syndrome (LVAS)	3 (5.36%)
Maternal Rubella	1 (1.79%)
Measles	1 (1.79%)
Meniere's Disease	4 (7.14%)
Meningitis	3 (5.36%)
Noise Exposure	4 (7.14%)
Otosclerosis	1 (1.79%)
Unknown	26 (46.43%)
Virus	1 (1.79%)
Type of Sound Processor	
Kanso	1 (1.79%)
N6	49 (87.50%)
N7	6 (10.71%)
Gender	
Female	22 (39.29%)
Male	34 (60.71%)
Ear Tested	
Left	29 (51.79%)
Right	27 (48.21%)

11.3.2 Supportive Efficacy Analyses

Additional summary statistics will be presented for the other auditory tests conducted including:

- AzBio Sentence Tests
- Speech audiometry
- Phoneme discrimination
- Loudness scaling
- Audiometry (using FOX Initial and Standard advice)
- Clinician survey

Table X: Supportive Efficacy Analyses (Group 1)

Variable	Statistic	Result
CNC Word Measure: Screening to 6-Month		
Improvement	n/N (%)	30/31 (96.8%)
	95% Exact CI	(83.3%, 99.9%)
No Change	n/N (%)	0/31 (0.0%)
	95% Exact CI	(0.0%, 11.2%)
Decrement	n/N (%)	1/31 (3.2%)
	95% Exact CI	(0.1%, 16.7%)
AzBio Sentence-in-Noise: 3-Month to 6-Month		
Improvement	n/N (%)	8/31 (25.8%)
	95% Exact CI	(11.9%, 44.6%)
No Change	n/N (%)	16/31 (51.6%)
	95% Exact CI	(33.1%, 69.8%)
Decrement	n/N (%)	7/31 (22.6%)
	95% Exact CI	(9.6%, 41.1%)

Table X: Supportive Efficacy Analyses (Group 2)

Variable	Statistic	Result
CNC Word Measure: Baseline to 1-Month		
Improvement	n/N (%)	4/56 (7.1%)
	95% Exact CI	(2.0%, 17.3%)
No Change	n/N (%)	46/56 (82.1%)
	95% Exact CI	(69.6%, 91.1%)
Decrement	n/N (%)	6/56 (10.7%)
	95% Exact CI	(4.0%, 21.9%)
AzBio Sentence-in-Noise: Baseline to 1-Month		
Improvement	n/N (%)	10/56 (17.9%)
	95% Exact CI	(8.9%, 30.4%)
No Change	n/N (%)	33/56 (58.9%)
	95% Exact CI	(45.0%, 71.9%)
Decrement	n/N (%)	13/56 (23.2%)
	95% Exact CI	(13.0%, 36.4%)

Table X: AzBio Sentences in Noise (Group 1)

	Statistics	3-Month	6-Month	Difference
AzBio Sentence Test	Mean ± SD (N) Median (Min, Max)	39.4 ± 23.0 (31) 33.1 (3.6, 84.8)	42.7 ± 21.5 (31) 38.7 (3.6, 76.6)	3.3 ± 16.7 (31) 0.4 (-21.7, 44.2)
	P-value*	N/A	N/A	0.285
*P-value for paired difference based on a two-sided, paired t-test with a type I error rate of 5%.				

Table X: AzBio Sentences in Noise (Group 2)

	Statistics	Baseline	1-Month	Difference
AzBio Sentence Test	Mean ± SD (N) Median (Min, Max)	46.7 ± 25.7 (56) 45.1 (0.7, 96.5)	46.9 ± 27.4 (56) 47.0 (0.0, 95.7)	0.2 ± 15.8 (56) -1.4 (-30.8, 61.5)
	P-value*	N/A	N/A	0.907
*P-value for paired difference based on a two-sided, paired t-test with a type I error rate of 5%.				

Table X: CNC Monosyllabic Phoneme Score (Group 1)

	Statistics	Screening	6-Month	Difference
CNC Phonemes	Mean ± SD (N) Median (Min, Max)	27.9 ± 23.6 (31) 18.7 (0.0, 81.7)	77.4 ± 12.8 (31) 80.3 (39.7, 93.7)	49.5 ± 26.8 (31) 57.7 (-21.3, 91.0)
	P-value*	N/A	N/A	< 0.001
*P-value for paired difference based on a two-sided, paired t-test with a type I error rate of 5%.				

Table X: CNC Monosyllabic Phoneme Score (Group 2)

	Statistics	Baseline	1-Month	Difference
CNC Phonemes	Mean ± SD (N) Median (Min, Max)	76.6 ± 17.4 (56) 82.2 (25.3, 96.0)	75.3 ± 17.5 (56) 81.5 (27.7, 97.7)	-1.3 ± 9.3 (56) -1.7 (-22.3, 47.7)
	P-value*	N/A	N/A	0.297
*P-value for paired difference based on a two-sided, paired t-test with a type I error rate of 5%.				

