bernafon
Your hearing · Our passion

Statistical Analysis Plan

A comparative, controlled clinical investigation of a new acoustic feedback cancellation strategy in comparison with the currently marketed system.

Study Type:	Clinical trial with Investigational Medical Device (MD)
Study Categorisation:	Category C; MD without CE mark
Study Registration:	SNCTP, EudraCT
Study Identifier:	BF004-1901
Sponsor, Sponsor	Bernafon AG Morgenstrasse 131, 3018 Bern
Investigational Product:	Hearing Instrument; Mermaid 9 S
Protocol Version and Date:	Version 2.0, Final Document

Table of Contents

1. ADMINISTRATIVE INFORMATION 3
1.1 Signatures 3
1.2 Roles and Responsibilities 4
2. INTRODUCTION 4
2.1 Background and Rationale 4
2.2 Purpose of statistical analysis plan 4
3. TRIAL OBJECTIVES AND OUTCOME MEASURES 4
3.1 Primary objectives 4
3.2 Primary Outcome Measure 5
3.3 Secondary Objectives 5
3.4 Secondary Outcome Measures 5
4. STUDY METHODS 6
4.1 Background and Rationale 6
4.2 Randomization and Stratification 6
4.3 Sample Size 6
4.4 Timing of interim and final analysis 8
5. STATISTICAL PRINCIPLES 8
5.1 Levels of confidence intervals and p-values 8
5.2 Analysis populations 8
6. ANALYSIS 8
6.1 Analysis methods 8
6.2 Missing data 10
6.3 Harms / Safety Data 10
6.4 Statistical software 10
7. REFERENCE 10

bernafon®
Your hearing · Our passion

1. ADMINISTRATIVE INFORMATION

1.1 Signatures

Study number	SNCTP; EudraCT, registration number (TBD)
Study Title	A comparative, controlled clinical investigation of a new acoustic feedback cancellation strategy in comparison with the currently marketed system.
Protocol Version	CIP v2.0

Printed name of Trial Statistician: Christophe Lesimple

Bern 2018.09.26 _____

Place/Date Signature

Printed name of Principle Investigator: Barbara Simon

Bern 2018.09.26 _____

Place/Date Signature

1.2 Roles and Responsibilities

Sponsor, Sponsor-Investigator

Bernafon AG, Morgenstrasse 131, 3018 Bern, Tel. +41 31 998 01 01

The role of the sponsor is to provide the site for the testing as well as the equipment used during testing. The sponsor will provide the hearing devices, the IMD, and the RMD used for the study. The results will be used by the sponsor to prove the performance of the IMD. The sponsor may audit the clinic as well as the processes and documentation performed by the investigators at that site.

Principal Investigator(s)

Barbara Simon, Research Audiologist, Morgenstrasse 131, 3018 Bern, Tel. +41 31 998 16 46

Email: bsim@bernafon.com

Statistician ("Biostatistician")

Christophe Lesimple, Morgenstrasse 131, 3018 Bern, Tel. +41 31 998 17 03

Email: cles@bernafon.com

Monitoring institution

Bernafon uses monitoring to oversee the study and verify that the conduct of the clinical investigation complies with the approved CIP, subsequent amendment(s), ISO14155, and the applicable regulatory requirement(s). There will be a specific person assigned as the Monitor (sec. 1.7).

Julie Tantau will monitor the investigation. She works within the Product Validation group at Bernafon. She is certified in GCP, and familiar with ISO 14155. She has also been certified in Clinical Monitoring and has a CAS I in Clinical Trial Practice and Management.

2. INTRODUCTION

2.1 Background and Rationale

The full background and rationale for the present trial is detailed in the study protocol [1]. In summary, Bernafon AG will carry out testing with participants who have hearing loss to validate the performance of the new feedback algorithm embedded in a hearing aid. The purpose of this study is to show that the performance of the new feedback cancellation system is better than the feedback system used in the currently CE marked devices. Speech understanding should not be negatively affected by the new system and there should be no consequential artefacts or unwanted noises caused by the new system.

2.2 Purpose of statistical analysis plan

The study protocol includes an outline of the statistical methods to be employed in the analysis of the trial data. The purpose of the Statistical Analysis Plan is to provide full details of the planned statistical methods to be used in the primary report of the trial results. It has been produced in line with the ICH E9 Guidelines [2] and guidelines from Gamble et al. (2017) [3].

3. TRIAL OBJECTIVES AND OUTCOME MEASURES

The results of the trial will be used to examine differences in benefit provided by the new feedback cancellation system in the IMD compared to that in the RMD, as well as identify further optimization of the tested products. In summary, the primary reason for this study is to evaluate the new hearing aid feedback cancellation algorithm. The secondary reason is to evaluate the audiological performance and safety of the new hearing aids before they're released to the market.

3.1 Primary objectives

The study seeks to assess the performance of the new feedback cancellation system in comparison to the current feedback system. Scores from a live feedback test of the IMD will be compared to those of the RMD.

3.2 Primary Outcome Measure

The primary outcome variable will be measured with a live feedback test that subjectively measures annoyance of feedback that the participants experience in simulated real-life situations that typically elicit feedback. It will be measured in two conditions: aided with the RMD and aided with the IMD.

3.3 Secondary Objectives

Secondary objectives are to assess the performance of the IMD using speech testing and questionnaires. The performance of the IMD should not be inferior to that of the RMD.

3.4 Secondary Outcome Measures

The secondary outcome will be measured with a standardized speech test and with questionnaires. The speech test will be measured in three conditions: unaided, aided with the RMD, and aided with the IMD.

The product questionnaire will be answered for both the RMD and the IMD, and then the answers compared. The questionnaire will focus on the performance of the device with specific questions regarding feedback and sound quality.

The Speech, Spatial, and Qualities comparative (SSQ-C) questionnaire is a standardized questionnaire of overall performance and directly compares the RMD to the IMD.

4. STUDY METHODS

4.1 Background and Rationale

For the current study, three behind-the-ear (BTE) styles with the new chip are the IMD, and the BTE styles that are currently sold on the market will be the RMD and comparator. All participants are hearing impaired persons and will be fit with the RMD and the IMD during the trial. Selection of the BTE style is made based on the degree of hearing loss. To make an effective comparison the test participants will wear sequentially the RMDs and the IMDs for approximately 10 days.

Feedback performances will be evaluated in a controlled environment for the live feedback test and in daily life situations during the wearing time. This approach should cover the range of use cases to have a thorough evaluation of the new feedback canceller.

4.2 Randomization and Stratification

The study is based on repeated measures with the RMD and with the IMD. Experienced acoustical feedback is an acoustical artefact that is not influenced by perceptual processes. Earmold misplacement and high amplification have been identified as factors that could trigger acoustical feedback. Therefore, it is not expected to have any learning or period effect for experienced hearing aid users. For each IMD style, eligible participants will test first the equivalent RMD and then the IMD.

The selection of the appropriate IMD device is based on the hearing loss configuration in three phases: first and second phases for mild to severe hearing loss and third phase for moderate to profound hearing loss. The first and second phases will test 2 IMD styles, i.e. miniRITE-T and miniRITE-T rechargeable, and the third phase, one style, the super power BTE. Number of participant ratio between the first two phases and the third phase is 2:1 because:

- Audiometrical configurations from phase one and two cover a wider range of hearing loss degree and acoustical coupling which has the potential to influence acoustical feedback,
- The feedback risk might be higher and more homogeneous in the third phase due to a strong dependency to amplification. This factor is directly reflected by the degree of hearing loss.

Randomization is only foreseen for selecting the list number of the speech test. It will be in consideration of the listening condition and the test condition.

4.3 Sample Size

The primary outcome, the live feedback test, measures the feedback annoyance on a visual analogue scale for different manipulations that could produce a feedback. An internal pilot test [4] found an average improvement of 1.54 points (SD 1.63 points) with the new developed feedback canceller (prototype) on the feedback annoyance scale. This effect was obtained on a group of listeners with severe hearing loss which should be close to the hearing loss of subjects recruited for the third phase of the study.