Table X: Primary Effectiveness Endpoint by Visit (Group 1)

	Visit	Statistics	Result at Visit	Change from Screening
CNC Word Recognition Test	Screening	Mean ± SD (N) Median (Min, Max)	13.9 ± 15.2 (31) 6.0 (0.0, 58.0)	N/A
	3 month	Mean ± SD (N) Median (Min, Max)	57.3 ± 17.7 (31) 59.0 (16.0, 86.0)	43.4 ± 24.1 (31) 43.0 (-17.0, 83.0)
		P-value*	N/A	< 0.001
	6 month	Mean ± SD (N) Median (Min, Max)	60.2 ± 17.2 (31) 64.0 (18.0, 85.0)	46.3 ± 22.0 (31) 49.0 (-15.0, 84.0)
		P-value*	N/A	< 0.001
*P-value for paired difference based on a two-sided, paired t-test with a type I error rate of 5%.				

12 General Statistical Considerations

12.1 Descriptive Statistics

Standard summary statistics will be calculated for all study variables to be reported. Continuous data will be summarized by statistics including means, standard deviation or error, median, minimum, maximum, and n. Categorical variables will be summarized by statistics including frequency counts and percentages.

For ordinal-scaled variables a combination may be applied where appropriate. Both categorical summary statistics (frequency and percentage) may be presented within a category and continuous summary statistics (means and standard deviations) may be presented of the scores in the category.

For categorical and ordinal variables, percentages will be calculated on non-missing data.

12.2 Protocol deviations

The number and type of protocol deviations will be summarized with descriptive statistics.

12.3 P-values

Unless otherwise specified statistical analyses will be performed using a two-sided hypothesis test at the overall 5% level of significance. P-values will be rounded to three decimal places. If a p-value is less than 0.001 it will be reported as < 0.001.

13 Interim Analyses

An interim report will be created to show descriptive statistics of group outcomes and report safety events. No formal statistical testing will be conducted.

14 Missing Data Analyses

All efforts will be made to complete follow-ups. All endpoints will be analyzed based on available data and the percentage of missing data will be taken into consideration.

15 Poolability of Data

This study is a multicenter clinical trial and all participating sites will be selected based upon the same criteria and trained according to the protocol. This will ensure generalizability of the study results.

Table X: Primary Effectiveness Endpoint by Site
6-Month Change from Screening in CNC Word Recognition Test (Group 1)

Site	Mean ± SD (N)	Median (Min, Max)	P-value[1] (t-test)	P-value[2] (ANOVA)
0029	38.0 ± 2.8 (2)	38.0 (36.0, 40.0)	0.033	0.591
1003	42.2 ± 34.9 (5)	59.0 (-15.0, 72.0)	0.054	
1050	44.3 ± 19.7 (3)	52.0 (22.0, 59.0)	0.060	
1065	61.1 ± 23.5 (7)	68.0 (20.0, 84.0)	< 0.001	
1507	36.3 ± 20.2 (6)	30.0 (14.0, 69.0)	0.007	
1523	46.6 ± 13.3 (5)	42.0 (30.0, 63.0)	0.001	
Sites with one subject (0848, 1141, 2180)	45.7 ± 14.3 (3)	49.0 (30.0, 58.0)	0.031	

[1]P-value for paired difference based on a two-sided, paired t-test with a type I error rate of 5%.
[2]P-value for ANOVA based on a one-way analysis of variance with 6-month change in CNC as the dependent variable and site ID as the independent variable. Sites with one subject were pooled together into one site.

Table X: Primary Effectiveness Endpoint by Site
1-Month Change from Baseline in CNC Word Recognition Test (Group 2)

Site	Mean ± SD (N)	Median (Min, Max)	P-value[1] (t-test)	P-value[2] (ANOVA)
0029	-4.0 ± 8.4 (9)	-4.0 (-21.0, 10.0)	0.189	0.487
0848	0.2 ± 6.8 (5)	1.0 (-10.0, 8.0)	0.951	
1003	3.0 ± 10.2 (9)	6.0 (-12.0, 18.0)	0.403	
1050	-6.4 ± 5.5 (7)	-6.0 (-18.0, -2.0)	0.021	
1065	0.6 ± 8.4 (9)	4.0 (-13.0, 12.0)	0.848	
1141	-4.0 ± 8.5 (2)	-4.0 (-10.0, 2.0)	0.626	
1507	3.9 ± 19.5 (8)	-1.5 (-14.0, 49.0)	0.591	
1523	-4.3 ± 8.1 (7)	-6.0 (-15.0, 6.0)	0.212	

[1]P-value for paired difference based on a two-sided, paired t-test with a type I error rate of 5%.
[2]P-value for ANOVA based on a one-way analysis of variance with 1-month change in CNC as the dependent variable and site ID as the independent variable.

16 Deviations from the Statistical Analysis Plan

Any deviation from the statistical methods planned, as specified above, will be documented and discussed in the final study report. Any rationale for such deviation will be presented in the final study report.