Observed changes in feedback annoyance were distributed only in favour to the new developed system regardless of the programmed gain. This observation makes sense regarding the technical implementation and verification of the new feedback canceler implemented in the IMD, i.e. as an improvement based on the RMD. As we test for superiority of the IMD in terms of feedback annoyance, we will use a one-sided hypothesis test.

The analysis will be done independently for each IMD style so that each implementation of the new feedback canceler is evaluated separately. Subjects allocated in the first phase can participate to the trial with the miniRITE-T rechargeable in the second phase.

The target between devices to be detected is 1.54 (SD 1.63) which is plausible and considered to be clinically relevant. A sample size of 11 participants is required to test the implementation of the feedback canceller in the third phase with a 2.5% significance level, 80% power, and a one-sided hypothesis (Chow, 2013 p.451 Table 11.3.1) [5].

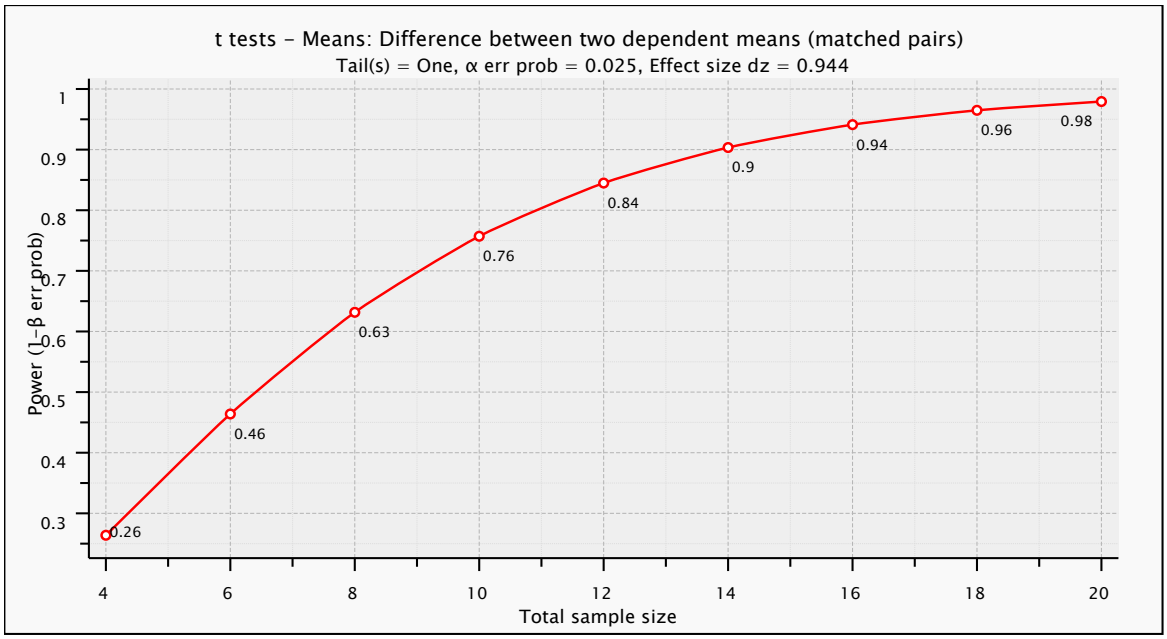

Figure 1: achieved power as a function of sample size for an effect size of 0.94 and the comparison of 2 dependant means in a one-sided test. This effect size is computed for the 3rd stage for subjects with a moderate to profound hearing loss (i.e. more gain and more feedback risk).

To respect the ratio between phases, 22 participants in the first two phases and 11 participants in the third phase must be recruited. As the feedback detection might be less consistent with the group from the first and second phases, we expect a smaller effect size in differences between IMD and RMD:

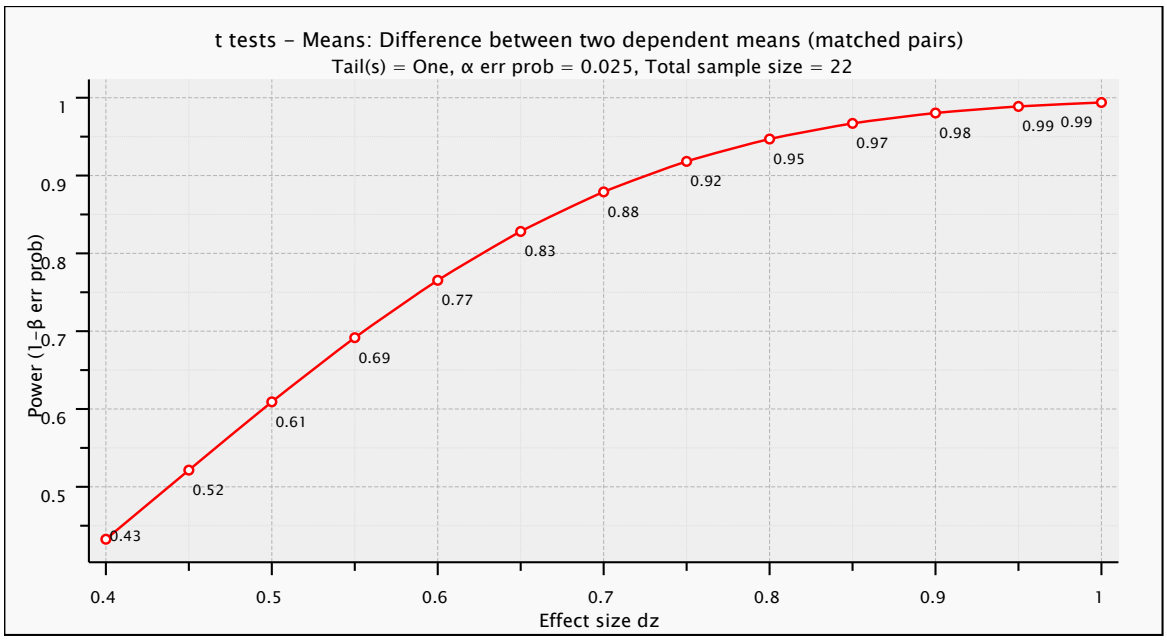

Figure 2: achieved power as a function of the effect size for a sample size of 22 subjects.

Test power is kept at 83% with an effect size reduction up to 0.65. There is no prior hypothesis about the magnitude of the change in effect size. The distribution of hearing loss degree for participants in the first and second phases is broader. Therefore, participants from these phases might not be equally affected by feedback:

- feedback risk might increase, leading to more feedback annoyance, with participants with a stronger hearing loss because the device is programmed with more gain,
- this risk might be partially compensated with a closer acoustical coupling which reduces the feedback risk during wearing time. However, for device insertion and removal, feedback

should be mainly dependent on the programmed gain, directly linked to the degree of hearing loss.

4.4 Timing of interim and final analysis

We plan to have 3 interim and one final analysis. A separate analysis will be performed after completed comparison performed between the RMD and each tested IMD style. These interim analyses are based on one-sided superiority hypothesis, i.e. there is less feedback annoyance with the IMD than with the RMD. However, there are no plans for early termination of the trial.

Valuable information may also be gained by summarising the results of all the conducted comparisons at the end of the study. It should be presented in an identical form to allow direct comparison on the estimates and their confidence limits. Meta-analytic techniques such as effect sizes will be used in the final analysis to provide overall evidence of efficacy via an overall hypothesis test.

5. STATISTICAL PRINCIPLES

5.1 Levels of confidence intervals and p-values

The statistical test for the feedback performance objective will be performed at one-sided $\alpha=0.025$. The statistical test for the speech test will be performed at two-sided $\alpha=0.05$. All reported p-values greater than or equal to 0.001 will be rounded to three decimal places and p-values less than 0.001 will be displayed as "<0.001".

5.2 Analysis populations

Primary analysis of the primary outcome will be based on the intention-to-treat (ITT). The ITT population will include all participants with associated primary outcome data, excluding only subjects who were deemed ineligible following screening visit, those who withdrew from the trial and were unwilling for their previously collected data to be utilised or those who failed to provide baseline.

The following demographic and baseline characteristics will be tabulated overall for the ITT and per protocol population: Age (years), Gender (categorical variable), Hearing Loss Degree (from/to categories), 4-frequency Pure Tone Average (in dB HL), Acoustical Coupling with the RMD/IMD (categorical variable), and history, i.e. hearing loss onset, ear surgery, otalgia, otorrhea, otitis, tinnitus, and noise exposure.

For continuous variables (e.g. age) descriptive statistics will be presented (mean, standard deviation, median, minimum, maximum, interquartile range and number of participants with data). For categorical variables (e.g. gender) frequencies, percentages and number of participants with data will be presented. The denominator for the percentages will be the number of patients with non-missing data.

6. ANALYSIS

6.1 Analysis methods

6.1.1 Live Feedback Test

The live feedback test provides a measure of feedback annoyance on a visual analogue scale during different manipulations that are known, from clinical experience, being able to trigger acoustical feedback, e.g. hearing aid insertion, cover the ear with the hand or a phone, and remove the hearing aid. The score can range from 0 (no feedback) to 10 (extremely annoying feedback).

In a first intention, a paired t-test will be used to evaluate if the observed difference (RMD - IMD) is higher than 0. However, following factors are identified and should be included within a second analysis based on mixed-effect regression:

- Test device RMD vs IMD which is the main fixed effect,
- Participant as a random effect for the repeated measures, and the side (left-right) as nested random factor within participant,
- Feedback risk with the acoustical coupling, the hearing loss degree, and maximum gain measured from the REM,
- Manipulation as fixed effect.

Data distribution will be evaluated prior the analysis. The possibility to apply a +1 and log 10 transformation should be evaluated because there is the possibility that the distribution is right skewed. The strategy to avoid collinearities is to compute the variation inflation factor (VIF) for each predictor and then drop the factor with the highest VIF above 3 [6].

6.1.2 Speech Test

Speech test is measured with a word recognition test in German [7]. The WAKO test is a closed word recognition test, i.e. one word is presented acoustically, and the subject must choose one answer out of five proposals. There are 10 lists and each list has 47 tested words.

There are three test conditions: unaided, aided with the IMD, and aided with the RMD. Each test condition produces 3 scores in different listening conditions, i.e. a specific signal-to-noise ratio. The listening condition represent the listening difficulty: in quiet for the easy listening condition, at +10 dB SNR for the moderate listening condition, and at 0 dB SNR for the difficult listening condition.

The result of one list can be expressed in percentage of word recognition. While this approach seems to be easily understandable, there are limitations to use percentages with confidence intervals and inference test [8]. The outcome of each tested item is either 0 and 1 which can be modelled by a logistic mixed effect regression. Fixed effects are test condition, listening condition, hearing loss, i.e. 4-frequencies average of the best ear. Random effects are the participants and the tested item. The analysis will be made with planned contrasts on the test condition:

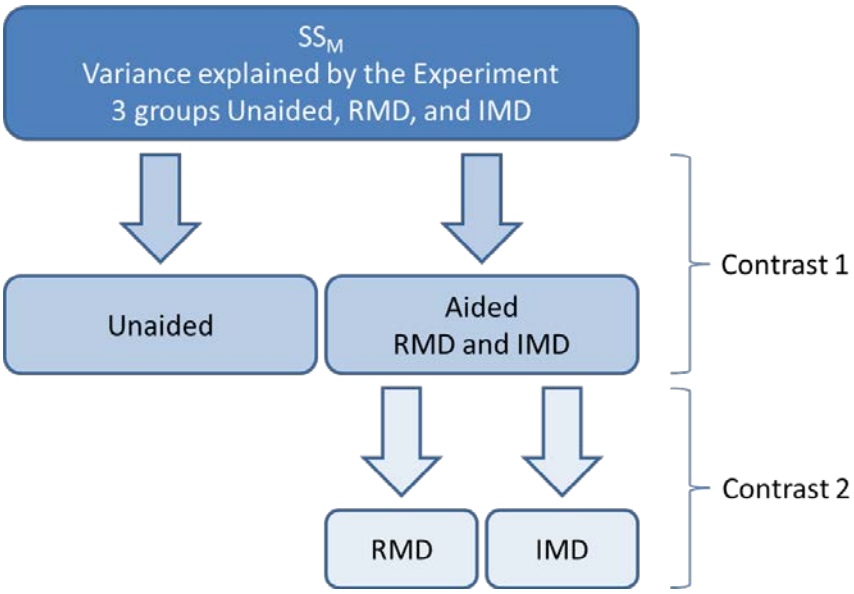

Figure 3: planned contrast for the speech test analysis. The first contrast is evaluating the benefit of amplification while the second contrast is focusing on the differences between devices.

The motivation for planned contrasts is given by the clinical differences between test conditions, both aided conditions are closer to the third, unaided, condition. A benefit with amplification in general can be reasonably expected [7]. The first contrast will test if there is an overall benefit with amplification and the second will specifically look at the differences of aided word recognition performances.

6.1.3 Product Questionnaire

Reported acoustical feedback: it is asked how often they experience acoustical feedback. Answer possibilities range from never to always on a 5-points Likert scale.

Reported sound quality: rating of the overall sound quality is reported on a 5-points Likert scale, from excellent to very bad.

Results from both questions will be analysed with Wilcoxon signed-rank test to evaluate if the mean rank differ between conditions.

Experienced artefacts, i.e. unexpected sounds or noises, must be reported and described. Each experience will be listed for each test condition.

6.1.4 Speech, Spatial, and Qualities (SSQ-C) comparative questionnaire

Twelve questions assess the subjects' performance with the hearing aids in real-life situations. The SSQ-C asks the respondent to compare the ability or experience with a current hearing aid to the ability or experience with a previous hearing aid. It uses a scale from −5 to 5, where −5 indicates that *performance was much better with the first hearing aid*, 0 indicates *no difference*, and 5 indicates that *performance was much better with the second hearing aid*.

Three subscales will be analysed separately in a first intention. Each one has 4 questions and they focus on different aspects of hearing: Speech (question 1 to 4), Spatial (question 5 to 8), and Qualities (question 9 to 12). The mean of all the tested item can also be used to get the overall performance with both devices. A one sample t-test is appropriate to detect if the average is significantly different from 0.

6.2 Missing data

Unless specified otherwise in each objective, no statistical techniques will be used to impute missing data. If a subject's data are missing for any reason, that subject will not be included in that portion of the analysis. The number of subjects included in each analysis will be reported so that the potential impact of missing data can be assessed.

6.3 Harms / Safety Data

The adverse event risks of taking part in the study have been assessed to be low [1]. Numbers of adverse events and serious adverse events will be cross-tabulated for each IMD style, and categorised by severity. No formal statistical analysis will be conducted, but AEs and SAEs will be closely monitored throughout the process.

Serious Adverse Events (SAEs) will be recorded and the site will notify the trial monitor of any SAE, who will then notify the project sponsor within one working day.

6.4 Statistical software

Data manipulation, statistical summaries and statistical analyses will be performed using R version 3.5.0 or higher [9].

7. REFERENCE

1. CIP_2018.07.06_V1, A comparative, controlled clinical investigation of a new feedback cancellation strategy in comparison with the currently marketed system, 2018.
2. ICH Harmonised Tripartite Guideline: Statistical Principles for Clinical Trials E9. 1998
3. Gamble, C., Krishan, A., Stocken, D., Lewis, S., Juszczak, E., Doré, C., … Loder, E. (2017). Guidelines for the content of statistical analysis plans in clinical trials. JAMA - Journal of the American Medical Association, 318(23), 2337–2343.
4. Internal concept report: Report for the lab test of Bee on HA users, CC1090, Andreea Micula, 2017.
5. Chow, S.-C., & Liu, J.-P. (2013). Sample Size Determination. In Design and Analysis of Clinical Trials: Concepts and Methodologies (pp. 441–518).
6. Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. Methods Ecol. Evol. 1 (1), 3–14.
7. Kollmeier B, Lenarz T, Winkler A, Zokoll MA, Sukowski H, Brand T, Wagener KC. [Indication for and verification of hearing aid benefit using modern methods of speech audiometry in German]. HNO. 2011 Oct;59(10):1012-21.
8. Jaeger TF. Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. J Mem Lang. 2008 Nov;59(4):434-446.
9. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.